



## Mapping two *Eucalyptus* subgenera using multiple endmember spectral mixture analysis and continuum-removed imaging spectrometry data

Kara N. Youngentob<sup>a,\*</sup>, Dar A. Roberts<sup>b</sup>, Alex A. Held<sup>c</sup>, Philip E. Dennison<sup>d</sup>,  
Xiuping Jia<sup>e</sup>, David B. Lindenmayer<sup>a</sup>

<sup>a</sup> The Fenner School of Environment and Society, The Australian National University, Canberra ACT 0200, Australia

<sup>b</sup> Department of Geography, University of California, Santa Barbara, CA 93106, United States

<sup>c</sup> Commonwealth Scientific and Industrial Research Organization, Division of Marine and Atmospheric Research, Canberra, ACT 2602, Australia

<sup>d</sup> Department of Geography, University of Utah, Salt Lake City, UT 84112, United States

<sup>e</sup> Australian Defense Force Academy, School of Electrical Engineering, Canberra, ACT 2600, Australia

### ARTICLE INFO

#### Article history:

Received 17 December 2009

Received in revised form 14 December 2010

Accepted 14 December 2010

#### Keywords:

MESMA

Classification

Continuum-removal

Ecology

Forest

Hyperspectral

Habitat mapping

Imaging spectroscopy

*Eucalyptus*

Plant functional-type

### ABSTRACT

Successful discrimination of a variety of natural and urban landscape components has been achieved with remote sensing data using multiple endmember spectral mixture analysis (MESMA). MESMA is a spectral matching algorithm that addresses spectral variability by allowing multiple reference spectra (i.e., endmembers) to represent each material class. However, materials that have a high-degree of spectral similarity between classes, such as similar plant-types or closely related plant species, and large variations in albedo present an ongoing challenge for accurate class discrimination with imaging spectrometry. Continuum removal (CR) analysis may improve class separability by emphasizing individual absorption features across a normalized spectrum. The spectral and structural characteristics common to most *Eucalyptus* trees make them notoriously difficult to discriminate in closed-canopy forests with imaging spectrometry. We evaluated whether CR applied to hyperspectral remote sensing data improved the performance of MESMA in classifying and mapping nine eucalypt tree species according to the two major *Eucalyptus* subgenera, *Eucalyptus* (common name “monocalypt”) and *Symphyomyrtus* (common name “symphyomyrtle”). Mixed-canopies comprised of monocalypts and symphyomyrtles are common in Australia, although their spatial distribution is not random. The ability to map these functional types on a landscape-scale could provide important information about ecosystem processes, landscape disturbance history and wildlife habitat. We created a spectral library of 229 pixels from 37 symphyomyrtle tree canopies and 406 pixels from 62 monocalypt tree canopies selected from HyMap imagery and verified with field data. Based on these reference data, we achieved overall classification accuracies at the subgenera-level of 75% (Kappa 0.48) for non-CR spectra and 83% (Kappa 0.63) for the CR spectra. We found that continuum-removal improved the classification performance of most endmember-models, although a larger portion of pixels remained unmodeled with the CR spectra (2%) compared to the non-CR spectra (0%). We utilized a new method for model optimization and created maps of monocalypt and symphyomyrtle distribution in our study area based on our best performing endmember-models. Our vegetation maps were largely consistent with our expectations of subgenera distribution based on our knowledge of the region.

© 2011 Elsevier Inc. All rights reserved.

### 1. Introduction

Our ability to understand, monitor and conserve native landscapes is limited by our lack of knowledge about the composition, structure and disturbance history of ecological communities. Recent advances in airborne and spaceborne hyperspectral remote sensing combined with improved algorithms for spectral discrimination are allowing researchers to map vegetation communities with increasing accuracy

(for reviews see, [Majeke et al., 2008](#); [Turner et al., 2003](#); [Xie et al., 2008](#)). However, vegetation types that have a high degree of spectral and structural similarity or high intra-species variability present an ongoing challenge for accurate classification and mapping ([Clark et al., 2005](#); [Cochrane, 2000](#); [Goodwin et al., 2005](#); [Hestir et al., 2008](#)). Variations in landscape topography, canopy structure and viewing geometry also can reduce the accuracies of classification techniques and spectral matching algorithms that are influenced by albedo ([Asner et al., 2000](#); [Dennison et al., 2004](#); [Wu, 2004](#)).

In this study, we investigated whether multiple endmember spectral mixture analysis (MESMA, [Roberts et al., 1998](#)) could be used to discriminate *Eucalyptus* trees from the two major subgenera, *Eucalyptus*

\* Corresponding author. Tel.: +61 2 6125 1035; fax: +61 2 6125 0757.

E-mail address: [kara.youngentob@anu.edu.au](mailto:kara.youngentob@anu.edu.au) (K.N. Youngentob).

(common name “monocalypt”) and *Symphyomyrtus* (common name “symphyomyrtle”), using hyperspectral remote sensing data collected over a forest in southeastern Australia. These two subgenera encompass the majority of eucalypt tree species and represent an important functional type based on general divisions among the physiology (e.g. concentrations of foliar biochemicals) and ecology (e.g. response to fire and salinity) of trees from these two groups (Noble, 1989). Despite these differences, most species within these two subgenera still possess the unfavorable structural characteristics and spectral variability that can frustrate attempts at discrimination on a canopy scale using imaging spectroscopy (Turner et al., 1998). The MESMA approach incorporates spectral variability within material classes, and therefore may be useful for this application.

We further investigated whether a spectral normalization technique, continuum removal (CR), could improve the performance of MESMA, given the variability in reflectance arising from the open-canopy architecture common to eucalypt trees.

We hypothesized that CR-spectra were likely to improve class discrimination by reducing brightness differences and accentuating subtle absorption features. We also tested the hypothesis that those species that were less confidently modeled had particular physiognomic characteristics common to both subgenera. Using our most successful models, we generated maps of the distribution of monocalypt and symphyomyrtle trees in our study area. This research is part of a larger study investigating the use of hyperspectral remote sensing to map habitat for arboreal marsupial folivores.

## 2. Background

### 2.1. Spectral mixture analysis

Spectral mixture analysis (SMA) is an analysis technique that models a mixed spectrum as the combination of two or more “pure” spectra, called endmembers (Adams et al., 1986; Gillespie, 1992; Roberts et al., 1993; Settle & Drake, 1993). Endmembers can be derived from the laboratory, field (Roberts et al., 1993), imagery (Dennison & Roberts, 2003a) or even radiative transfer (Painter et al., 1998; Sonnentag et al., 2007). Although multiple near-infrared (NIR) scattering results in non-linear mixing (e.g., Borel & Gerstl, 1994; Roberts et al., 1993; Somers et al., 2009a,b), most often a linear model is assumed. Typical endmembers include one to three bright spectra paired with a dark “shade” or photometric shade endmember to control for brightness. Using a linear mixing model, a mixed spectrum is modeled as the sum of the reflectance of each material within a pixel multiplied by its spectral fraction (Eq. 1):

$$\rho'_\lambda = \sum_{i=1}^N f_i \rho_{i\lambda} + \varepsilon_\lambda \quad (1)$$

where  $\rho'_\lambda$  is the reflectance of a modeled spectrum,  $\rho_{i\lambda}$  is the reflectance for endmember  $i$ ,  $f_i$  is the fraction contributed by the endmember,  $N$  is the number of endmembers, and  $\varepsilon_\lambda$  is the residual term—for a specific wavelength  $\lambda$ . Fractional abundance can be estimated using a variety of approaches, including least squares (Shimabukuro & Smith, 1991), modified Gram–Schmidt orthogonal decomposition (Adams et al., 1993) or singular value decomposition (Boardman et al., 1995) as three common approaches. Model fit is often assessed using a root mean square error (RMSE) error metric:

$$\text{RMSE} = \sqrt{\frac{\sum_{\lambda=1}^M (\varepsilon_\lambda)^2}{M}} \quad (2)$$

where  $M$  is the number of bands (Dennison et al., 2004).

Most often, SMA is implemented using a fixed-set of endmembers applied to an entire scene. While this approach has proven to be effective for mapping fractional cover, it fails to account for within-class

spectral variability or spatial variability in the spectral dimensionality of the data (Roberts et al., 1998; Sabol et al., 1992; Song, 2005). In an attempt to overcome this limitation, various techniques to address endmember variability and similarity have been developed (e.g., Asner & Lobell, 2000; Bateson et al., 2000; Roberts et al., 1998; Somers et al., 2009a,b). For example, Bateson et al. (2000) incorporated endmember variability into SMA by representing each endmember as a bundle of spectra constructed from the data. Somers et al. (2009a,b) presented an alternative to SMA, Integrated Spectral Unmixing (InSU), that combined reflectance and derivative reflectance features using an automated waveband selection protocol. This method reduced the variability within endmember classes by focusing on a subset of wavelengths. Similarly, Asner and Lobell (2000) used a spectral unmixing algorithm in combination with waveband selection and a normalization technique to reduce endmember variability.

MESMA is an extension of SMA that addresses spectral and spatial variability within material classes by allowing the number and type of endmembers to vary on a per pixel basis (Roberts et al., 1998). Rather than using waveband selection or spectral transformation techniques to reduce endmember variability, MESMA enables the user to select multiple endmembers to represent each material class. Spectral matching can be accomplished with the two-endmember case of MESMA, which is comprised of one class endmember coupled with a shade endmember (Dennison et al., 2007; Roberts et al., 1998). The MESMA approach has been widely used for mapping minerals (Bedini et al., 2009; Li & Mustard, 2003), snow cover and grain size (Painter et al., 1998), fire properties (Dennison et al., 2006; Eckmann et al., 2008), continental-scale land-cover type (Ballantine et al., 2005), urban environments (Powell et al., 2007; Rashed, 2008) and vegetation type and biophysical properties (Dennison & Roberts, 2003a; Roberts et al., 1998; Sonnentag et al., 2007). Recently, techniques to improve MESMA endmember selection (Dennison et al., 2004; Dennison & Roberts, 2003b; Dennison et al., 2007) and an open-source software application (Roberts et al., 2007) have been developed to facilitate its use. However, a limitation of MESMA over spectral matching algorithms that use a similarity metric derived from spectral angles rather than overall reflectance (i.e. Spectral Angle Mapper, SAM, Kruse et al., 1993), is that classification accuracy can be strongly influenced by variations in albedo (Dennison et al., 2004).

