

Journal of
Applied Remote Sensing

RemoteSensing.SPIEDigitalLibrary.org

**Spectral discrimination of bulloak
(*Allocasuarina luehmannii*) and
associated woodland for habitat
mapping of the endangered bulloak
jewel butterfly (*Hypochrysops
piceata*) in southern Queensland**

Wan Nor Zanariah Zainol Abdullah
Armando A. Apan
Tek N. Maraseni
Andrew F. Le Brocque

Spectral discrimination of bulloak (*Allocasuarina luehmannii*) and associated woodland for habitat mapping of the endangered bulloak jewel butterfly (*Hypochrysops piceata*) in southern Queensland

Wan Nor Zanariah Zainol Abdullah,^{a,b,c,*} Armando A. Apan,^{a,b}
Tek N. Maraseni,^a and Andrew F. Le Brocque^b

^aUniversity of Southern Queensland, International Centre for Applied Climate Science,
West Street, Toowoomba, Queensland, 4350, Australia

^bUniversity of Southern Queensland, Faculty of Health, Engineering and Science,
West Street, Toowoomba, Queensland, 4350, Australia

^cUniversiti Putra Malaysia, Faculty of Engineering, Department of Biological and Agricultural
Engineering, 43400, Serdang, Selangor, Malaysia

Abstract. The bulloak jewel butterfly (*Hypochrysops piceata*) is an endangered species due to a highly restricted distribution and complex life history, yet little is known of the availability of suitable habitat for future conservation. The aim of this study was to examine the potential of hyperspectral reflectance data for the discrimination of woodland species in support of bulloak jewel butterfly's habitat mapping. Sites from known butterfly sightings in Leyburn, Southern Queensland, Australia, were examined using hyperspectral scanning and vegetation species discrimination. Reflectance data of eight woodland vegetation species (*Allocasuarina luehmannii*, *Eucalyptus crebra*, *Eucalyptus populnea*, *Callitris glauca*, *Corymbia maculata*, *Angophora leicarpa*, *Acacia sparsiflora*, and *Jacksonia scoparia*) were collected at the leaf and canopy levels using a full-range (350 to 2500 nm) hand-held nonimaging spectroradiometer. Partial least square (PLS) regression was used to interpret the bulloak tree spectra against other vegetation species. The PLS results indicated high-prediction accuracies ranging from 78% to 95% and 52% to 5% for canopy and leaf levels, respectively. The highest spectral separability was observed at the near-infrared bands (approximately at 700 to 1355 nm), followed by selected ranges in the short-wave infrared band where separability peaked at 1670 and 2210 nm. The results confirmed the feasible use of hyperspectral sensing for discriminating vegetation species and its potential use for habitat mapping of the endangered bulloak jewel butterfly. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JRS.8.083561](https://doi.org/10.1117/1.JRS.8.083561)]

Keywords: hyperspectral sensing; vegetation species discrimination; partial least square regression; bulloak jewel butterfly.

Paper 14200 received Apr. 9, 2014; revised manuscript received Jul. 16, 2014; accepted for publication Jul. 25, 2014; published online Aug. 25, 2014.

1 Introduction

The assessment and mapping of high biodiversity areas are vital components of conservation planning.^{1,2} The location and extent of wildlife habitat, as well as the spatial distribution of wildlife, are important considerations in developing conservation strategies to mitigate the habitat loss and degradation.³ In assessing and mapping habitat areas, a number of bioclimatic and environmental factors need to be considered at the appropriate level of thematic details and spatial scales.^{4,5} As regards to vegetation, the assessment of its structure and composition presents certain challenges whenever vast tracts of land are involved.⁶ Traditional field survey and mapping

*Address all correspondence to: Wan Nor Zanariah Zainol Abdullah, E-mail: zanariah.zainolabdullah@usq.edu.au

techniques for vegetation classification cannot always provide the required information in an appropriate time and cost-effective manner.^{7,8} Geographic information systems (GIS) and remote sensing technologies offer potential solutions to broad-scale vegetation assessment and mapping of wildlife habitat.^{9–11}

In the context of conservation mapping and management of wildlife, habitat is often a species-specific concept.⁵ This means that the management of species, including their respective habitat, demands that the mapping be conducted at the individual species level.^{12,13} This could be a challenging task, given that the individual species may have very specific habitat requirements which cannot be easily mapped. For instance, the bulloak jewel butterfly (*Hypochrysops piceata*),¹⁴ an endangered species in southern Queensland, has close associations with the bulloak tree (*Allocasuarina luehmannii*)¹⁵ such that the habitat mapping at the tree species or plant-association level is highly preferred. Yet, mapping at the tree species level will be difficult even for the contemporary remotely sensed data captured by multispectral sensors.

Recent development of hyperspectral remote sensing systems, such as the Hyper-X and HypsIRI missions, may provide significantly improved mapping of vegetation at the species or plant association levels.^{16,17} Data from hyperspectral remote sensing technology increase the capability to accurately map vegetation characteristics which were formerly not measurable with broadband multispectral bands.^{18–20} For instance, vegetation discrimination studies identified that the hyperspectral imagery has accurately mapped and differentiated vegetation species.^{18,21} Some studies achieved classification with high accuracies in vegetation species discrimination analysis.^{17,18,21}

Mapping the habitat of the bulloak jewel butterfly (*H. piceata*) using hyperspectral remote sensing data is currently unexplored, with no reported work published in the literature. Identified as endangered by Queensland's Nature Conservation Act 1992 and classified as a high precedence under the Department of Environment and Heritage Protection (EHP), the bulloak jewel butterfly uniquely depends on the species of bulloak tree (*A. luehmannii*). Having a mutual connection with bulloak trees, ant species (*Anonychomyrma* sp.), and also with the attendance of scale insects named *Rhyzococcus* sp.,^{12,13,15,22} this butterfly occurs in bulloak or mixed bulloak woodland on sandy and alluvial soils in southern Queensland.^{14,22}

Remnant bulloak woodland appears to provide important resources not only as habitat for the bulloak jewel butterfly, but also as a food resource for a number of bird species.^{22,23} Bulloak woodland is also categorized as an endangered ecological community.²⁴ Thus, there is an urgent need for habitat mapping of woodlands associated with the bulloak tree. Furthermore, Lundie-Jenkins and Payne¹⁵ highlighted in the *Recovery Plan for the Bulloak Jewel Butterfly (Hypochrysops piceatus) 1999–2003* the crucial need to (a) discover and confirm additional populations of the butterfly and (b) conduct spatial mapping and predictive modeling. These suggestions were initially proposed by Dunn and Kitching¹² based on their previous butterfly surveys.

