DISCRIMINATION OF REMNANT TREE SPECIES AND REGENERATION STAGES IN QUEENSLAND, AUSTRALIA USING HYPERSPECTRAL IMAGERY

Armando Apan[1], Stuart Phinn[2], and Tek Maraseni[1]

[1]Australian Centre for Sustainable Catchments & Faculty of Engineering and Surveying, University of Southern Queensland, Toowoomba QLD 4350, Australia (apana@usq.edu.au)

[2] Centre for Remote Sensing and Spatial Information Science, University of Queensland, Brisbane QLD 4072, Australia (s.phinn@uq.edu.au)

ABSTRACT

This study assessed the utility of hyperspectral imagery in discriminating remnant tree species and stand regeneration stages in Southeast Queensland, Australia. Reflectance data of three species of woody vegetation (i.e. *Eucalyptus populnea*, *Acacia pendula* and *Eucalyptus orgadophila*), acquired using a HyMap™ airborne system, were analysed using partial least squares (PLS) regression. Three groups of *E. orgadophila* species, representing stand regeneration status, were also evaluated. For discriminating such tree species, the PLS results showed high prediction accuracy ranging from 83-88%. The most significant spectral bands span from the visible region (peak at 558nm and 689nm), near-infrared region (peak at 987nm), and shortwave infrared region (peak at 1788nm). Hyperspectral data was able to discriminate the old stand of *E. orgadophila* from the young stand, with a moderate accuracy of 72%. Results such as these confirmed the potential utility of hyperspectral data in vegetation mapping and stand characterisation.

Index Terms— hyperspectral, vegetation, species, regeneration, Australia

1. INTRODUCTION

Remnant vegetation pertains to remaining vegetation after an area has been cleared or modified. Patches of remnant vegetation are important for a wide variety of functions, such as a) habitat of wildlife [1], and b) control of soil erosion and dryland salinity [2]. As their extent and distribution continue to decline due to anthropogenic activities and natural causes, reliable information about remnant vegetation is necessary for effective resource and environmental management. Among the information frequently needed by scientists and managers refers to species richness (e.g. [3]) and regeneration stages (e.g. [4]). Hyperspectral remote sensing offers some potential to map vegetation species and regeneration stages. Previous studies have demonstrated that hyperspectral data and derived spectral vegetation indices (SVIs) can achieve better results in mapping vegetation. For instance, it was found that hyperspectral data from CASI sensor acquired at 1m spatial resolution could provide discrimination of woodland species in the Southern Brigalow Belt, Queensland [5].

The objectives of this study were to: a) examine if selected remnant vegetation species and regeneration stages can be adequately discriminated from airborne measurements of hyperspectral reflectance, b) determine the best spectral bands relevant to discrimination, and c) compare the prediction accuracies and errors produced from using raw hyperspectral data, spectral vegetation indices (SVIs), and other data transformations.

2. RESEARCH METHODS

2.1. Study area and image data acquisition

The study area was located on the property "Well Park" (approximately 27.105°S, 151.345°W), on the Darling Downs region of Southeast Queensland, Australia (Figure 1). It is primarily an agricultural landscape, with vast alluvial plains being planted with cereal crops, such as wheat and sorghum. The riparian areas on the eastern side of the image are dominated by poplar box (*Eucalyptus populnea*), myall (*Acacia pendula*) and river red gum (*Eucalyptus camaldulensis*). On the basaltic upland areas to the west of the image, mountain coolibah (*Eucalyptus orgadophila*) dominates the remnant vegetation communities.

This study acquired an aerial hyperspectral image on 16 April 2004 using the Hyperspectral Mapper (HyMap™) system [6]. The sensor utilised four 32-element detector arrays to provide 126 spectral channel covering the 450nm to 2500nm spectral range over a 512 pixel swath. The spatial

978-1-4244-4687-2/09/$25.00 ©2009 IEEE

resolution achieved at that mission was about 3m, covering a swath width of approximately 2km. The data provider performed radiometric calibration to deliver data in apparent surface reflectance unit.

Figure 1. Location map of the study area

2.2. Field data collection

Three native vegetation species were included in this study: poplar box (*Eucalyptus populnea*), myall (*Acacia pendula)* and mountain coolibah (*Eucalyptus orgadophila*) (Table 2 and Figure 2). A total of 107 samples across the study area was collected using a combination of simple random sampling and purposive selection method. A number of selected tree attributes, i.e. diameter at breast height (dbh), tree height, and canopy diameter at the major and minor axis, were measured correspondingly.