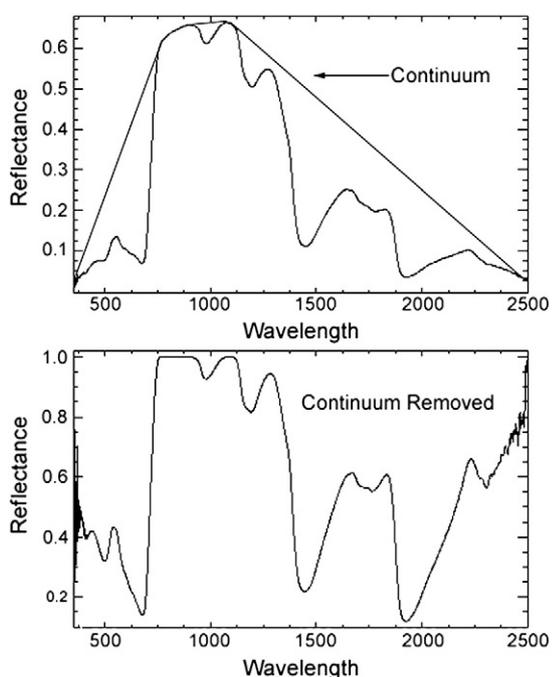
### 2.2. Continuum removal

Continuum removal (CR) is a spectral processing technique that normalizes brightness while emphasizing absorption features (Clark & Roush, 1984). A convex hull, or continuum, is fitted over a spectrum to connect the points of maximum reflectance with a straight line (Fig. 1). The continuum can be applied to selected segments or across the entire spectrum. The peak reflectance points where the actual spectrum meets the continuum line are standardized to a value of one and this value decreases towards zero as the distance between the original spectrum and continuum line increases. The continuum is removed by dividing reflectance value ( $\rho$ ) of a specific wavelength ( $\lambda$ ) by the reflectance value of the continuum ( $\rho_{c\lambda}$ ) at the corresponding wavelength (Eq. 3):

$$\text{CR} = \frac{\rho_\lambda}{\rho_{c\lambda}} \quad (3)$$

In remotely sensed data, the effects of field of view and photon scattering can alter radiance reaching the sensor (Richards & Jia, 2006). Subtle absorption features in reflectance data are enhanced in the normalization process of CR and their depth and position are not influenced by variations in albedo (Schmidt & Skidmore, 2003).

CR is most often used in spectral data analysis to identify the geological composition of materials and to quantify vegetation biochemistry (Huang et al., 2004; Kruse & Lefkoff, 1993; Mutanga &



**Fig. 1.** A continuum line fit over a *Eucalyptus* leaf reflectance spectrum (upper) and the resulting continuum-removed spectrum (lower).

Skidmore, 2004; van der Meer, 2004). Several other spectral transformation methods such as derivative analysis and wavelet transformation are also commonly used to facilitate these analyses (e.g., Huang et al., 2004; Leung et al., 1998). Derivative analysis and wavelet transformation involve spectral smoothing and feature reduction algorithms that can eliminate unnecessary signal components and emphasize subtle absorption features. They are sometimes combined with CR to take advantage of its normalization technique, which is not a feature of these other methods (e.g., Huang et al., 2004). Derivative analysis and wavelet transformation can be useful for class discrimination (e.g., Castro-Esau et al., 2004; Li, 2004; Somers et al., 2009a,b; Zhang et al., 2004). However, unlike CR, they introduce the risk of discarding important spectral features.

The application of CR for mapping vegetation types is less common than its use in quantifying vegetation biochemistry from spectral data; however, a few studies have demonstrated its potential value. A normalized spectral mixture analysis (NSMA) developed by Wu (2004), improved the performance of a vegetation–impervious surface–soil (V–I–S) model. Likewise, normalized reflectance values were helpful in differentiating the spectral features of lichen and rock in mixed pixel spectra (Zhang et al., 2005) and vegetation types in a coastal wetland (Schmidt & Skidmore, 2003). Spectral feature fitting (SFF) is a spectral matching algorithm that uses continuum-removed spectra and has been applied with success to mapping vegetation species, functional types and condition (Kokaly et al., 2003, 2007). Hestir et al. (2008) used CR to isolate water absorption features from spectra and this enabled them to map an invasive, succulent plant species with SMA. Filippi and Jensen (2007) were less successful in using CR in conjunction with artificial neural networks to classify coastal vegetation. However, their method did not rely on end-member selection. To our knowledge, the application of continuum-removed spectra to MESMA has never been investigated for vegetation mapping.

### 2.3. *Eucalyptus* forests

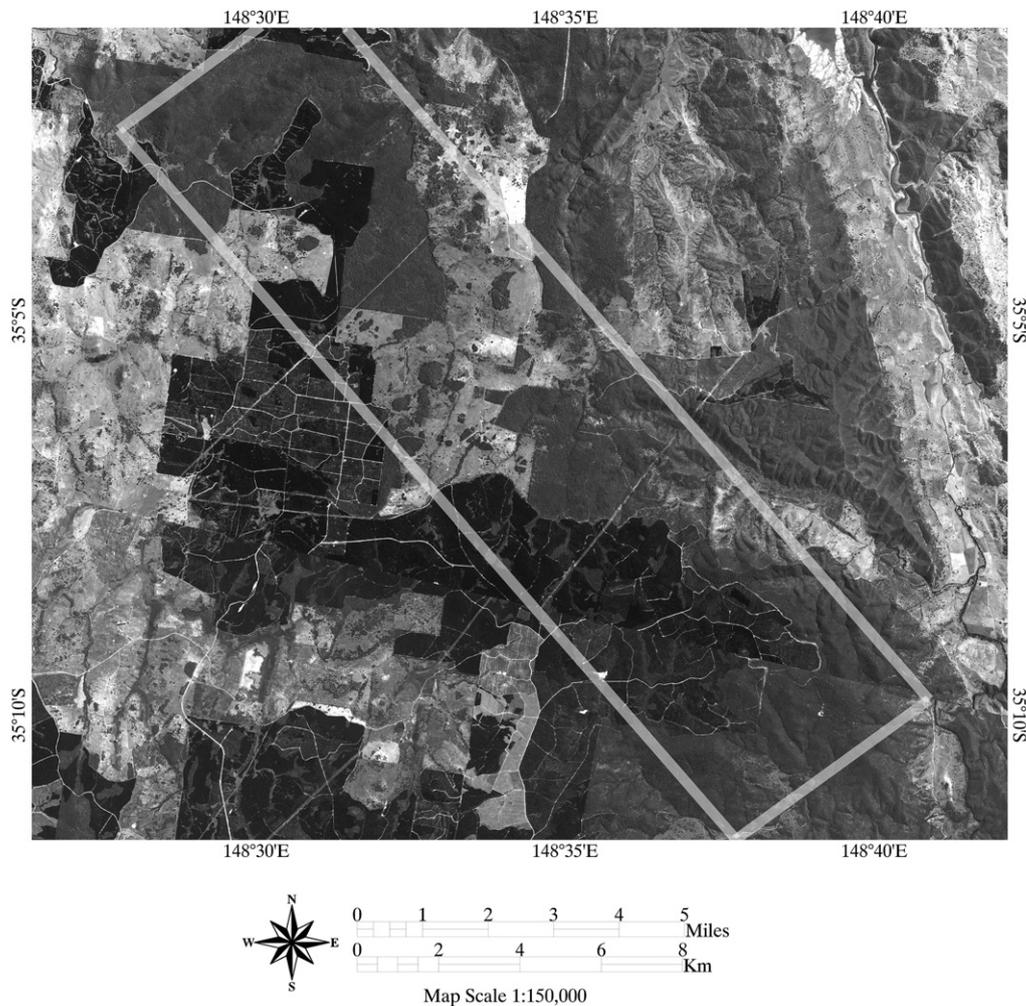
*Eucalyptus* (L'Herit) is a genus of broad-leaved, evergreen trees that dominate most forest and woodland landscapes in Australia

(Williams & Woinarski, 1997). Several eucalypt species are also important exotic plantation trees and invasive species in the Americas, Africa and Asia (Davidson, 1993; Rouget et al., 2002). Unfavorable structural characteristics, such as pendulous leaves and open, overlapping canopies combined with a high degree of spectral variability within species and similarity between species make *Eucalyptus* tree species notoriously difficult to discriminate using remote sensing data (Coops et al., 2004; Goodwin et al., 2005). A recent attempt to map eucalypt species with hyperspectral remote sensing data was largely successful but it required the use of multiple sensors (the Compact Airborne Spectrographic Imager (CASI) and Hyperspectral Mapper (HyMap)) and an open woodland structure to facilitate crown delineation (Lucas et al., 2008). Youngentob et al. (2008) suggested that mapping *Eucalyptus* subgenera could provide critical information about landscape scale processes and wildlife habitat and that this may be more feasible and affordable with current, single-sensor technology.

There are over 700 species of *Eucalyptus* trees. Approximately 440 of those belong to the subgenus *Symphomyrtus* and 140 to *Eucalyptus* (Noble, 1989). Monocalypts and symphyomyrtles occur throughout most of the range of environments tolerated by the *Eucalyptus* genus as a whole, although tree species from specific subgenera are characteristic of particular ecosystem types and elevations (Austin et al., 1983; Davidson & Reid, 1980; Hughes et al., 1996). Mixed stands comprised of these two groups are also common in native forests (Pryor, 1959), but their spatial distributions are not random (Austin et al., 1983) and they tend to deviate from one another in their early growth characteristics (Davidson & Reid, 1980), response to fire (Duff et al., 1983; Noble, 1984), response to salinity (Marcar, 1989; Noble, 1989), concentrations of foliar chemicals (Eschler et al., 2000; Gleadow et al., 2008; Hawkins & Polglase, 2000; Hill et al., 2001) and susceptibility to predation from insect and mammalian folivores (Moore et al., 2004; Stone et al., 1998). A Bhattacharyya distance (B-dis) calculation applied to hyperspectral data collected from the main canopy species in our study area suggested that the spectral properties of those monocalypts and symphyomyrtles are sufficiently distinct to enable identification and mapping with data from a high resolution, hyperspectral sensor that includes NIR and shortwave infrared (SWIR) wavelengths (Youngentob et al., 2008). The bands identified in the B-dis approach combined with a maximum likelihood classification was highly successful for discriminating eucalypt subgenera based on sunlit pixels collected from HyMap reflectance data (testing Kappa 0.92) (Youngentob et al., 2008). However, even when additional “shade” classes comprised of lower reflectance pixels for each subgenera were included, this method produced a map that did not match expectations of field distributions based on our knowledge of the region (Youngentob et al. unpublished).

### 3. Study area

Our research site is located in the Tumut region of southern New South Wales, Australia (midpoint = 148°30'E, 35°10'S) (Fig. 2). Over 50,000 ha of native *Eucalyptus* forests were cleared from this region between 1930 and 1985 to establish an exotic pine (*Pinus radiata*) plantation. This research focused on the remaining native eucalypt forest blocks and isolated remnant forests and paddock trees in and around Buccleuch, Bungongo and Bondo State Forests and Brindabella National Park. The native forest range from dry sclerophyll with characteristic apple box (*Eucalyptus bridgesiana*) and red stringy bark (*E. macrorhynca*) interspersed with broad-leaved peppermint (*E. dives*) and manna gum (*E. viminalis*) to the more prevalent tall-open and montane forest where manna gum, mountain gum (*E. dalrympleana*) and narrow-leaved peppermint (*E. radiata*) thrive in the cooler, wetter environment. Snow gum (*E. pauciflora*) is present and increases in abundance at higher elevations. Mountain swamp gum (*E. camphora*) and black sallee (*E. stellulata*) are common in low-lying, marshy areas.



**Fig. 2.** A Systeme Observation de la Terra (SPOT) image of a portion of the study site in the Tumut region of New South Wales, Australia. The lighter rectangle outlines the approximate area encompassed by the five northern HyMap flight-lines. The darkest portions of the image are *Pinus radiata* forest and these are surrounded by blocks of native *Eucalyptus* forest (medium gray) and grazed paddocks (light gray).

Brown-barrel (*E. fastigata*) and alpine ash (*E. delegatensis*) are present but relatively uncommon in our study area. The topography of the region features low to moderate relief undulating hills ranging from 600 to 1200 m above sea level and becomes mountainous to the east (Brindabella Ranges) and south (Snowy Mountains). The underlying geology is comprised of granite plutons interbedded with sandstone and shale, which formed during the middle Paleozoic as part of a larger geologic subdivision known as the Lachlan Fold Belt. Annual temperatures average 15.3 °C–29.3 °C in summer to 3.0 °C–11.5 °C in winter (mean min and mean max; Burrinjuck Climatic Station). Precipitation is distributed evenly across the year and typically ranges from 785 to 1385 mm annually (BIOCLIM, Nix, 1986).