The goal of this study was to evaluate the possible use of hyperspectral data in discriminating bulloak trees from other woodland vegetation species in a woodland remnant in southern Queensland. The specific objectives were to (a) examine whether bulloak tree species can be effectively differentiated from the selected woodland vegetation species at the leaf and canopy levels from the full-range (350 to 2500 nm) hand-held spectroradiometer data and (b) compare the prediction accuracies and errors of the species discrimination produced from raw and transformed data. This study is part of a broader assessment focused on habitat mapping and suitability modeling of the endangered bulloak jewel butterfly in southern Queensland.

2 Materials and Methods

2.1 Study Area

The study area (33,400 ha) is located near Leyburn, Queensland, a town in the Darling Downs Region (approximately 28.0167°S, 151.5667°E) as depicted in Fig. 1. A small settlement,

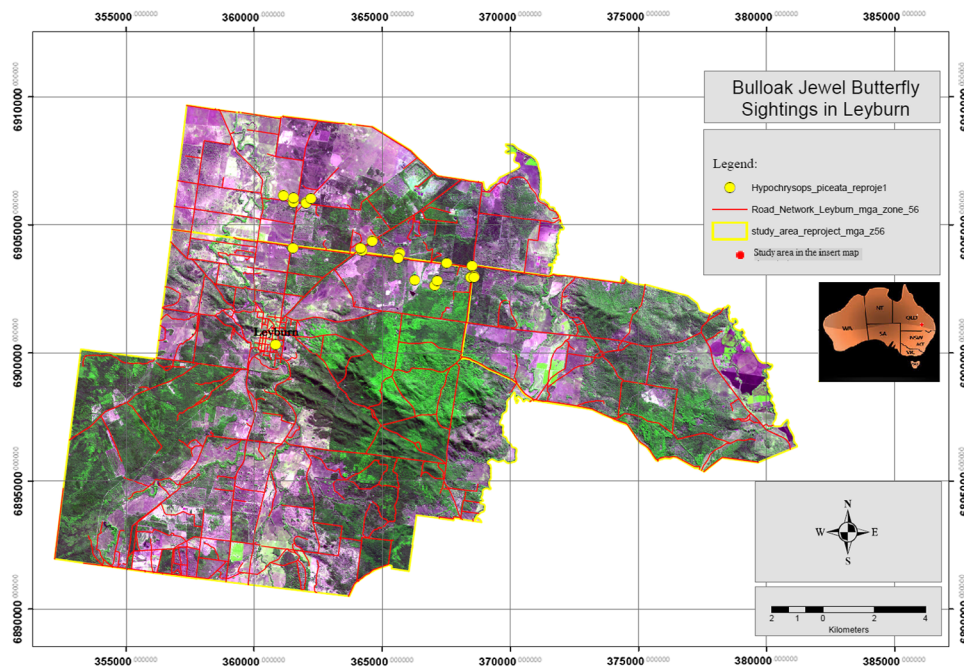


Fig. 1 Location and extent of the study area.

Leyburn is situated 219 km southwest of Brisbane with an altitude of approximately 423 m above mean sea level. The woodland species in this area are generally mixed in composition and occur in multiple stages of regeneration and degradation. The dominant vegetation species in the survey sites included bulloak (*A. luehmannii*), narrow-leaf ironbark (*Eucalyptus crebra*), poplar box (*Eucalyptus populnea*), white cypress pine (*Callitris glauca*), spotted gum (*Corymbia maculata*), apple gum (*Angophora leiocarpa*), currawang (*Acacia sparsiflora*), and dogwood (*Jacksonia scoparia*).

The study area is dominated by livestock grazing, representing approximately 85% of the total area.²⁵ Other land use types include cropping, production forestry, mining, reservoir/dam, residential, intensive animal production, and other minimal uses.²⁵ There are 19 regional ecosystems (REs)²⁶ types in the study area, of which five combinations of these REs (5730 ha in total) are associated with the bulloak trees species.

The study area is a natural habitat of the bulloak jewel butterfly, and this species has been sighted in the locality^{12,14} as marked in Fig. 1. The data collection conducted by the Queensland's EHP²⁶ from 1969 to 2003 resulted in a total of 19 sightings of the bulloak jewel butterfly in this locality. Data collection dates vary from 1969 until 2003. All sightings are very restricted within this area. Hence, it is selected as the study area for this project.

2.2 Field Data Collection

Field data collection was conducted in June 2013 to measure the reflectance of eight woodland vegetation species (e.g., *A. luehmannii*, *E. crebra*, *E. populnea*, *C. glauca*, *C. maculata*, *A. leiocarpa*, *A. sparsiflora*, and *J. scoparia*) separately at the leaf and canopy levels as depicted in Fig. 2 and Table 1. RE maps, produced by Queensland's EHP,²⁶ served as a guide in the purposive selection of sample locations at the plant community level. The target sites at this level were selected based on the presence of bulloak trees, accessibility, and proximity to plant communities where the butterfly was sighted. The information about the target sites was collected during the preliminary survey prior to the spectrometer reading data collection.