Table 2. Vegetation species included in the study

Common Name	**Scientific Name**	**Attributes**
Mountain Coolibah	*Eucalyptus orgadophila*	Open-woodland species; mature tree can range from 15-20m height and up to 1m dbh; open, straggly crown; leaf size and colour shades vary between juvenile and adult trees; green to greyish green leaves [7]
Poplar box	*Eucalyptus populnea*	Open-woodland species; 8-20m in height; up to 0.8m dbh; medium sized compact crown; glossy green, concolorous; elliptical to ovate, elliptical, broad-lanceolate [7]
Myall	*Acacia pendula*	A species of wattle which grows up to 10 metres in height; pendulous (drooping) in form with grey green narrow phyllodes (modified petioles serving leaf's purpose) which are about 8 cm in length; Yellow ball-shaped flowers appear in spring [8]

The regeneration status ("*young stand*", "*middle-aged stand*" and "*old stand*") of the mountain coolibah species was visually identified on the ground (Table 3), and was confirmed by using the tree variables mentioned above.

2.3 Data processing

The image data processing involved the following main components: a) calculation of spectral vegetation indices (SVI), b) training sample selection, and c) assembling textual data into spreadsheet for regression analysis. For SVI calculation, the ENVI™ 4.3 *Vegetation Index Calculator* module [9] was used to generate 14 vegetation index layers (Table 4).

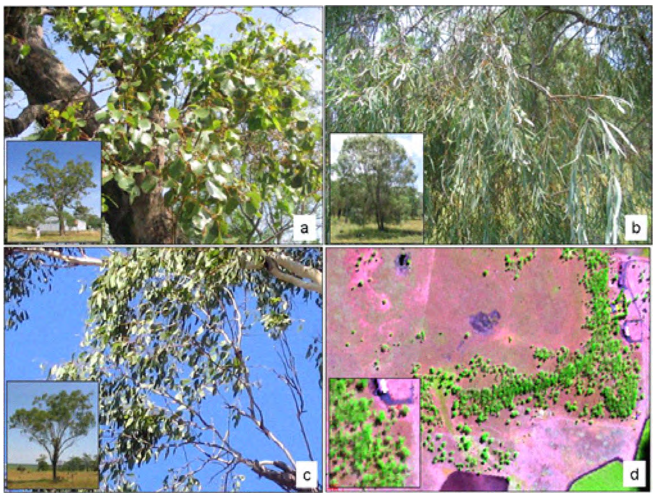

Figure 2. Tree species (a =poplar box; b = myall; and c = coolibah) included in the study and a close up image (d) of the basaltic upland.

Table 3. Regeneration attributes of mountain coolibah trees

Regeneration Stage	**Foliage Attributes**	**Structural Attributes** (Based from Field Data)
Young stand	Alternate, petiolate, ovate to lanceolate; 7-15 x 2-4cm; green to greyish green, becoming concolorous [7]	Ave DBH = 20.9 cm Ave height = 13.9 m Ave crown diameter = 4.4 m
Old stand	Alternate, petiolate, lanceolate to narrow-lanceolate; 8-17 x 0.8-2.5cm; green to greyish green, concolorous [7]	Ave DBH = 80.4 cm Ave height = 19.7 m Ave crown diameter = 17.0 m
Middle-aged stand	Attributes in between the young stand and old stand	Ave DBH = 46.8 cm Ave height = 13.3 m Ave crown diameter = 10.6 m

The selection of training samples for each tree species was started by identifying two "starting pixels" representing the centre of the tree canopy. Then, a "region growing" facility of ENVI software was utilised, employing a threshold of 2 standard deviations away from the mean of the starting pixel value, and by using the 4 neighbouring pixels to determine the growth pattern. The final selection of pixels was determined iteratively, based from this spectral-based technique and the knowledge of the crown canopy diameter measured on-site.

Table 4. Spectral vegetation indices used in this study.

Name
1. Normalized Difference Vegetation Index
2. Simple Ratio Index
3. Atmospherically Resistant Vegetation Index
4. Red Edge Normalized Difference Vegetation Index
5. Vogelmann Red Edge Index 1
6. Red Edge Position Index
7. Photochemical Reflectance Index
8. Red Green Ratio Index
9. Carotenoid Reflectance Index 1
10. Carotenoid Reflectance Index 2
11. Anthocyanin Reflectance Index 1
12. Water Band Index
13. Moisture Stress Index
14. Normalized Difference Infrared Index

2.4 Partial least squares regression

PLS regression is a bilinear modelling method for relating the variations in one or several response variables to the variations of several predictors [11]. Using this technique, information in the original X-data (independent variables) is projected onto a small number of underlying ("latent") variables called PLS components or factors. The PLS method was originally developed as an extension of the more familiar principal component analysis (PCA).