#### 4. Methods

Due to the large number of protocols in our methodology, we have provided a schematic diagram to briefly outline each step (Fig. 3). The steps are explained in detail in the following sub-sections.

##### 4.1. Hyperspectral data acquisition and preprocessing

We collected HyMap data on March 8th, 2007 (HyVista Corporation Pty Ltd). HyMap is an airborne imaging spectrometer that captures 126 wavebands over a spectral range of 446.1 nm to 2477.8 nm at bandwidth intervals of 10 nm in the visible (VIS) and NIR wavelengths and 15–20 nm in the SWIR (Cocks et al., 1998). At 1500 m flying

altitude, the ground instantaneous field of view was approximately 3.5 m with a swath width of 1.8 km (512 pixels). Imagery was collected between noon and 3 pm under clear-sky conditions. Five adjacent, NE–SW flight-lines ranging from 15 to 20 km in length were flown in the north of the study area and six flight-lines in the south. This study focuses on the region within the five northern flight-lines. The HyMap data were provided in a geo-corrected format based on positional data (UTM-WGS-84, Zone 55 S) and atmospherically corrected to reflectance using HyCorr, a version of ATREM for HyMap (CSES, 1992; Gao & Goetz, 1990). This method uses a radiative transfer model to calibrate to absolute reflectance based on the atmospheric water absorption features at 940 nm and 1140 nm. The apparent reflectance data were further corrected for residual noise using Empirical Flat Field Optimal Reflectance Transformation (EFFORT, Boardman, 1998). One corrupted waveband at 446.1 nm and two additional bands (1389.1 nm and 1403.9 nm) located in a spectral region that is strongly influenced by water vapor were subsequently removed, resulting in 123 bands.

We performed our image analyses using ENVI software (Research Systems, Inc., Boulder, Colorado). We used a normalized difference vegetation index (NDVI)-based mask to remove pixels dominated by paddock fields and other grasses (non-photosynthetic vegetation in late summer) as well as soil, roads and quarries (Xiao et al., 2004). We easily identified the pine compartments in a black and white image (Fig. 2) because they appeared darker than the native eucalypt forest. We excised the pine compartments from the image with a “regions of interest” (ROI)-based mask. We also found that *P. radiata* trees had

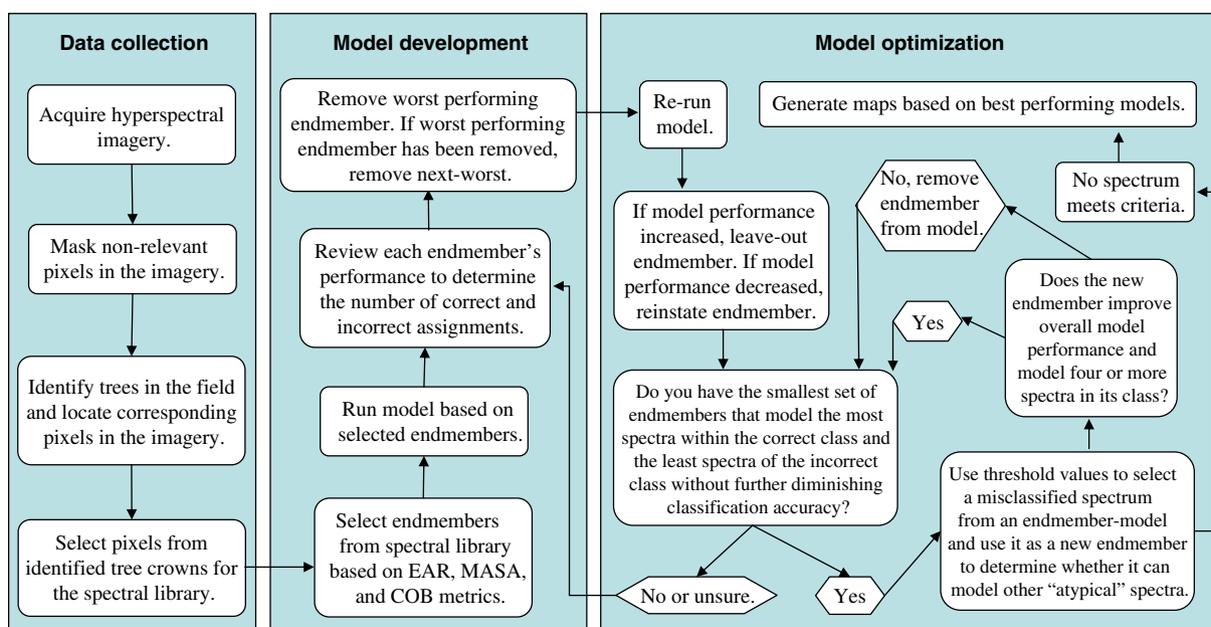


Fig. 3. A schematic diagram showing the sequence of steps in our classification methodology.

higher NDVI values than eucalypt trees, possibly owing to denser canopies with higher, overall nitrogen content (Myers et al., 1996; Poggiani, 1985). We masked NDVI values greater than 0.93 to remove pine wildlings from the native eucalypt forest. The remaining unmasked portion of the image was comprised primarily of eucalypt forest pixels and isolated eucalypt paddock tree crowns. Continuum-removal analysis was applied to the whole spectrum of every pixel in the masked imagery in order to create a separate "normalized" reflectance image from the hyperspectral data.

#### 4.2. Identification of individual *Eucalyptus* trees

We located isolated eucalypt paddock trees ( $n=99$ ) in the field that we could match to tree crowns in the HyMap imagery. These trees were then identified to the subgenera and species level (Table 1). We attempted to obtain at least six representatives from each of the nine most common forest tree species (listed in Section 3). However, *E. pauciflora*, *E. dives*, *E. macrorhyncha* and *E. stellulata* were relatively rare as isolated trees or homogenous tree clusters in the paddocks and this resulted in a smaller number of trees identified. Those species and *E. bridgesiana* (common in the paddocks) were also less abundant in the contiguous forest than *E. viminalis*, *E. dalrympleana* and *E. radiata*. A large number of *E. radiata* were identified to insure that a suitable quantity of spectra could be collected to represent the *Eucalyptus* subgenera, since this species was the most abundant monocalypt in the contiguous forest and the paddocks.

#### 4.3. Tree pixel selection for spectral library

To identify relatively pure eucalypt tree-canopy spectra from the imagery, we first displayed the masked HyMap reflectance and continuum-removed images in three wavelengths from the SWIR (1.65 nm), NIR (0.84 nm) and VIS red-edge (0.67 nm) regions of the electromagnetic spectrum. Viewed in these wavelengths, green pixels indicate high concentrations of chlorophyll containing vegetation (e.g. canopy leaves) and purple, blue and white pixels are either not as photosynthetically active (e.g. bark and branches) or highly shaded (Fig. 4). We selected individual pixels ( $n=531$ ) from the previously identified eucalypt paddock trees that had high NIR values relative to SWIR and VIS values in both the continuum-removed and the reflectance image. Separate libraries were created for the reflectance and continuum-removed spectra. However, the spectra in the continuum-removed library

had the same pixel locations in the imagery as the spectra in the reflectance library. The selected pixels were assigned to the appropriate eucalypt species-level and subgenera-level classes based on *a priori* knowledge from field verification (Table 1). We also selected shaded, non-green pixels ( $n=104$ ) from the identified paddock trees. These pixels were typically located on the heavily shaded side of the canopy relative to the viewing geometry of the sensor and the angle of the sun. The shade spectra were often "noisier" due to lower reflected radiance and were assigned to a separate species-level class, "shade/noise", within the appropriate subgenera-level class. We did not select any of the darkest pixels (dark blue or black in reflectance image). Duplicate pixels in the imagery can be created as an artifact of the geo-referencing process so spectra that had identical reflectance values as a previously selected spectrum from the same tree were not included in the library.

#### 4.4. Endmember selection for models

We used ViperTools (Roberts et al., 2007) for endmember-model selection, MESMA and class-mapping. Endmember selection is an important aspect of MESMA that should take into consideration the

Table 1

The number of individual paddock trees (listed by species) from the two major eucalypt subgenera *Eucalyptus* (common name "monocalypt") and *Symphomyrtle* (common name "symphyomyrtle") and the corresponding number of HyMap pixels selected for the endmember spectral library.

Vegetation class	Individual trees (n)	Selected pixels (n)
<i>Eucalyptus bridgesiana</i>	11	38
<i>Eucalyptus camphora</i>	6	19
<i>Eucalyptus dalrympleana</i>	7	51
<i>Eucalyptus viminalis</i>	13	71
Shade/noise		50
Symphomyrtle (total)	37	229
<i>Eucalyptus dives</i>	2	8
<i>Eucalyptus macrorhyncha</i>	2	9
<i>Eucalyptus pauciflora</i>	3	18
<i>Eucalyptus radiata</i>	52	299
<i>Eucalyptus stellulata</i>	3	18
Shade/noise		54
Monocalypt (total)	62	406

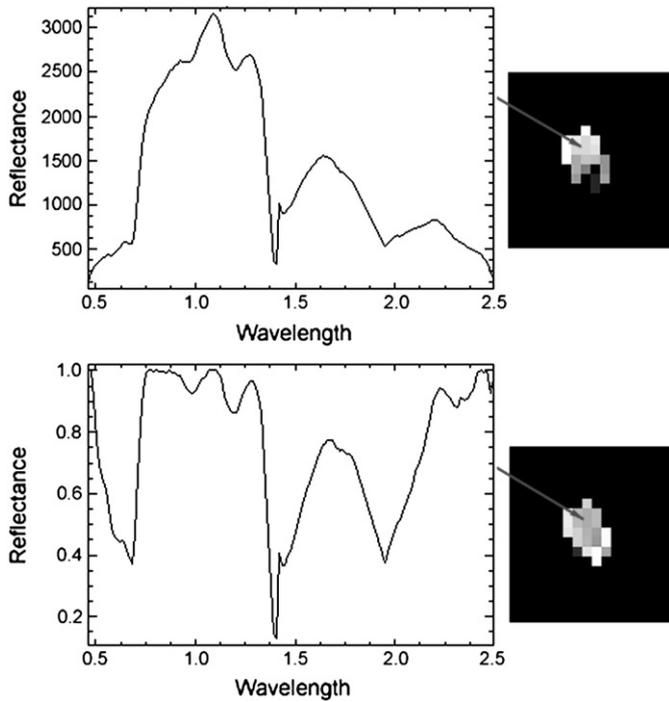


Fig. 4. Spectral profiles for a pixel collected from a reflectance (upper) and continuum-removed (lower) HyMap image of an isolated eucalypt tree crown.

spectral diversity of the library and computational efficiency, since a spectral library can be comprised of hundreds of spectra for each material class (Dennison et al., 2004; Dennison & Roberts, 2003b). We used two average spectral similarity metrics, endmember average root mean square error (EAR; Dennison & Roberts, 2003a) and minimum average spectral angle (MASA; Dennison et al., 2004), to select endmembers from the spectral library that were most representative of their species-level vegetation class. Both of these methods model the spectra within a vegetation class testing each spectrum as an endmember. The spectrum with the lowest average error metric is the endmember that best models the other spectra in its class, on average (Dennison & Roberts, 2003a).