Using a simple random sampling method, a total of 40 sample trees were initially considered within the RE sites. However, due to the reasons related to accessibility, height of tree, apparent

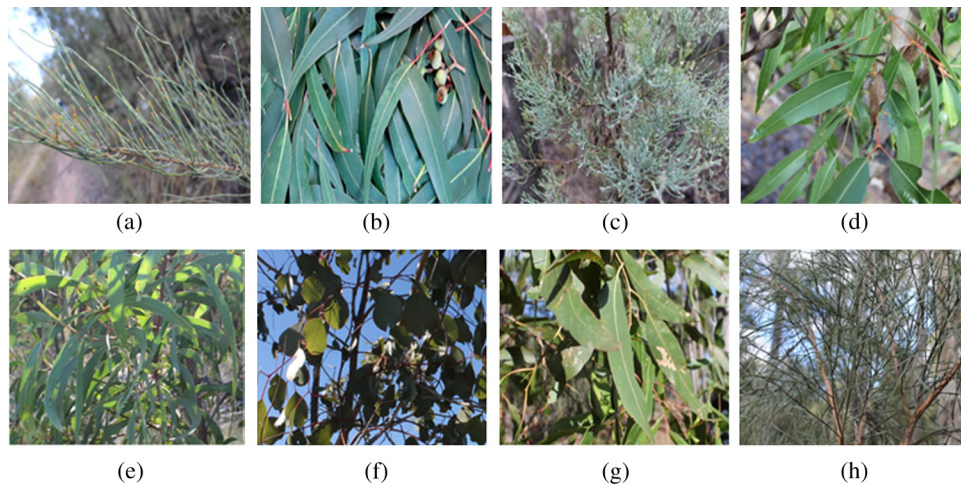


Fig. 2 Woodland vegetation species for species discrimination study (taxonomic designation and details are found in Table 1). (a) Bulloak. (b) Narrow-leaf ironbark. (c) White cypress pine. (d) Apple gum. (e) Currawang. (f) poplar box. (g) Spotted gum. (h) Dogwood.

tree health condition, and exposure to sun, only 32 trees were sampled. Four trees were sampled for each woodland species. For each woodland species, 45 to 60 spectral measurements were collected at the leaf level, with the same amount collected at the canopy level. The identities of the selected species were defined by comparing the physical appearance of leaves, flowers, fruits, and bark with those respective characteristics found in the field guide book.²⁷ They were further verified by consulting with the local experts during the preliminary survey work.

Table 1 List of woodland species for the spectroradiometer readings and brief descriptions of foliage/leaf.²⁴

Item	Common name	Scientific name	Family name	Descriptions of foliage/leaf
a	Bulloak	<i>Allocasuarina luehmannii</i>	Casuarinaceae	Leaves reduced to tiny pointed scales or “leaf teeth” which sheath the branchlets, foliage branchlets rather wiry and erect, up to 2 mm in diameter ²⁴
b	Narrow-leaf ironbark	<i>Eucalyptus crebra</i>	Myrtaceae	Adult leaves disjunct, narrow-lanceolate or lanceolate, 7- to 15-cm long, 0.9- to 1.7-cm wide, green or gray-green, dull, concolorous ²⁴
c	White cypress pine	<i>Calitris columellaris</i>	Cupressaceae	Foliage usually green, rarely glaucous, leaves 1- to 3-mm long, the dorsal surface not keeled ²⁴
d	Apple gum	<i>Angohpora leiocarpa</i>	Myrtaceae	Juvenile leaves ovate or elliptic, to 12.5-cm long, 6.5-cm wide. Adult leaves lanceolate, 9- to 17-cm long, 2- to 3.5-cm wide, apex acute or acuminate ²⁴
e	Currawang	<i>Acacia sparsiflora</i>	Fabaceae	Branchlets angled at extremities, glabrous or with scattered hairs (hairy on juvenile plants), narrowed at ends, 8- to 16-cm long, 5- to 10-mm wide, thinly coriaceous, gray-green ²⁴
f	Poplar box	<i>Eucalyptus populnea</i>	Myrtaceae	Juvenile leaves disjunct, ovate to orbiculate, dull gray-green. Adult leaves disjunct, broad-lanceolate, elliptic, ovate or rhomboidal, 5- to 11-cm long, 2- to 7-cm wide, green, glossy, concolorous ²⁴
g	Spotted gum	<i>Eucalyptus maculata</i>	Myrtaceae	Juvenile leaves disjunct, elliptic to ovate, glossy green. Adult leaves disjunct, lanceolate, 10- to 21-cm long, 1.5- to 3-cm wide, green, concolorous ²⁴
h	Dogwood	<i>Jacksonia scoparia</i>	Fabaceae	Branches and branchlets erect or pendulous, strongly angled or winged, leaves usually reduced to scales, occasionally a few leaves present at base of branches or on young plants, obovate, petiolate ²⁴

A portable full-range (350 to 2500 nm) spectroradiometer from the Analytical Spectral Devices, Inc., Boulder, Colorado,²⁸ was used to record the reflectance data at 1-nm wavelength intervals. This device is a hand-held battery-powered spectrometer with a fore optic cable for light collection and a notebook computer for data logging.²⁶ For canopy-level measurement, the sensor was positioned about 70 cm above the canopy to record the average spectra within a 10-cm diameter on the canopy. A twin-step ladder was used to measure the canopy reflectance. On the other hand, for leaf-level measurement, a distance of 17-cm height was used to record the average spectra for a 2.5-cm diameter of leaf. A precalibrated yard stick was used to measure the distance between the sensor and the leaves. These calculations were based on the 8-deg field of view (FOV) toward the targeted sample. The FOV is used to express the solid angle through which light incident on the fore optic will arrive to the detector system.²⁸

Reliable field spectrometry data collection depends upon accurate calibration of the devices used. Thus, the sensor was calibrated using a white reference plate prior to the scanning for both levels of measurement. The scanning was carried out on the unshadowed portion of the leaf or canopy on a clear sunny day. For canopy measurement, the most challenging task was to choose a part of the canopy with minimal shadow coverage from other parts of the same tree or from adjoining trees. Thus, the best practice to cater for this issue was through a selection of an exposed (no incident shadow) part of the canopy or a tree that has a suitable distance from other adjacent trees.