The reflectance data was transformed by three procedures: a) first derivative of raw reflectance spectra b) smoothing by moving average, and c) using the wavelengths or SVIs identified by the Martens' Uncertainty Test [12].

Similar with traditional regression analysis, PLS regression requires that the response ("Y" dependent) variable be a quantitative data. Because of this, the values for species type variable were coded "0" and "1", and consequently limited the analysis to only two classes at a time. A full cross-validation (leave-one-out) technique was applied for this study.

The performance of each PLS model was measured by calculating the root mean square error of prediction (RMSEP). A measure of the average difference between predicted and measured response values, it can be interpreted as the average prediction error, expressed in the same unit as the original response value [10].

3. RESULTS AND DISCUSSION

3.1. Prediction of tree species and regeneration stages

The PLS regression results demonstrated the potential to predict the tree species from HyMap hyperspectral data. Between mountain coolibah and poplar box, there was a high correlation between predicted and measured values for the validated samples, i.e. r=0.90 to 0.97 (Table 5). The root mean square error of prediction (RMSEP) was relatively low (i.e. from 0.12 - 0.21 in a range of 0-1), indicative of the good prediction accuracy (79-88%) of the regression models. Among the three species combinations, it appears that spectral discrimination between myall and poplar box has the least accuracy (75-83%), although this value is still relatively high.

The spectral discrimination between old and young stands achieved a cross-validated maximum accuracy of 72%, with correlations between the predicted and measured values ranging from r=0.78 to 0.83. The result between young and middle-aged stands was basically the same, as it produced an accuracy of 73%. However, the discrimination between middle-aged and old stand is relatively lower in magnitude (highest at 67% accuracy and correlation of r=0.74).

Table 5. PLS regression results of tree species and hyperspectral data

Data	Correlation (r)	RMSEP	Accuracy of Prediction %
A. Coolibah vs. Poplar box (n=72)			
1. Raw Spectra	0.96	0.1303	87
2. First Derivative of Raw Spectra	0.96	0.1303	87
3. Smoothing (Moving Average)	0.96	0.1272	87
4. Smoothed Raw Spectra (Marten's Significant Variables)	0.97	0.1235	**88**
5. Spectral Vegetation Indices (SVI)	0.90	0.2125	79
6. SVI (Marten's Significant Variables)	0.91	0.2046	80
B. Coolibah vs. Myall (n=69)			
1. Raw Spectra	0.93	0.1717	83
2. First Derivative of Raw Spectra	0.91	0.1965	80
3. Smoothing (Moving Average)	0.93	0.1802	82
4. Smoothed Raw Spectra (Marten's Significant Variables)	0.96	0.1434	**86**
5. Spectral Vegetation Indices (SVI)	0.90	0.2119	79
6. SVI (Marten's Significant Variables)	0.90	0.2100	79
C. Myall vs. Poplar box (n=51)			
1. Raw Spectra	0.88	0.2485	75
2. First Derivative of Raw Spectra	0.91	0.2102	79
3. Smoothing (Moving Average)	0.86	0.2545	75
4. Smoothed Raw Spectra (Marten's Significant Variables)	0.94	0.1721	**83**
5. Spectral Vegetation Indices (SVI)	0.93	0.1818	82
6. SVI (Marten's Significant Variables)	0.88	0.2487	75

Correlation (r) is between predicted and measured values; RMSEP – root mean square error of prediction; n = the number of samples

3.2. Comparison of raw, SVIs, and transformed data

For tree species discrimination involving a) mountain coolibah vs. poplar box and b) mountain coolibah vs. myall, the prediction accuracy of raw spectra is relatively higher than the accuracy of spectral vegetation indices (SVIs) (Table 5). These were a) 87% vs. 79%, and b) 83% and 79%, respectively. For myall vs. poplar box, the accuracy of SVIs is higher than raw spectra, i.e. 75% vs. 82%. These varying results indicate that the benefit of using vegetation indices is case-specific.

With regards to regeneration stages, the prediction performance of the raw reflectance spectra was generally similar to that of the vegetation indices. In this case, it can be generalised that the raw spectra and SVI datasets have comparable predictive power for regeneration stages when analysed using PLS regression. This suggests that the transformations applied to vegetation indices did not make a significant difference in model prediction.

Considering the PLS results of species discrimination, the first derivative transformation and the smoothing transformation did not bring significant improvement to the RMSEP and prediction accuracy values. In few cases, there was noticeable reduction in prediction accuracy when these data transformations were applied. However, it was the use of "significant variables" from the Marten's Uncertainty Test that brought a consistent increase in the accuracy for species discrimination. For example, for myall vs. poplar box results, the 75% accuracy of raw spectra has increased to 83% when data from Marten's Uncertainty Test was used.