The error metric for EAR is RMSE so that:

$$EAR_i = \frac{\sum_{j=1}^n RMSE_{ij}}{n-1} \quad (4)$$

where  $i$  is an endmember,  $j$  is the modeled spectrum, and  $n$  is the number of modeled spectra. The “−1” corrects for the zero error resulting from an endmember modeling itself.

MASA is similar to EAR, but uses a spectral angle ( $\theta$ ) as the error metric. Spectral angle is calculated as:

$$\theta = \cos^{-1} \left( \frac{\sum_{\lambda=1}^M \rho_{\lambda} \rho'_{\lambda}}{L_{\rho} L_{\rho'}} \right) \quad (5)$$

where  $L_{\rho}$  is the length of the endmember vector and  $L_{\rho'}$  is the length of the modeled spectrum vector calculated as:

$$L_{\rho'} = \sqrt{\sum_{\lambda=1}^M \rho_{\lambda}^2} \quad (6)$$

MASA is then calculated as:

$$MASA_i = \frac{\sum_{j=1}^n \theta_{ij}}{n-1} \quad (7)$$

As EAR is based on RMSE from a linear spectral mixing model, it is influenced by albedo. We applied minimum (0%) and maximum (100%) shade-fraction constraints to decrease the likelihood that very light or very dark spectra would be identified as highly representative endmembers for their class by increasing the RMSE of spectra that exceeded the shade fraction thresholds (Dennison & Roberts, 2003b). MASA is not influenced by variations in albedo. The algorithms were run separately for the reflectance and the continuum-removed spectral libraries.

A third metric, Count-based Endmember Selection (COB; Roberts et al., 2003) is also included in Viper Tools. COB uses the MESMA concept to select endmembers based on the number of library spectra each endmember models. COB determines the number of spectra modeled by an endmember within the endmember's class (InCOB) and outside of the endmember's class (OutCOB). The optimum model would have the highest InCOB and lowest OutCOB. InCOB as implemented in Viper Tools also includes an iterative element, in which spectra that are modeled by the endmember with the highest InCOB are subsequently removed from the selection process, creating a smaller library consisting only of spectra that did not meet the fit criteria for the highest InCOB. Next, the second highest InCOB, calculated using the remaining spectra is reported, followed by InCOBs calculated for progressively smaller numbers of spectra unmodeled by earlier selections. Constraints used for MESMA can also be used for COB, including fraction and RMSE constraints. A RMSE threshold was applied to exclude reflectance spectra endmembers that exceeded 0.025 RMSE (Roberts et al., 1998). A RMSE threshold of 0.06 was applied to the continuum-removed spectra. Numerous studies have found 2.5% to be an acceptable threshold value for reflectance spectra (e.g., Roberts et al., 1998); however, no such studies have been done with CR spectra. We investigated several thresholds for CR and selected the value that provided the best discrimination with the fewest number of endmembers excluded. Given that this was determined empirically, a different spectral library with a different set of materials to discriminate might require a different threshold.

#### 4.5. Model optimization

The final set of endmembers was determined through an interactive, iterative process by which endmembers were both added and subtracted from the set to vary overall and class accuracy. The total number of endmembers selected for each species-level class varied depending on the size of each class. We did not select more than half of the spectral library for any class. We initially selected a minimum of 4 and a maximum of 10 candidate endmembers for each species-level class by choosing the endmembers with the lowest EAR values. Minimum EAR endmembers often included one or more of the lowest MASA endmembers as well. However, when the lowest EAR value endmembers failed to capture the lowest MASA value, we selected up to 4 additional lowest MASA value endmembers for each class. Where there were multiple endmembers with nearly identical lowest EAR or lowest MASA values, we gave preference to those endmembers that had the highest InCOB. In a few cases, additional endmembers were selected using InCOB where the lowest EAR and lowest MASA endmembers failed to model a subset of spectra within a class. This was only done if the total number of endmembers selected had not exceeded half of the total spectral library for that species-class.

This selection process does not account for spectral confusion between classes. The endmembers with the lowest EAR or MASA values may also model another class better than that class's own endmembers (Powell et al., 2007). To determine which endmembers were “over-representative”, we applied the selected endmembers to the entire spectral library and calculated overall species classification accuracy. Each endmember was reviewed to determine the number of correct and incorrect assignments. This enabled us to identify those endmembers that modeled a large number of spectra from outside their species-class. As expected, there was considerably more class confusion at the species-level than at the subgenera-level. Since we were primarily concerned with our ability to discriminate between the two eucalypt subgenera, we focused on excluding those endmembers that modeled a large number of spectra as the wrong subgenera class. This was an iterative process that began by removing an endmember with the greatest model error at the subgenera-level. We then re-ran the models with the smaller number of selected endmembers. If removing an endmember resulted in an increase in overall accuracy or no change at all, we left that model out. If removing an endmember resulted in a decrease in overall accuracy, we reinstated that endmember-model. We repeated this process until we had the smallest set of endmembers that modeled the most spectra within the correct subgenera class and the least spectra of the incorrect subgenera class without further diminishing our classification accuracy. We retained at least one endmember from each species-level class in the final model. Due to a high degree of spectral diversity in the shade/noise category, we selected a sufficient number of shade/noise endmember-models to ensure that at least 75% of the shade/noise spectra in the library were modeled. We required that the accuracy of each species-level class in the final model exceed 50% (producer's accuracy) for modeling the correct subgenera for all model classes except *E. radiata* and *E. viminalis* (the most abundant forest species), which we required to have at least 70% producer's accuracy.

Some endmembers may not be highly representative of their class but still important for modeling other “atypical” spectra in their class. Once we had completed the process described above, we investigated whether we could further improve the performance of the final model by selecting misclassified spectra from the worst performing endmember-models and using them as new endmembers. We looked at each of the final endmember-models and selected a misclassified spectrum from a model if the RMSE between the model-endmember and the misclassified spectrum approached a threshold value (0.025 reflectance and 0.06 continuum-removed). The logic was that if these spectra were poorly fit by that endmember-model, then they may be a better endmember for modeling other misclassified spectra from their own class. We then re-ran the model with the additional endmember. We left the new endmember in the model if it improved overall model performance and modeled four or more spectra in its class. We then repeated this process with other endmembers until we were unable to improve model performance or further reduce the total number of endmembers in the final model according to the criteria above.

## 5. Results and discussion

### 5.1. Error assessment

We used a Kappa analysis to measure the agreement between the reference data and the supervised classifications (Monserud & Leemans, 1992). We focused our results on the ability of selected endmembers to correctly model other library spectra by the appropriate subgenera class, although we report Kappa statistics for classification accuracies at both the species-level and the subgenera-level. A confusion matrix and associated accuracy assessments at the subgenera-level are presented in Tables 2 and 3 for the best performing models from the reflectance and continuum-removed

data. The continuum-removed data performed better than the non-transformed data in the majority of models.

We found low species-level accuracy for both the non-transformed and the continuum-removed hyperspectral reflectance data. Overall accuracy for the non-transformed data was 35% (Kappa = 0.25), and overall accuracy for the continuum-removed data was 46% (Kappa = 0.31). Accuracy was much higher at the subgenera-level. Overall accuracy for the non-CR reflectance spectra was 75% (Kappa = 0.48) based on the best 30 endmember-model (Table 2) and 83% (Kappa 0.63) for the best 34 endmember continuum-removed model (Table 3). The spectra from the selected endmember-models are displayed in Fig. 5.

The ability of endmember-models from each eucalypt tree species to appropriately model other spectra at the subgenera-level varied widely in both producer's accuracy (determined by errors of omission) and user's accuracy (based on errors of commission). For the majority of models, CR models were less likely to exclude spectra from the appropriate subgenera class than non-CR models and less likely to model spectra from the wrong subgenera than non-CR models. However, the best performing CR model left a higher percentage of the spectral library unmodeled (2%) than the best performing non-CR model (0%). Models that used CR-spectra required a higher RMSE threshold because the transformation process can increase the signature to noise interference and subsequently increase the diversity of spectra within-groups and impact model-fit (Abilio et al., 2007). This is also why a larger number of endmember-models were required to obtain the best classification accuracies for the CR data compared to the non-CR data.

### 5.2. Classification error in relation to eucalypt tree species physiognomy

Plant species that share certain structural and spectral characteristics, such as leaf shape and color or foliage density, are generally more difficult to discriminate with imaging spectrometry than more distinct species or plant types (Castro-Esau et al., 2006; Clark et al., 2005; Goodwin et al., 2005; Hestir et al., 2008). Shade resulting from open-canopy structure or viewing geometry can also increase confusion among tree-class types due to spectral similarities resulting from the lower reflectance values of these shaded regions (Asner et al., 2000; Lucas et al., 2008). Most eucalypt tree species have a relatively open canopy structure. An imaging spectrometer flown over a eucalypt forest is likely to collect radiation reflecting from the ground, trunk, and upper branches of the trees in addition to the canopy foliage. Therefore, similarities among eucalypt tree-classes in canopy density, and bark, branch and foliage color are all likely to influence classification accuracies. For example, the brown fibrous bark that covers the trunk and thick branches of *E. bridgesiana* (*Symphyomyrtus*) is more similar in appearance to the bark of the monocalypt species in our study area than to the smooth gum and ribbony bark of the three other symphyomyrtles. We expected to see more confusion between canopy spectra of *E. bridgesiana* and the monocalypts for this reason. Non-CR *E. bridgesiana* models performed the poorest for subgenera-level classification in user's accuracy (57%), although the producer's accuracy was somewhat higher (66%). *E. bridgesiana* models from the continuum-removed data had a considerably higher user's and producer's accuracy (78% and 80% respectively).

Similarly, *E. pauciflora* has a smoother and lighter bark than the other monocalypt species in this region. The leaves are also more similar in color to the lighter green leaves of the three symphyomyrtles than to the other four monocalypts. For these reasons, we expected relatively high classification errors for *E. pauciflora*. We found that *E. pauciflora* models had the poorest subgenera-level classification in producer's accuracy for reflectance data (56%) and the second poorest for continuum-removed data (59%).

**Table 2**  
Confusion matrix and classification accuracies for hyperspectral reflectance data of 9 eucalypt tree species from 2 subgenera, *Eucalyptus* (common name “monocalypt” (M)) and *Symphomyrtus* (common name “symphyomyrtle” (S)).

Correct matrix		Reference pixels											User's accuracy subgenera-level (%)	
		Eubr (S)	Euca (S)	Euda (S)	Euvi (S)	Shade (S)	Eust (M)	Eupa (M)	Eudi (M)	Euma (M)	Eura (M)	Shade (M)		
Mapped pixels	Eubr (S)	10	2	4	7	1	2	1	0	0	15	0	57	
	Euca (S)	3	2	3	1	0	0	1	0	0	2	0	75	
	Euda (S)	5	6	12	12	3	0	3	0	2	22	0	59	
	Euvi (S)	5	1	11	39	0	0	2	0	1	29	0	64	
	Shade (S)	0	0	0	0	33	0	0	0	0	0	20	62	
	Eust (M)	4	0	3	2	0	1	1	0	0	23	0	74	
	Eupa (M)	5	3	7	2	0	0	7	1	2	40	0	75	
	Eudi (M)	0	1	2	1	0	10	1	3	0	81	1	96	
	Euma (M)	0	0	1	0	0	0	0	2	2	10	0	92	
	Eura (M)	3	2	5	2	0	1	0	2	1	74	2	87	
	Shade (M)	0	0	0	1	8	2	0	0	0	2	26	77	
	Producer's accuracy subgenera (%)		66	65	63	88	82	81	56	100	63	77	59	
	Correct matrix		Monocalypt					Symphyomyrtle					Overall user's acc.	
Monocalypt		293					52					85%		
Symphyomyrtle		100					160					62%		
Overall producer's acc.		75%					75%					0 unmodeled		
Subgenera level overall accuracy = 75%							Species level overall accuracy = 35%							
Kappa statistic = 0.48							Kappa statistic = 0.25							

Pixels collected from the shaded portion of eucalypt tree canopies (S/N, shade-noise) were assigned to their own class within the appropriate subgenera designation. Eucalypt tree species names are abbreviated (Eu+the first two letters of the species name). Self-modeled endmembers are not included in the table or accuracy calculations.