Reflectance data were collected between 10.00 am and 2.30 pm local time for optimal performance of the sun's azimuth and elevation. Moreover, additional calibration was performed against a white reference plate if there was cloud interference during the recording session. This helps in calibrating ongoing differences between multiple sources and standardizing the measurement.

2.3 Data Preprocessing

For both levels of measurement, 10 spectra were internally averaged by the spectrometer for each sample. All spectral datasets were stored in a computer and processed using the RS3 software intended for use with a graphical user interface. The reflectance data were then transformed into ASCII format. The sample datasets from the ASCII format constituted 16 raw files of reflectance data for the eight selected vegetation species sampled at the leaf and canopy levels. These datasets were structured in spreadsheets to produce the data array of 45 to 60 samples for every species by 2151 wavebands. Furthermore, bulloak species was compared with the other species independently, since there was a need to discriminate between bulloak and other vegetation species (e.g., bulloak versus apple gum; bulloak versus white cypress pine; and bulloak versus narrow-leaf ironbark). Thus, the variables for the seven pairs of vegetation species were labeled properly before exporting into the Unscrambler 9.2²⁹ software for partial least squares (PLSs) regression analysis.

Before the calibration stage was undertaken, the spectral reflectance was preprocessed for an optimal performance. A series of "cleaning" operations was applied for the elimination of: (a) very short wavelengths (350 to 399 nm) and strong water vapor absorption bands: 1356 to 1480, 1791 to 2021, and 2396 to 2500 nm,^{18,21} (b) outliers that indicated abnormal reflectance response as compared with other samples.^{20,30,31}

Pretreatment or transformation of the spectral data was a significant component of a number of spectral analyses to improve the accuracy of results. In this study, two chemometrics pretreatment methods [moving average smoothing^{8,18,20} and multiplicative scatter correction (MSC)^{20,31}] were applied to reduce the noise and to normalize the data. The goals of smoothing or the averaging method were to decrease the number of variables in the dataset, to eliminate uncertainty in measurement, and to reduce the effect of noise.³² This method replaces each point of the spectra with the average of x adjacent points, where x is a positive integer in a matrix.³² In this case, a 3 by 3 matrix or a kernel window of moving average was selected to transform the raw bands prior to the PLS regression analysis. On the other hand, the purpose of MSC is to treat undesired scatter effects for both multiplicative (amplification) and additive (offset) effects.³²

2.4 PLS Regression

The most frequently used regression techniques in spectroscopy analysis are principal component analysis (PCA) and PLS regression.^{30,31,33} The PLS model is comparatively better than the PCA because it does not include latent variables that are less important to describe the variance of the quality measurement.^{20,30,31,34} The PLS regression can be defined as a bilinear modeling method for relating the variations in one or numerous response variables (Y -variable) to numerous predictors (X -variable).^{33,35} In PLS regression, information of the independent (X -variable) variables is projected onto a small number of latent (Y -variable) variables named PLS factors or components.^{30,35} This meant that the Y -data are used in estimating the latent variables to ensure that the first components are those that are most relevant for predicting the Y -variables.

The PLS comprises regression and classification tasks as well as reduction techniques and modeling tools.^{31,33} Furthermore, PLS reduces the entire reflectance spectra to a few relevant factors and regresses them to the measured parameter of a given sample.^{36,37} Thus, the PLS model is considered to be more robust than a multiple linear regression calibration model.³⁶

This method performs well when the various “ X -variables” articulate common information, i.e., when there is a great amount of correlation or collinearity. Thus, the use of PLS regression in discrimination analysis offers better capabilities than traditional regression techniques in analyzing hyperspectral data because of its inherent high collinearity.³⁸ Other statistical approaches (e.g., discriminant function analysis) are not applicable for analysis involving over 1000 X -variables. These alternatives have limitations in handling multicollinearity of predictor variables.

Furthermore, PLS requires the response of a dependent variable Y (species) which is a categorical data to quantitatively analyze the data similar to the other traditional regression technique. In this situation, the values for the species-type variable were coded as “0” and “1” [e.g., bulloak (0) versus apple gum (1)]. Thus, in this study, it limited the analysis to only two species at a time with the intention of exploring an interspecies difference.

The dataset was divided into two sets which were comprised of calibration and validation sets for PLS regression analysis. Seventy-five percent of the dataset was used to develop a prediction equation (calibration set), while the remaining 25% was used for the validation of the predictive equation.^{20,38} In the development of the PLS model, a full cross-validation (leave-one-out) method was used to calculate the quality of prediction and to prevent over-fitting of the calibration model.^{32,35} This means that only one sample at a time is kept out of the calibration. Furthermore, the performance of the PLS models was evaluated by the root-mean-square error of prediction (RMSEP) and the coefficient of determination (r^2) of the model. An accurate PLS model should have a high regression coefficient of determination (r^2) and a low RMSEP between the predicted and measured values of each regression analysis.^{20,31,35}

Regression analysis was led by identification and elimination of outliers using an outlier list tool offered by Unscrambler 9.2 software. The outliers were also identified from the influence plot which displays the sample residual-variance against leverages. It means that the outliers were listed under the PLS components (PCs) list as signified by the influence plot. The selection of an outlier always starts with the ones that appear first in the earliest components.²⁹ Furthermore, samples with high-residual variance are likely to be outliers. Thus, if there were six samples identified as outliers in the raw bands of interspecies discrimination, then all of them will be removed prior to the PLS regression analysis. This approach was also applied for the smoothing and MSC techniques with the same number of outliers encountered as in the raw bands in order to compare these three techniques. It was found that less than 5% of the dataset was considered outliers.