3.3. Significant spectral bands and indices

Based on the regression coefficient plot and by the results of Martens' Uncertainty Test, the most significant spectral bands for tree species discrimination span from the visible region (peak at 558nm and 689nm), near-infrared region (peak at 987nm), and shortwave infrared region (peak at 1788nm). Those wavebands in the NIR region attained the highest regression coefficient values. With regards to SVIs, the most significant variables were identified, in descending order, as follows: *Red Green Ratio Index (RGRI)*, *Simple Ratio Index*, and *Normalized Difference Infrared Index*.

For regeneration stage prediction, the results show that the most significant spectral bands are those from the visible region (peak at 704nm) and near-infrared region (peak at 1289nm). On the other hand, the identified SVIs to be most significant in the prediction model were *Anthocyanin Reflectance Index*, *Carotenoid Reflectance Index* and *Water Band Index*.

4. CONCLUSIONS

Cross-validated PLS regression models produced high prediction accuracies, confirming the usefulness of narrow-band spectral reflectance data for tree species discrimination. Vegetation indices produced varying accuracy results, indicating that their potential use for species discrimination is case-specific. The data transformation techniques (i.e. first derivative transformation and smoothing) did not bring significant improvement to prediction accuracy. In contrast, the use of "significant variables" identified from the Marten's Uncertainty Test brought a consistent increase in the accuracy of the regression model.

The spectral separability of mountain coolibah trees under different regeneration stages indicated a good separation between the young stand and old stand in all regions of the spectrum, except the far end of the shotwave infrared bands. In addition to the NIR region, the spectral bands in the visible region appeared to exert a key role in the separation of the regeneration stages. This study found that the raw spectra and vegetation indices datasets have comparable predictive power for regeneration stages when analysed using PLS regression.

This study confirmed the potential utility of hyperspectral data to discriminate selected tree species and regeneration stages that can help improve vegetation mapping and stand characterisation. More work is being done to classify the image using per-pixel (e.g. support vector machine) and object-based (image segmentation) approaches.

References

[1] R.E. Major, F.J. Christie, and G. Gowing, G., 2001, "Influence of Remnant and Landscape Attributes on Australian Woodland Bird Communities", *Biological Conservation*, 102 (1), 47-66, 2001.

[2] R.J. Hobbs, "Effects Of Landscape Fragmentation On Ecosystem Processes In The Western Australian Wheatbelt", *Biological Conservation*, 64 (3), 193-201, 1993.

[3] W.M. Block, A.B. Franklin, J.P. Ward, Jr., J.L.Ganey, and G.C. White, "Design and Implementation of Monitoring Studies to Evaluate the Success of Ecological Restoration on Wildlife", *Restoration Ecology*, 9 (3), 293–303, 2001.

[4] N. Gagné, L. Bélanger, and J. Huot, "Comparative Responses of Small Mammals, Vegetation, and Food Sources to Natural Regeneration and Conifer Release Treatments in Boreal Balsam Fir Stands of Quebec", *Canadian Journal of Forest Research*, 29 (7), 1128–1140, 1999.

[5] M. Paterson, R.M. Lucas, and L. Chisholm, "Differentiation of Selected Australian Woodland Species Using CASI Data, *Proceedings of the IEEE 2001 International Geoscience and Remote Sensing Symposium*, University Of New South Wales, Sydney, Australia, 643-645, 2001.

[6] Hyvista Corporation, *Hymap Data Acquisition Report* – May 2004, Hyvista Corporation, North Ryde, Nsw, Australia, 2004.

[7] D.J. Boland, M.I.H. Brooker, G.M. Chippendale, N. Hall, B.P.M. Hyland, R.D. Johnston, D.A. Kleinig & J.D. Turner, 1994, *Forest Trees of Australia*, CSIRO Publishing, Collingwood, Australia, 1994.

[8] J.W. Wrigley & M. Fagg, *Australian Native Plants: Propagation, Cultivation and Use in Landscaping*, 3rd Edition, Angus and Robertson, Sydney, 1988.

[9] ITT Visual Information Solutions, *ENVI 4.3*, ITT Visual Information Solutions, Boulder Co., 2006.

[10] CAMO, *The Unscrambler 9.1*, CAMO Process As, Oslo, 2004.

[11] K. Esbensen, *Multivariate Data Analysis – In Practice*, Camo Process As, Oslo, 2002.

[12] H. Martens and M. Martens, "Modified Jack-Knife Estimation of Parameter Uncertainty in Bilinear Modelling by Partial Least Squares Regression (PLSR)", *Food Quality and Preference*, 11, 5-16, 2000.