*E. camphora* was one of the few models to perform worse in both producer's and user's accuracies in the CR models (53% and 55% respectively) than in the non-CR models (65% and 75% respectively). *E. camphora* has a very open canopy structure, relatively short stature (generally <10 m) and prefers low lying, marshy areas with poorly drained soils. This may have contributed to mixed pixels and lower overall reflectance values for the spectra from this species (Fig. 5) and also resulted in reduced classification accuracies. The use of lidar in combination with hyperspectral remote sensing imagery may improve classification accuracies for species, such as *E. camphora*, that are

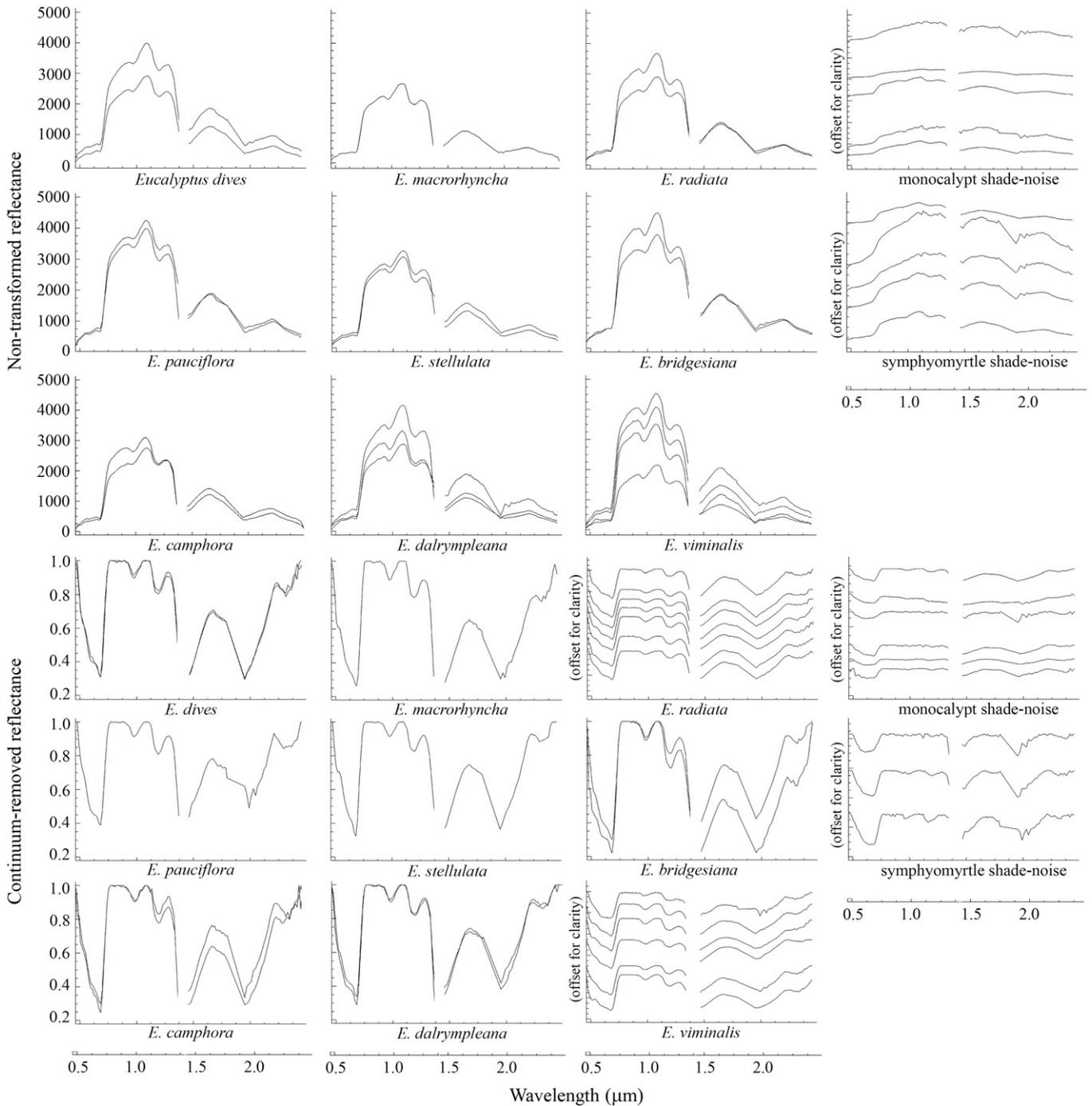
distinct in their size relative to the other species but have spectral reflectance properties that are less distinct (Gillespie et al., 2004).

*E. viminalis* and *E. dalrympleana* are closely related, white, ribbon-barked trees that differ from one another primarily in the shape of their young leaves. *E. radiata* and *E. dives* are two peppermint species that are also very similar, except that the latter has broader adult leaves. A high degree of species-level confusion within these pairs was expected. When these pairs were combined, species-level accuracy improved to 58% (kappa 0.42) in the continuum-removed data and 52% (kappa 0.38) in the reflectance data.

**Table 3**  
Confusion matrix and classification accuracies for continuum-removed hyperspectral reflectance data of 9 eucalypt tree species from 2 subgenera, *Eucalyptus* (common name “monocalypt” (M)) and *Symphomyrtus* (common name “symphyomyrtle” (S)).

Correct matrix		Reference pixels											User's accuracy subgenera-level (%)	
		Eubr (S)	Euca (S)	Euda (S)	Euvi (S)	Shade (S)	Eust (M)	Eupa (M)	Eudi (M)	Euma (M)	Eura (M)	Shade (M)		
Mapped pixels	Eubr (S)	6	3	7	12	1	0	0	0	0	8	0	78	
	Euca (S)	0	1	2	3	0	0	1	0	0	4	0	55	
	Euda (S)	5	3	12	2	0	0	2	0	0	6	0	73	
	Euvi (S)	17	1	21	34	6	0	3	0	0	17	0	80	
	Shade (S)	0	1	0	1	25	0	1	0	0	0	13	66	
	Eust (M)	0	2	0	0	0	1	2	0	0	1	0	67	
	Eupa (M)	0	0	0	1	1	0	4	0	0	5	0	82	
	Eudi (M)	0	0	2	1	0	1	2	0	3	46	0	95	
	Euma (M)	1	1	1	1	0	1	0	1	0	12	0	78	
	Eura (M)	6	2	4	9	0	7	1	5	5	162	3	90	
	Shade (M)	0	3	0	0	11	7	1	0	0	31	25	82	
	Producer's accuracy subgenera (%)		80	53	86	81	73	100	59	100	100	88	68	
	Correct matrix		Monocalypt					Symphyomyrtle					Overall user's acc.	
Monocalypt		326					46					88%		
Symphyomyrtle		55					163					75%		
Overall producer's acc.		86%					78%					10 (2%) unmodeled		
Subgenera level overall accuracy = 83%							Species level overall accuracy = 46%							
Kappa statistic = 0.63							Kappa statistic = 0.31							

Pixels collected from the shaded portion of eucalypt tree canopies (S/N, shade-noise) were assigned to their own class within the appropriate subgenera designation. Eucalypt tree species names are abbreviated (Eu+the first two letters of the species name). Self-modeled endmembers are not included in the table or accuracy calculations.

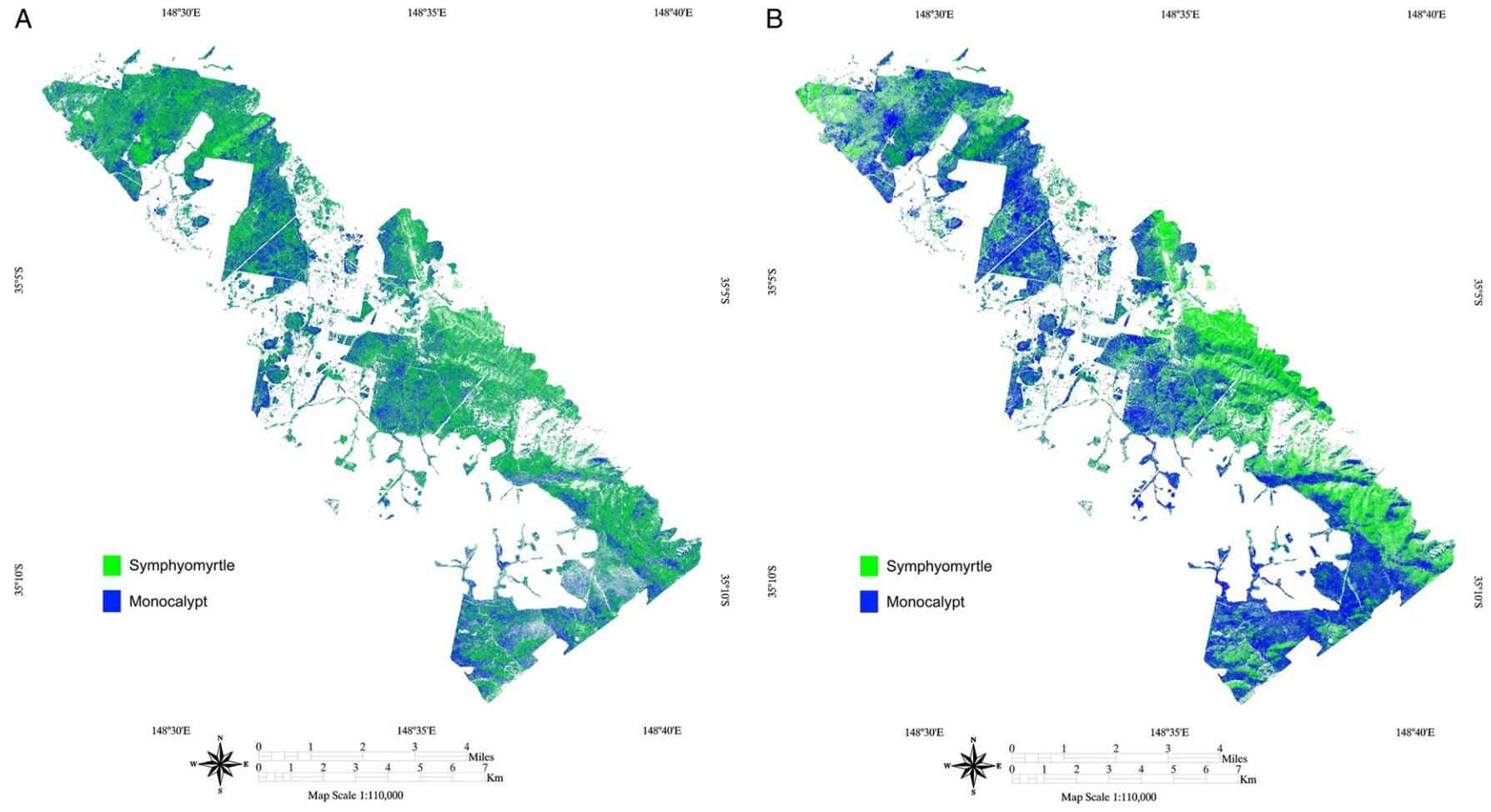


**Fig. 5.** Selected imagery endmember spectra for models based on non-transformed (top) and continuum-removed (bottom) HyMap reflectance data of eucalypt tree-crowns. Individual tree species names are listed under each graph.