3 Result and Discussion

3.1 Vegetation Reflectance Properties

Leaves of woodland species display the typical spectral curves of green healthy vegetation: high reflectance in the near-infrared (NIR) and relatively low reflectance in the visible and

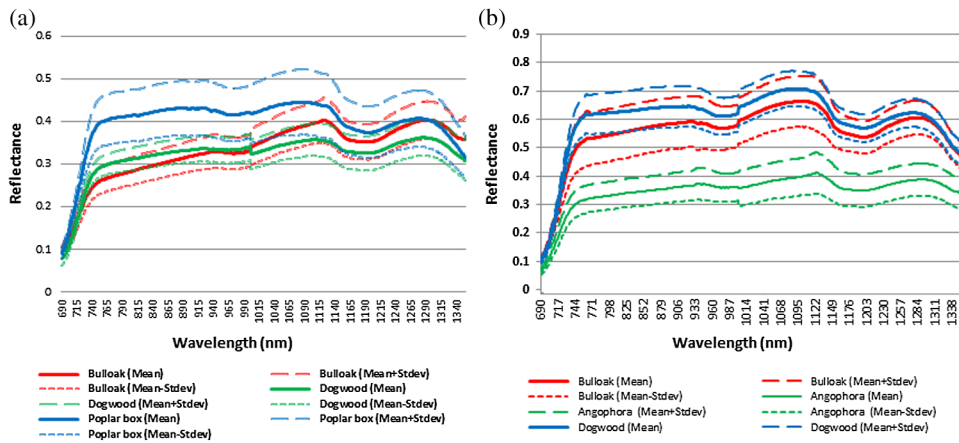


Fig. 3 Mean reflectance spectra at near-infrared (NIR) band (700 to 1355 nm) of three vegetation species at (a) canopy and (b) leaf levels.

short-wave infrared (SWIR). Visual interpretation of raw spectra from Figs. 3–5 indicated the differences in ranges across vegetation species. For example, bulloak species was displayed and compared with dogwood (the closest spectral response) and poplar box (highly separated spectral response).

At both leaf and canopy levels, the highest separability of the spectral magnitude was clearly shown at the NIR band (approximately at 700 to 1355 nm) as depicted in Fig. 3. These were followed by the selected ranges in the SWIR band (where separability peaked at 1670 and 2210 nm), green band (550 nm), and red band (650 nm). This result agreed with other researchers^{19,34,35} who concluded that the spectral difference between the species is insignificant in the visible band but is truly notable in the NIR and SWIR bands. It means that the differences in pigment (chlorophyll) absorption between species pairs were not the discriminating variables, but those related to leaf internal structure (NIR-related) and leaf water content and other biochemicals (SWIR-related) were discriminating variables.^{19,34,39}

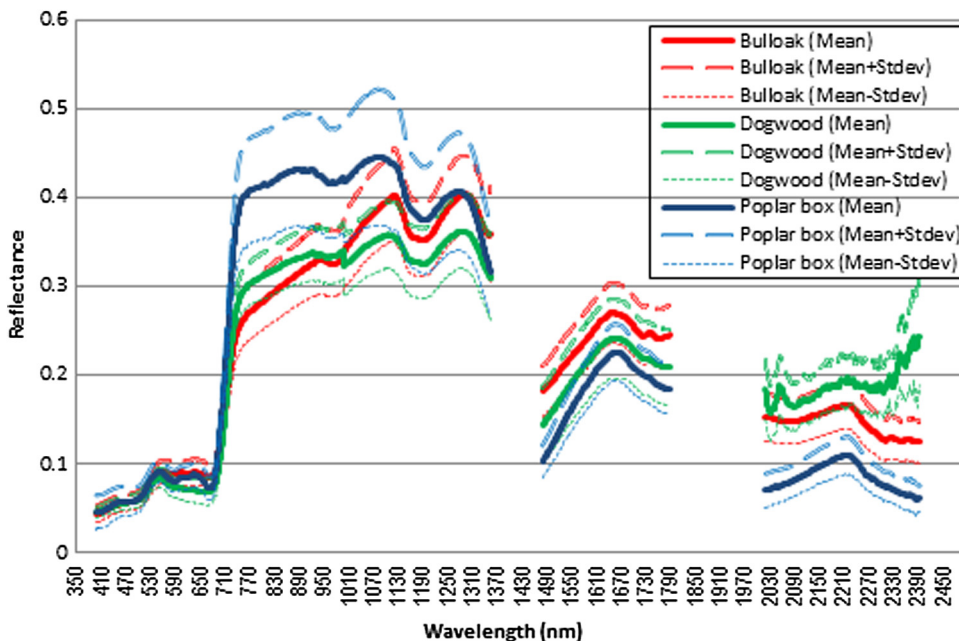


Fig. 4 Mean reflectance spectra (350 to 2500 nm) of three vegetation species at canopy level.