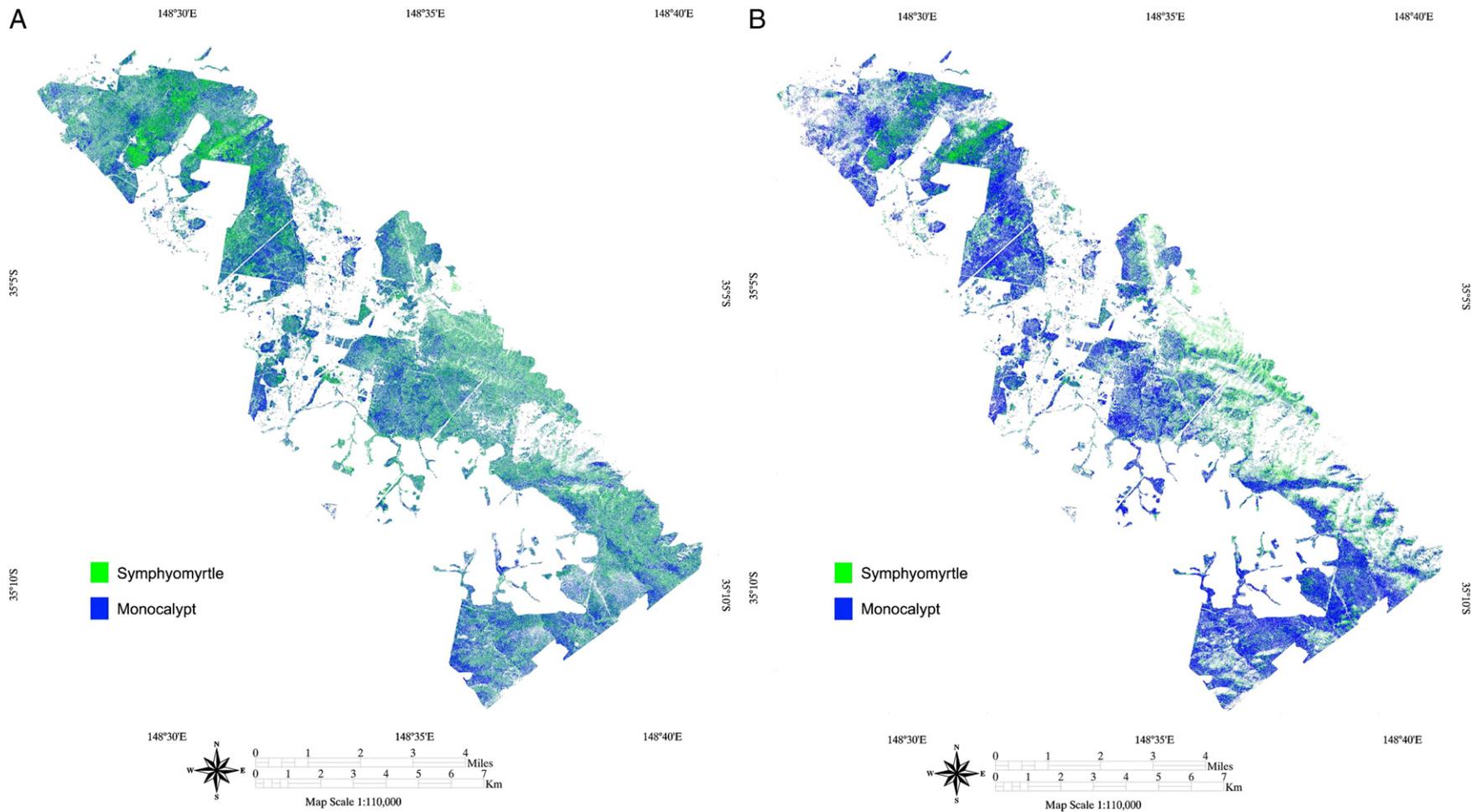
5.3. Map generation

We used the best performing continuum-removed and reflectance endmember-models (Fig. 5) to generate maps of the distribution of monocalypts and symphyomyrtles in our study area (Fig. 6A and B). We applied the same shade fraction and RMSE constraints to model the full-scene that we used to model the spectral library. Based on our individual model assessment (Section 4.5), we identified those endmembers that reported <75% accuracy in modeling spectra from the appropriate subgenera and masked-out the pixels in the images that were modeled by those endmembers in order to create two “higher-confidence” maps (Fig. 7A and B). Excluding pixels that were

modeled by the minority subset of endmembers that fell below this 75% threshold provided an interesting, additional end-product that should reflect a higher confidence in predicted class distributions than the maps based on the full-suite of selected endmember-models. However, removing the pixels that were less confidently mapped meant that a larger portion of the final image remained unmodeled. We also must emphasize that our accuracy assessments were based on the reference data (spectral library) and the final prediction accuracies may be different if the spectral diversity of the spectral library was not representative of the spectral diversity in the full-image (Powell et al., 2007). For example, if image spectra possessed a higher abundance of frequently miss-modeled spectra than the library



**Fig. 6.** Vegetation maps of the two major eucalypt subgenera, Eucalyptus (monocalypt) and Symphyomyrtus (symphyomyrtle) based on A) the best-performing continuum-removed model comprised of 34 spectral endmembers and B) the best-performing reflectance model comprised of 30 spectral endmembers. Regions in the imagery that were not comprised of eucalypt forest have been masked.



**Fig. 7.** Vegetation maps of the two major eucalypt subgenera, *Eucalyptus* (monocalypt) and *Symphyomyrtus* (symphyomyrtle) based on pixels modeled by endmembers that reported  $\geq 75\%$  accuracy from A) the best-performing continuum-removed model comprised of 34 spectral endmembers and B) the best-performing reflectance model comprised of 30 spectral endmembers. Regions in the imagery that were not comprised of eucalypt forest have been masked.

did, the actual accuracy of the classified image would be lower than our estimated accuracy.

In our northeastern-most flight-line, we observed that symphyomyrtles tended to be over-modeled at higher elevations and monocalypts appeared to be over-modeled on the southern aspects of high-relief areas. An analysis of change in vegetation composition performed by Shugart and Nobel (1981) in the Brindabella Ranges found that higher altitude sites (above 1200 m) were dominated by *E. pauciflora*, a monocalypt species. Due to physiological characteristics discussed in the previous section, this species may be more spectrally similar to the most common symphyomyrtle species in our regions, *E. viminalis* and *E. dalrympleana*, than to the other monocalypts. This likely contributed to the classification error in the higher elevation regions along our northeastern-most flightline.

We expected to see higher densities of the most dominant monocalypt, *E. radiata*, on southern slopes because they prefer cooler and wetter environments (Williams & Woinarski, 1997) and this was clearly visible in the imagery. However, the near-homogenous monocalypt densities that are reported along some of the southern slopes suggested that there also was some confusion between monocalypts and symphyomyrtles that, based on the confusion matrix, most likely resulted from errors of commission between shaded symphyomyrtle and shaded monocalypt pixels. Shaded pixels are a common source of error in classifications based on imaging spectrometry and for this reason, shaded pixels are often excluded from classification exercises (e.g., Lucas et al., 2008). We included shaded pixels, although we intentionally segregated these pixels into their own class to assess their contribution to model error separately. The “shade-noise” spectra exhibited a high degree of variability and several of these endmembers only modeled a small number of other “shade-noise” spectra, which made it difficult to assess how they might perform in the full image. Had we included more “shade/noise” spectra in our library, we may have observed a higher-degree of model error in some of the “shade-noise” endmember-models, resulting in their exclusion from the final model and/or a failure to meet the minimum threshold for the “higher accuracy” maps.

Otherwise, based on our knowledge of the forest types in the region, the predictions generated by the maps are largely consistent with our expectations of subgenera distribution. There is also a high-degree of consistency between maps, although our endmember-model assessments and field knowledge suggest that the most accurate representation is provided by the “high-accuracy” continuum-removed map (Fig. 7A). The higher-accuracy reflectance map did not perform as well because several of the symphyomyrtle endmember models had accuracies lower than 75% and this resulted in a clear under-representation of the symphyomyrtle species (Fig. 7B).

As expected, the maps show that the majority of eucalypt forests in our study area are comprised of mixed-canopy (monocalypt and symphyomyrtle) stands (Pryor, 1959). However, there are also regions that we know to be dominated by tree species from one subgenera and this provided us with a qualitative indication of overall map accuracy. The monocalypt, *E. macrorhyncha*, dominates a familiar patch in the north (centered at 35°2'7S, 148°29'12E) and this is visible to varying degrees in all four of the maps that we generated from our models. We know that symphyomyrtle species dominate other locations in this northern region, particularly in two *E. viminalis* dominated swaths centered at (35°1'59S, 148°30'1E and 35°2'25S, 148°31'25E) and an area comprised largely of *E. camphora* (centered at 35°3'54S, 148°31'24E). It is also important to consider that *E. viminalis* and *E. dalrympleana* can grow taller than the monocalypt species in our area and this could further increase the apparent dominance of symphyomyrtle canopies relative to the actual proportion of symphyomyrtles and monocalypts on the ground. Predictions over the wetter forest regions around 850 m elevation such as (35°9'47S, 148°38'11E and 35°9'57S, 148°38'52E) correctly indicate a dominance of monocalypts (primarily, *E. radiata*). However,

the high-relief in the valley within the second area (35°9'57S, 148°38'52E) may have contributed to more unmodeled pixels at that location and the under-representation of *E. viminalis* (prevalent along the slopes there as well) due to shaded and mixed pixels.

## 6. Summary

We have demonstrated that it is possible to map the two major *Eucalyptus* subgenera, *Eucalyptus* (“monocalypt”) and *Symphyomyrtus* (“symphyomyrtle”), on a landscape-scale using hyperspectral data collected with a single, airborne sensor (HyMap). This is significant because these subgenera represent important functional types and the traits that are typically shared among members within each group may contribute to their non-random distribution within a landscape (Austin et al., 1983). Mapping monocalypts and symphyomyrtles on a landscape-scale could help to identify habitat for specialist folivores (Moore et al., 2004; Youngentob et al., 2010), provide useful information about the disturbance history of a region (Davidson & Reid, 1980; Duff et al., 1983) and help predict the response of large forest stands to fire (Noble, 1984), fungal pathogens (Stone et al., 1998), variations in nutrient-cycling and salinity (Hawkins & Polglase, 2000; Marcar, 1989; Noble, 1989) and herbivore predation (Stone et al., 1998). The tree species on which we based our classifications are the most abundant species in our study area and are common in several forest-types in NSW and northern Victoria. Future research will focus on investigating the factors that may be responsible for the observed distributions of subgenera within our mapped areas (e.g., the interesting patterns observed in the area centered at 35°4'9S, 148°31'52E) and their relationship to arboreal folivore distributions.

The MESMA approach for classifying and mapping eucalypt subgenera enabled us to capture a larger degree of spectral diversity within our classes than methods that rely on a single endmember to represent each class. We found that continuum-removal analysis improved model accuracy based on reference data, although a slightly higher portion of the pixels remained unmodeled. We used the similarity metrics EAR and MASA and count-based selection (COB) included in the Viper Tools software (Roberts et al., 2007) for our initial endmember-model selections. We then used an iterative process that involved individual endmember-model assessment for model optimization. The procedure that we used to refine our model selection and improve classification accuracies has the potential to be largely automated. This would greatly improve the efficiency of this process because model optimization was the most time-intensive component. Some *a priori* knowledge of tree species in the field was required to collect endmembers for our spectral libraries. The presence of isolated tree-crowns in the paddocks within our flight-lines enabled us to easily match trees in the imagery to trees in the field. However, high-accuracy GPS equipment has been used to identify individual tree-crowns in closed-forests where isolated canopies were not an option (Huber et al., 2008). A limitation to our approach is that the costs associated with collecting hyperspectral data can be prohibitive. It may be possible to use spectral unmixing methods to map eucalypt subgenera with less expensive, coarser-resolution spatial hyperspectral satellite imagery. As spaceborne technology improves and airborne missions become less expensive, our knowledge about the spatial distribution and abundance of forest-types and individual species will increase along with our ability to understand, monitor and conserve natural landscapes. In this article, we have presented a new approach that combines two powerful spectral analysis techniques, MESMA and continuum-removal, to map eucalypt subgenera distributions in a closed-canopy forest with imaging spectrometry.

## Acknowledgements

The authors thank Nicole Coggan, Sarah Ugalde, Jeffery Alexander, Paul Daniel and other volunteers for their assistance in the field. Three

anonymous reviewers provided helpful comments that improved earlier versions of this manuscript. This research was made possible by the generous support of The Hermon Slade Foundation, The Wilderness Society, Ecological Society of Australia, The Fenner School of Environment and Society at the Australian National University and The Australian Commonwealth Scientific and Research Organization (CSIRO) Division of Marine and Atmospheric Research. This research was conducted with the permission of State Forests New South Wales (permit C032438) and National Parks (permit S12036).

## References

- Abilio de Carvalho, O., Jr., & Guimaraes, R. F. (2007). Employment of the multiple endmember spectral mixture analysis (MESMA) method in mineral analysis. *XIII annual SBSR Brazilian remote sensing symposium, Brazil*.
- Adams, J. B., Smith, M. O., & Johnson, P. E. (1986). Spectral mixture modeling: A new analysis of rock and soil types at the Viking Lander 1 site. *Journal of Geophysical Research*, *91*, 8098–8112.
- Adams, J. B., Smith, M. O., & Johnson, P. E. (1993). Imaging spectroscopy: Interpretation based on spectral mixture analysis. In C. M. Pieters, & P. A. J. Englert (Eds.), *Remote geochemical analysis: Elemental and mineralogical composition* (pp. 145–164). Cambridge, England: Cambridge University Press.
- Asner, G. P., & Lobell, D. B. (2000). A biogeophysical approach for automated SWIR unmixing of soils and vegetation. *Remote Sensing of Environment*, *74*, 99–112.
- Asner, G. P., Wessman, C. A., Bateson, C. A., & Privette, J. L. (2000). Impact of tissue, canopy, and landscape factors on the hyperspectral reflectance variability of arid ecosystems. *Remote Sensing of Environment*, *74*, 69–84.
- Austin, M. P., Cunningham, R. B., & Wood, J. T. (1983). The subgeneric composition of Eucalypt forest stands in a region of southeastern Australia. *Australian Journal of Botany*, *31*, 63–71.
- Ballantine, J. -A. C., Okin, G. S., Prentiss, D. E., & Roberts, D. A. (2005). Mapping North African landforms using continental scale unmixing of MODIS imagery. *Remote Sensing of Environment*, *97*, 470–483.
- Bateson, C. A., Asner, G. P., & Wessman, C. A. (2000). Endmember bundles: A new approach to incorporating endmember variability into spectral mixture analysis. *IEEE Transactions on Geoscience and Remote Sensing*, *38*, 1083–1094.
- Bedini, E., van der Meer, F., & van Ruitenbeek, F. (2009). Use of HyMap imaging spectrometer data to map mineralogy in the Rodalquilar caldera, southeast Spain. *International Journal of Remote Sensing*, *30*, 327–348.
- Boardman, J. W. (1998). Post-ATREM polishing of AVIRIS apparent reflectance data using EFFORT: A lesson in accuracy versus precision. *Summaries of the seventh JPL airborne earth science workshop* (pp. 53). Pasadena, California: JPL Publication.
- Boardman, J. W., Kruse, F. A., & Green, R. O. (1995). Mapping target signatures via partial unmixing of AVIRIS data. *Summaries of the 5th JPL airborne earth science workshop* (pp. 23–26). Pasadena, California: JPL Publication.
- Borel, C. C., & Gerstl, S. A. W. (1994). Nonlinear spectral mixing models for vegetative and soil surfaces. *Remote Sensing of Environment*, *47*, 403–416.
- Castro-Esau, K. L., Sánchez-Azofeifa, G. A., & Caelli, T. (2004). Discrimination of lianas and trees with leaf-level hyperspectral data. *Remote Sensing of Environment*, *90*, 353–372.
- Castro-Esau, K. L., Sánchez-Azofeifa, G. A., Rivard, B., Wright, S. J., & Quesada, M. (2006). Inter and intra species variability of spectral reflectance of tropical trees. *American Journal of Botany*, *93*, 517–530.
- Clark, M. L., Roberts, D. A., & Clark, D. B. (2005). Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote Sensing of Environment*, *96*, 375–398.
- Clark, R. N., & Roush, T. L. (1984). Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. *Journal of Geophysical Research*, *89*, 6329–6340.
- Cochrane, M. A. (2000). Using vegetation reflectance variability for species level classification of hyperspectral data. *International Journal of Remote Sensing*, *21*, 2075–2087.
- Cocks, T., Jenssen, R., Stewart, A., Wilson, I., & Shields, T. (1998). The HyMap airborne hyperspectral sensor: the system, calibration and performance. In: *Proceedings of the 1st EARSeL Workshop on Imaging Spectroscopy* (pp. 37–43). Zurich, Switzerland: EARSeL.
- Coops, N. C., Stone, C., Culvenor, D. S., & Chisholm, L. (2004). Assessment of crown condition in eucalypt vegetation by remotely sensed optical indices. *Journal of Environmental Quality*, *33*, 956–964.
- CSES (1992). *Atmosphere REMoval Program (ATREM) user's guide, version 1.1* (pp. 1–24). Boulder, Colorado: Center for the Study of Earth from Space.
- Davidson, J. (1993). Ecological aspects of *Eucalyptus* plantations. In K. White, J. Ball, & M. Kashio (Eds.), *Proceedings regional expert consultation of Eucalyptus*, vol. 1, Bangkok: Food and Agriculture Organization, United Nations.
- Davidson, N. J., & Reid, J. B. (1980). Comparison of the early growth-characteristics of the *Eucalyptus* subgenera *Monocalyptus* and *Symphyomyrtus*. *Australian Journal of Botany*, *28*, 453–461.
- Dennison, P. E., Charoensiri, K., Roberts, D. A., Peterson, S. H., & Green, R. O. (2006). Wildfire temperature and land cover modeling using hyperspectral data. *Remote Sensing of Environment*, *100*, 212–222.
- Dennison, P. E., Halligan, K. Q., & Roberts, D. A. (2004). A comparison of error metrics and constraints for multiple endmember spectral mixture analysis and spectral angle mapper. *Remote Sensing of Environment*, *93*, 359–367.
- Dennison, P. E., & Roberts, D. A. (2003a). The effects of vegetation phenology on endmember selection and species mapping in southern California chaparral. *Remote Sensing of Environment*, *87*, 295–309.
- Dennison, P. E., & Roberts, D. A. (2003b). Endmember selection for multiple endmember spectral mixture analysis using endmember average RMSE. *Remote Sensing of Environment*, *87*, 123–135.
- Dennison, P. E., Roberts, D. A., & Peterson, S. H. (2007). Spectral shape-based temporal compositing algorithms for MODIS surface reflectance data. *Remote Sensing of Environment*, *109*, 510–522.
- Duff, G. A., Reid, J. B., & Jackson, W. D. (1983). The occurrence of mixed stands of the *Eucalyptus* subgenera *Monocalyptus* and *Symphyomyrtus* in south-eastern Tasmania. *Australian Journal of Ecology*, *8*, 405–414.
- Eckmann, T. C., Roberts, D. A., & Still, C. J. (2008). Using multiple endmember spectral mixture analysis to retrieve subpixel fire properties from MODIS. *Remote Sensing of Environment*, *112*, 3773–3783.
- Eschler, B. M., Pass, D. M., Willis, R., & Foley, W. J. (2000). Distribution of foliar formylated phloroglucinol derivatives amongst *Eucalyptus* species. *Biochemical Systematics and Ecology*, *28*, 813–824.
- Filippi, A. M., & Jensen, J. R. (2007). Effect of continuum removal on hyperspectral coastal vegetation classification using a fuzzy learning vector quantizer. *IEEE Transactions on Geoscience and Remote Sensing*, *45*, 1857–1869.
- Gao, B. -C., & Goetz, A. F. H. (1990). Column atmospheric water-vapour and vegetation liquid water retrievals from airborne imaging spectrometer data. *Journal of Geophysical Research, [Atmospheres]*, *95*, 3549–3564.
- Gillespie, A. (1992). Spectral mixture analysis of multispectral thermal infrared images. *Remote Sensing of Environment*, *42*, 137–145.
- Gillespie, T. W., Brock, J., & Wright, C. W. (2004). Prospects for quantifying structure, floristic composition and species richness of tropical forests. *International Journal of Remote Sensing*, *25*, 707–715.
- Gleadow, R. M., Haburjak, J., Dunn, J. E., Conn, M. E., & Conn, E. E. (2008). Frequency and distribution of cyanogenic glycosides in *Eucalyptus* L'Hérit. *Phytochemistry*, *69*, 1870–1874.
- Goodwin, N., Turner, R., & Merton, R. (2005). Classifying *Eucalyptus* forests with high spatial and spectral resolution imagery: An investigation of individual species and vegetation communities. *Australian Journal of Botany*, *53*, 337–345.
- Hawkins, B., & Polglase, P. J. (2000). Foliar concentrations and desorption of nitrogen and phosphorus in 15 species of eucalypts grown under non-limited water and nutrient availability. *Australian Journal of Botany*, *48*, 597–602.
- Hestir, E. L., Khanna, S., Andrew, M. E., Santos, M. J., Viers, J. H., Greenberg, J. A., Rajapakse, S. S., & Ustin, S. L. (2008). Identification of invasive vegetation using hyperspectral remote sensing in the California Delta ecosystem. *Remote Sensing of Environment*, *112*, 4034–4047.
- Hill, J., Attiwill, P. M., Uren, N. C., & O'Brien, N. D. (2001). Does manganese play a role in the distribution of the eucalypts? *Australian Journal of Botany*, *49*, 1–8.
- Huang, Z., Turner, B. J., Dury, S. J., Wallis, I. R., & Foley, W. J. (2004). Estimating foliage nitrogen concentration from HYMAP data using continuum removal analysis. *Remote Sensing of Environment*, *93*, 18–29.
- Huber, S., Kneubühler, M., Psomas, A., Itten, K., & Zimmermann, N. E. (2008). Estimating foliar biochemistry from hyperspectral data in mixed forest canopy. *Forest Ecology and Management*, *256*, 491–501.
- Hughes, L., Cawsey, E. M., & Westoby, M. (1996). Geographic and climatic range sizes of Australian eucalypts and a test of Rapoport's rule. *Global Ecology and Biogeography Letters*, *5*, 128–142.
- Kokaly, R. F., Despain, D. G., Clark, R. N., & Livo, K. E. (2003). Mapping vegetation in Yellowstone National Park using spectral feature analysis of AVIRIS data. *Remote Sensing of Environment*, *84*, 437–456.
- Kokaly, R. F., Rockwell, B. W., Haire, S. L., & King, T. V. V. (2007). Characterization of post-fire surface cover, soils, and burn severity at the Cerro Grande Fire, New Mexico, using hyperspectral and multispectral remote sensing. *Remote Sensing of Environment*, *106*, 305–325.
- Kruse, F. A., & Lefkoff, A. B. (1993). Knowledge-based geologic mapping with imaging spectrometers. *Remote Sensing Reviews*, *8*, 3–28.
- Kruse, F. A., Lefkoff, A. B., Boardman, J. W., Heidebrecht, K. B., Shapiro, A. T., Barloon, P. J., & Goetz, A. F. H. (1993). The spectral image processing system (SIPS)-interactive visualization and analysis of imaging spectrometer data. *The earth and space science information system (ESSIS)* (pp. 192–201). Pasadena, California (USA): AIP.
- Leung, A. K., Chau, F., & Gaob, J. (1998). A review on applications of wavelet transform techniques in chemical analysis. *Chemometrics and Intelligent Laboratory Systems*, *43*, 165–184.
- Li, J. (2004). Wavelet-based feature extraction for improved endmember abundance estimation in linear unmixing of hyperspectral signals. *IEEE Transactions on Geoscience and Remote Sensing*, *42*, 644–649.
- Li, L., & Mustard, J. F. (2003). Highland contamination in lunar mare soils—Improved mapping with multiple end-member spectral mixture analysis (MESMA). *Journal of Geophysical Research*, *108*, 5053–5067.
- Lucas, R., Bunting, P., Paterson, M., & Chisholm, L. (2008). Classification of Australian forest communities using aerial photography, CASI and HyMap data. *Remote Sensing of Environment*, *112*, 2088–2103.
- Majeke, B., van Aardt, J. A. N., & Cho, M. A. (2008). Imaging spectroscopy of foliar biochemistry in forestry environments. *Southern Forests*, *70*, 275–285.
- Marcar, N. E. (1989). Salt tolerance of frost-resistant eucalypts. *New Forests*, *3*, 141–149.
- Monserud, R. A., & Leemans, R. (1992). Comparing global vegetation maps with the Kappa statistic. *Ecological Modelling*, *62*, 275–293.
- Moore, B. D., Wallis, I. R., Marsh, K. J., & Foley, W. J. (2004). The role of nutrition in the conservation of the marsupial folivores of eucalypt forests. In D. Lunney (Ed.),