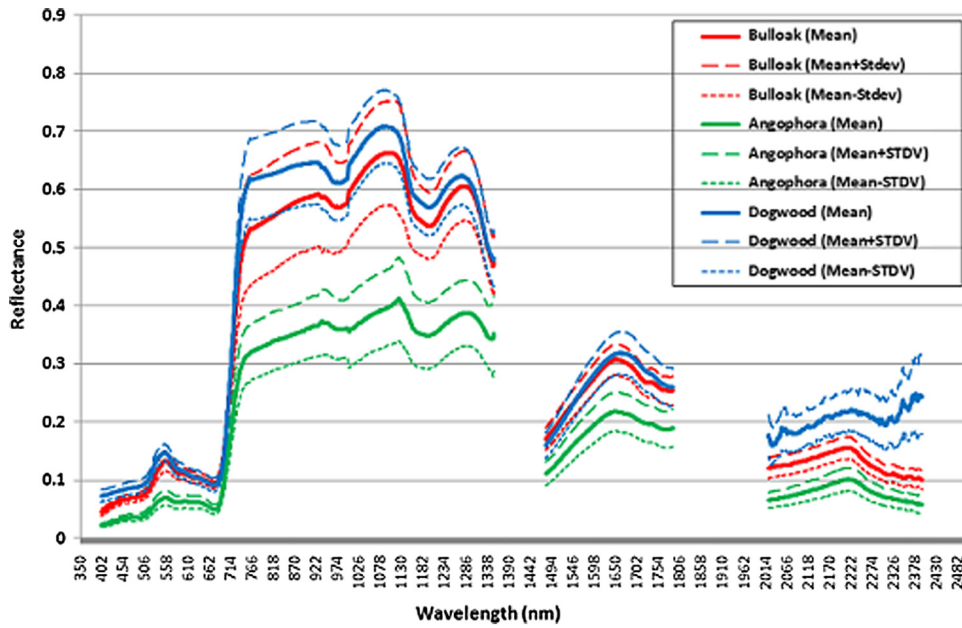


Fig. 5 Mean reflectance spectra (350 to 2500 nm) of three vegetation species at leaf level.

At the canopy level, the bulloak-dogwood species comparison had shown the lowest separability, whereas the bulloak-poplar box pair exhibited the highest spectral discrimination (Fig. 4). Likewise, score plots in PLS demonstrated these patterns as shown in Figs. 6 and 7. These plots clearly show that the bulloak-poplar box pair has high discrimination. Similarly, at the leaf level, bulloak-dogwood species revealed the lowest spectral separability between them, while bulloak-angophora exhibited the highest separability.

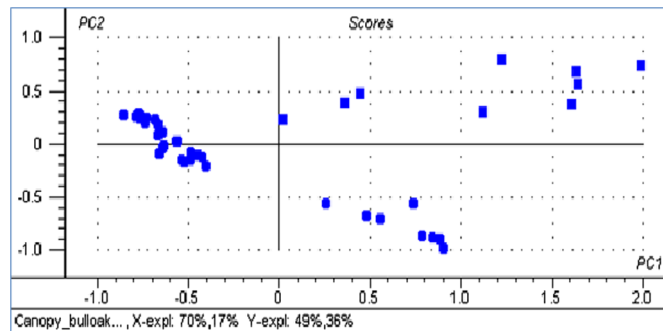


Fig. 6 Score plot showing discrimination between bulloak (rounded square) and dogwood (square) at the canopy level.

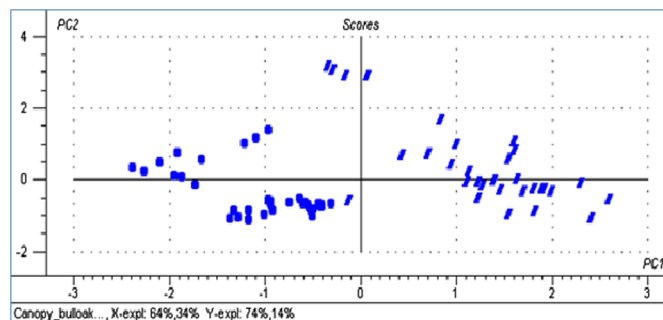


Fig. 7 Score plot showing discrimination between bulloak (rounded square) and poplar box (rectangle) at the canopy level.

The spectral separability of different species pairs, such as bulloak-dogwood and bulloak-poplar box, can be explained by their leaf attributes (e.g., color, size, texture, shape, etc.) as described in Table 1. For instance, the needle-like modified leaves or foliage of the bulloak tree closely resemble that of the dogwood. In comparison, the glossy and rounded dark green leaves of an adult polar box tree are distinctly different from bulloak. These examples also reflect the same situation in describing the spectral separability of the bulloak-angophora species pair: bulloak's foliage is needle-like, whereas angophora has oppositely arranged lanceolate leaves. These findings are in agreement with Zhang et al.¹⁹ and Taylor et al.²¹ who demonstrated that the variations due to leaf attributes, along with other biophysical properties, leaf chemical composition, and leaf water content, were the main factors which could contribute to the differences in spectral separability of plant foliage.

There are two significant issues for every study involving the discrimination and mapping of vegetation: (a) the variability of biological types and (b) the spectral similarity (characteristics) of most vegetation.⁴⁰ Some of this variability can be accredited to geographical and environmental backgrounds, some to phenology and seasonal circumstances, but most of them still remain at the scale of communities and individual vegetation in their habitat.⁴¹ However, the spectral properties of all types of vegetation are suppressed by the similar set of pigments, structures, and biochemicals throughout the year.^{40,42} In this study area, vegetation communities are nondeciduous and exhibit little morphological changes.

3.2 Prediction Accuracies of PLS Regression

3.2.1 Bulloak tree versus other woodland species at leaf and canopy levels

Prediction accuracies of PLS results can be obtained from the subtraction of value one with the RMSEP. Tables 2 and 3 summarize the PLS results for the canopy and leaf levels of raw spectra. There were high correlations between predicted and measured values for the validated samples, i.e., $r = 0.985$ to 0.997 and $r = 0.985$ to 0.996 for canopy and leaf levels, respectively. The RMSEP was reasonably low (i.e., from 0.0433 to 0.1086 for canopy level and from 0.0542 to 0.0827 for leaf level) in range of 0 to 1, indicating good prediction accuracies.

Figure 8 reveals the accuracies of raw spectra at the canopy level range from 89.14% to 95.67%, while the accuracies at the leaf level range from 91.73% to 94.58%. Among the seven pairs of species combination, the lowest accuracy pairs correspond to bulloak versus narrow-leaf ironbark (89.14%) and bulloak versus apple gum (91.73%) for canopy and leaf levels, respectively. These values were found to be relatively high. This means that the accuracy ranges obtained in this study were relatively close to those attained by Huang and Apan³¹ (93.57% to 94.27%) who used PLS regression for detecting *Sclerotinia* rot disease on celery.