- Conservation of Australia's forest fauna (pp. 549–575). (2nd ed.). NSW, Australia: Royal Zoological Society of NSW.
- Mutanga, O., & Skidmore, A. K. (2004). Integrating imaging spectroscopy and neural networks to map grass quality in the Kruger National Park, South Africa. *Remote Sensing of Environment*, 90, 104–115.
- Myers, J. B., Theiveyanathan, S., O'Brien, N. D., & Bond, W. J. (1996). Growth and water use of *Eucalyptus grandis* and *Pinus radiata* plantations irrigated with effluent. *Tree Physiology*, 16, 211–219.
- Nix, H. A. (1986). A biogeographic analysis of Australian elapid snakes. In R. Longmore (Ed.), *Atlas of elapid snakes of Australia*. Canberra: Australian Government Publishing Service.
- Noble, I. R. (1984). Mortality of lignotuberous seedlings of *Eucalyptus* species after an intense fire in montane forest. *Austral Ecology*, 9, 47–50.
- Noble, I. R. (1989). Ecological traits of the *Eucalyptus* L'Herit subgenera *Monocalyptus* and *Symphyomyrtus*. *Australian Journal of Botany*, 37, 207–224.
- Painter, T. H., Roberts, D. A., Robert, G. O., & Dozier, J. (1998). The effects of grain size on spectral mixture analysis of snow-covered area from AVIRIS data. *Remote Sensing of Environment*, 65, 320–332.
- Poggiani, F. (1985). Nutrient cycling in *Eucalyptus* and *Pinus* plantations ecosystems, silvicultural implications. *Instituto de Pesquisas e Estudos Florestais*, 31, 33–40.
- Powell, R. L., Roberts, D. A., Dennison, P. E., & Hess, L. L. (2007). Sub-pixel mapping of urban land cover using multiple endmember spectral mixture analysis: Manaus, Brazil. *Remote Sensing of Environment*, 106, 253–267.
- Pryor, L. D. (1959). Species distribution and association in *Eucalyptus*. In A. Keast, R. L. Crocker, & C. S. Christian (Eds.), *Biogeography and ecology in Australia* (pp. 461–471). Junk: The Hague.
- Rashed, T. (2008). Remote sensing of within-class change in urban neighborhood structures. *Computers, Environment and Urban Systems*, 32, 343–354.
- Richards, J. A., & Jia, X. (2006). *Remote sensing digital image analysis: An introduction* (pp. 1–439). (4th ed.). Berlin, Germany: Springer-Verlag.
- Roberts, D. A., Dennison, P. E., Gardner, M., Hetzel, Y., Ustin, S. L., & Lee, C. (2003). Evaluation of the potential of hyperion for fire danger assessment by comparison to the airborne visible/infrared imaging spectrometer. *IEEE Transactions On Geoscience and Remote Sensing*, 41(6 Part 1), 1297–1310.
- Roberts, D. A., Gardner, M., Church, R., Ustin, S. L., Scheer, G., & Green, R. O. (1998). Mapping Chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models. *Remote Sensing of Environment*, 65, 267–279.
- Roberts, D. A., Halligan, K. Q., & Dennison, P. E. (2007). ViperTools. <http://www.vipertools.org/>
- Roberts, D. A., Smith, M. O., & Adams, J. B. (1993). Green vegetation, nonphotosynthetic vegetation, and soils in AVIRIS data. *Remote Sensing of Environment*, 44, 255–269.
- Rouget, M., Richardson, D. M., Nel, J. L., & van Wilgen, B. W. (2002). Commercially important trees as invasive aliens: Towards spatially explicit risk assessment at a national scale. *Biological Invasions*, 4, 397–412.
- Sabol, D. E., Adams, J. B., & Smith, M. O. (1992). Quantitative subpixel spectral detection of targets in multispectral images. *Journal of Geophysical Research*, 97, 2659–2672.
- Schmidt, K. S., & Skidmore, A. K. (2003). Spectral discrimination of vegetation types in a coastal wetland. *Remote Sensing of Environment*, 85, 92–108.
- Settle, J. J., & Drake, N. A. (1993). Linear mixing and the estimation of ground cover proportions. *International Journal of Remote Sensing*, 14, 1159–1177.
- Shimabukuro, Y. E., & Smith, J. A. (1991). The least-squares mixing models to generate fraction images derived from remote sensing multispectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 29, 16–20.
- Shugart, H. H., & Nobel, I. R. (1981). A computer model of succession and fire response of the high-altitude *Eucalyptus* forest of the Brindabella Range, Australian Capitol Territory. *Australian Journal of Ecology*, 6, 149–164.
- Somers, B., Cools, K., Delalieux, S., Stuckens, J., Van der Zande, D., Verstraeten, W. W., & Coppin, P. (2009). Nonlinear hyperspectral mixture analysis for tree cover estimates in orchards. *Remote Sensing of Environment*, 113, 1183–1193.
- Somers, B., Delalieux, S., Stuckens, J., Verstraeten, W. W., & Coppin, P. (2009). A weighted linear spectral mixture analysis approach to address endmember variability in agricultural production systems. *International Journal of Remote Sensing*, 30, 139–147.
- Song, C. (2005). Spectral mixture analysis for subpixel vegetation fractions in the urban environment: How to incorporate endmember variability? *Remote Sensing of Environment*, 95, 248–263.
- Sonnentag, O., Chen, J. M., Roberts, D. A., Talbot, J., Halligan, K. Q., & Govind, A. (2007). Mapping tree and shrub leaf area indices in an ombrotrophic peatland through multiple endmember spectral unmixing. *Remote Sensing of Environment*, 109, 342–360.
- Stone, C., Simpson, J. A., & Gittins, R. (1998). Differential impact of insect herbivores and fungal pathogens on the *Eucalyptus* subgenera *Symphyomyrtus* and *Monocalyptus* and genus *Corymbia*. *Australian Journal of Botany*, 46, 723–734.
- Turner, B. J., Dibley, G., & Dury, S. J. (1998). Hyperspectral characteristics of Australian native eucalypt forest. In: *Proceedings of the 1st EARSeL Workshop on Imaging Spectroscopy* (pp. 317–330). Zurich, Switzerland: EARSeL.
- Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., & Steininger, M. (2003). Remote sensing for biodiversity science and conservation. *Trends in Ecology & Evolution*, 18, 306–314.
- van der Meer, F. (2004). Analysis of spectral absorption features in hyperspectral imagery. *International Journal of Applied Earth Observation and Geoinformation*, 5, 55–68.
- Williams, J. E., & Woinarski, J. (1997). *Eucalypt ecology: Individuals to ecosystems* (pp. 1–430). Cambridge: Cambridge University Press.
- Wu, C. (2004). Normalized spectral mixture analysis for monitoring urban composition using ETM+ imagery. *Remote Sensing of Environment*, 93, 480–492.
- Xiao, Q., Ustin, S. L., & McPherson, E. G. (2004). Using AVIRIS data and multiple-masking techniques to map urban forest tree species. *International Journal of Remote Sensing*, 25, 5637–5654.
- Xie, Y., Sha, Z., & Yu, M. (2008). Remote sensing imagery in vegetation mapping: A review. *Journal of Plant Ecology*, 1, 9–23.
- Youngentob, K., Jia, X., Held, A., Renzullo, L., & Lindenmayer, D. B. (2008). Seeing the forest for the trees: Classifying *Eucalyptus* subgenera using HyMap hyperspectral remote sensing data. *14th annual Australasian remote sensing and photogrammetry conference*. Australia: Darwin.
- Youngentob, K., Wallis, I. R., Lindenmayer, D. B., Wood, J. T., Pope, M., & Foley, W. J. (2010). Foliage chemistry influences tree choice and landscape use of a gliding marsupial folivore. *Journal of Chemical Ecology*. doi:10.1007/s10886-010-9889-9
- Zhang, J., Rivard, B., & Sanchez-Azofeifa, A. (2004). Derivative spectral unmixing of hyperspectral data applied to mixtures of lichen and rock. *IEEE Transactions on Geoscience and Remote Sensing*, 42, 1934–1940.
- Zhang, J., Rivard, B., & Sánchez-Azofeifa, A. (2005). Spectral unmixing of normalized reflectance data for the deconvolution of lichen and rock mixtures. *Remote Sensing of Environment*, 95, 57–66.