Table 2 Results of PLS model from raw spectra at canopy level.

Data	Correlation (r)	RMSEP	Accuracy of prediction (%)
Bulloak versus apple gum	0.994	0.0733	92.67
Bulloak versus currawang	0.993	0.0802	91.99
Bulloak versus cypress	0.997	0.0433	95.67
Bulloak versus dogwood	0.997	0.0878	91.22
Bulloak versus narrow-leaf ironbark	0.985	0.1086	89.14
Bulloak versus poplar box	0.996	0.0685	93.15
Bulloak versus spotted gum	0.996	0.0622	93.78

r = correlation between predicted and measured values.

RMSEP = root-mean-square error of prediction.

Table 3 Results of PLS model from raw spectra at leaf level.

Data	Correlation (r)	RMSEP	Accuracy of prediction (%)
Bulloak versus apple gum	0.985	0.0827	91.73
Bulloak versus currawang	0.995	0.0542	94.58
Bulloak versus cypress	0.988	0.0783	92.17
Bulloak versus dogwood	0.991	0.0755	92.45
Bulloak versus narrow-leaf ironbark	0.996	0.0550	94.50
Bulloak versus poplar box	0.991	0.0693	93.07
Bulloak versus spotted gum	0.993	0.0639	93.61

r = correlation between predicted and measured values.

RMSEP = root-mean-square error of prediction.

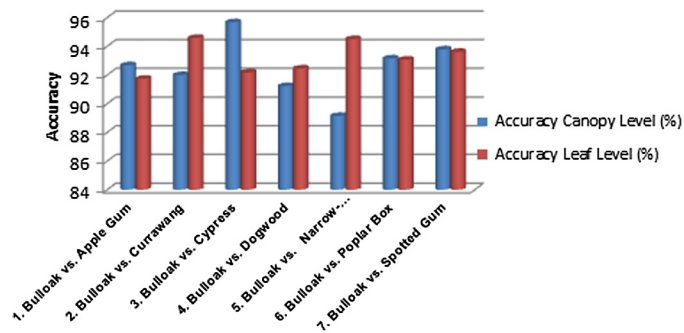


Fig. 8 Partial least square (PLS) accuracies of raw spectra at canopy and leaf levels.

In most cases of PLS regression analysis of raw spectra as conducted here, the prediction accuracies of canopy-level samples produced higher accuracies than the leaf-level sets, except for the bulloak versus currawang, bulloak versus dogwood, and bulloak versus narrow-leaf ironbark (Fig. 8). These higher accuracies of canopy samples may be best explained by the differentiating reflectance of mixed materials integrated within the wider FOV of the sensor. From the point of view of satellite radiation, a forest or woodland environment will reflect an integrated signal from leaves, branches, and trunks of trees as well as from soil and leaf litters.^{21,38} Each of these elements has substrative and additive effects on the spectral curve, thus the ultimate reflectance will be a mixture of all of them.¹⁸

3.2.2 Raw spectra versus transformed data

For vegetation species discrimination involving bulloak tree spectra against the other woodland vegetation species, the prediction accuracy of raw spectra is reasonably higher than the accuracy of transformed spectra (smoothing and MSC) as depicted in Tables 4 and 5 for both canopy and leaf levels. Overall, the prediction accuracies of raw spectra and smoothing were nearly the same, whereas MSC's prediction accuracies consistently produced the lowest accuracies. For example, prediction accuracies of raw spectra and the smoothing method for bulloak versus apple gum, as well as for bulloak versus cypress (canopy level), were nearly the same. They were 92.67% (raw) and 92.64% (smoothing) for bulloak versus apple gum, while they were 95.67% (raw) and 95.67% (smoothing) for bulloak versus cypress. At the leaf level, similar patterns were observed: for bulloak versus dogwood, the accuracies were 92.45%, 92.54%, and 81.82% for raw spectra, smoothing, and MSC, respectively.

Table 4 Results of PLS model from raw and transformed data at canopy level.

Data	Correlation (<i>r</i>)	RMSEP	Accuracy of prediction (%)
A. Bulloak versus apple gum			
1. Raw spectra	0.994	0.073349	92.67
2. Smoothing	0.994	0.073603	92.64
3. MSC	0.903	0.238930	79.17
B. Bulloak versus currawang			
1. Raw spectra	0.993	0.080193	91.99
2. Smoothing	0.993	0.081650	91.84
3. MSC	0.928	0.208762	79.12
C. Bulloak versus cypress			
1. Raw spectra	0.997	0.043289	95.67
2. Smoothing	0.997	0.043308	95.67
3. MSC	0.970	0.140115	85.99
D. Bulloak versus dogwood			
1. Raw spectra	0.997	0.087771	91.22
2. Smoothing	0.997	0.082857	91.71
3. MSC	0.923	0.213178	78.68
E. Bulloak versus narrow-leaf ironbark			
1. Raw spectra	0.985	0.108632	89.14
2. Smoothing	0.985	0.106842	89.32
3. MSC	0.947	0.193268	80.67
F. Bulloak versus poplar box			
1. Raw spectra	0.996	0.068451	93.15
2. Smoothing	0.996	0.069287	93.07
3. MSC	0.983	0.103679	89.63
G. Bulloak versus spotted gum			
1. Raw spectra	0.996	0.062182	93.78
2. Smoothing	0.996	0.062745	93.73
3. MSC	0.971	0.138749	86.13

r = correlation between predicted and measured values.

RMSEP = root-mean-square error of prediction.

This situation indicated that the raw and smoothing spectra have equivalent predictive power for species discrimination when evaluated by PLS regression.

In most cases, the PLS regression accuracies at the canopy level produced higher values than the leaf level either for raw spectra or transformed data. However, a reverse observation was found for the bulloak versus currawang, bulloak versus dogwood, and bulloak versus narrow-leaf ironbark as shown in Fig. 9.

Table 5 Results of PLS model from raw spectra and transformed data at leaf level.

Data (leaf level)	Correlation (r)	RMSEP	Accuracy of prediction (%)
A. Bulloak versus apple gum			
1. Raw spectra	0.985	0.082697	91.73
2. Smoothing	0.985	0.082770	91.72
3. MSC	0.845	0.260188	73.98
B. Bulloak versus currawang			
2. Raw spectra	0.995	0.054240	94.58
3. Smoothing	0.995	0.054335	94.57
4. MSC	0.8405	0.272395	72.76
C. Bulloak versus cypress			
1. Raw spectra	0.9885	0.078305	92.17
2. Smoothing	0.989	0.078605	92.14
3. MSC	0.178	0.479132	52.09
D. Bulloak versus dogwood			
1. Raw spectra	0.991	0.075504	92.45
2. Smoothing	0.991	0.074623	92.54
3. MSC	0.959	0.181769	81.82
E. Bulloak versus narrow-leaf ironbark			
1. Raw spectra	0.996	0.054975	94.50
2. Smoothing	0.996	0.055004	94.50
3. MSC	0.710	0.357798	64.22
F. Bulloak versus poplar box			
1. Raw spectra	0.991	0.069275	93.07
2. Smoothing	0.991	0.069220	93.08
3. MSC	0.689	0.355788	64.42
G. Bulloak versus spotted gum			
1. Raw spectra	0.993	0.063883	93.61
2. Smoothing	0.993	0.063838	93.62
3. MSC	0.732	0.322934	67.71

r = correlation between predicted and measured values.

RMSEP = root-mean-square error of prediction.

Figure 9 reveals that the accuracies of raw spectra and transformed data at the canopy level range from 78.68% to 95.67%, while accuracies at the leaf level range from 52.09% to 94.58%. The least accurate pairs correspond to bulloak versus dogwood (78.68%) and bulloak versus cypress (52.09%) at canopy and leaf levels, respectively. Both values were generated from the MSC's prediction method. This suggests that the transformations executed to the raw spectra did not generate a substantial modification in model prediction.

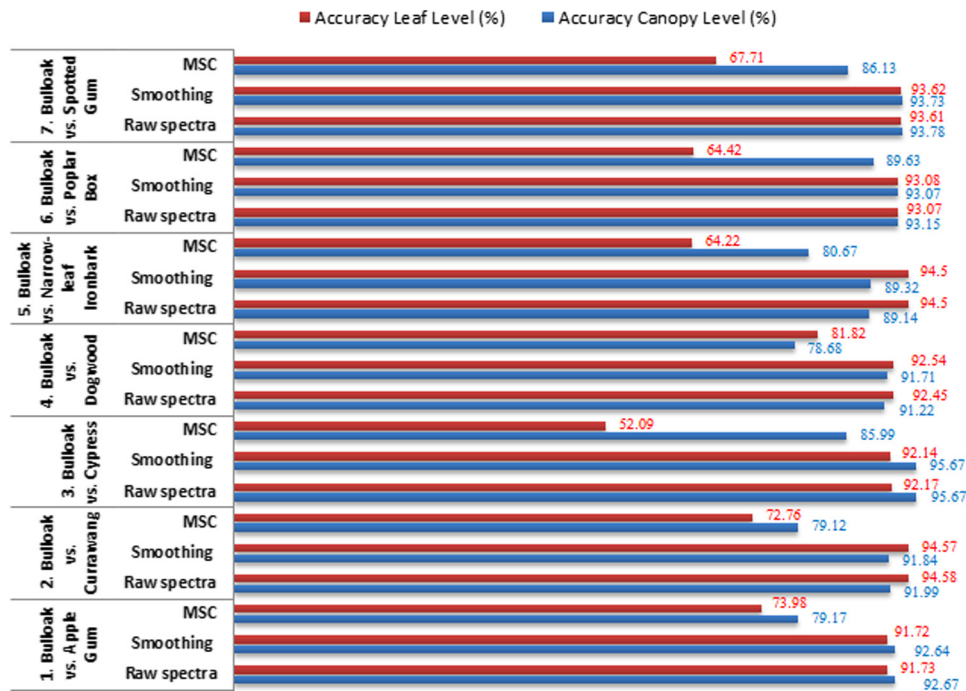


Fig. 9 Accuracies of raw and transformed data at canopy and leaf levels.

4 Conclusions

The results from PLS regression confirm the effectiveness of narrow-band spectral reflectance data for vegetation species discrimination sampled at the study area. Among the seven pairs of species combinations of raw spectra, the least discrimination was observed between bulloak versus narrow-leaf ironbark (89.14%) and bulloak versus apple gum (91.73%) for canopy and leaf levels, respectively. In comparison with the raw spectra and transformed data, the least accurate pairs correspond to bulloak versus dogwood (78.68%) and bulloak versus cypress (52.09%) at canopy and leaf levels, correspondingly. These values were produced by the MSC's prediction method.

This study revealed that the raw spectra and smoothing (transformed) datasets have a corresponding predictive power for discriminating between species at the canopy and leaf levels. However, the transformation techniques of MSC applied to the raw data did not produce significant enhancement to the accuracy of prediction. We conclude that a full cross-validation technique in PLS regression produced high-prediction accuracies for raw spectra and smoothing datasets.

The spectral separability of bulloak tree against other woodland vegetation species indicated good discrimination between selected regions of the spectrum. The NIR region (700 to 1355 nm) appeared to play a key role in the discrimination between species in PLS regression. However, there is a limitation for using PLS for this kind of study as it confines the analysis to only two species at a time. In addition, studies on interspecies spectral differences have been largely descriptive or quantitative. Discrimination was possible but was statistically based, acknowledging that there is variability within a plant group or species. There were relative differences between spectra rather than an individual signature for each species. Variability can be attributed to phenology, geographical, and environmental settings, but much variation still remains at the scale of communities and individual species in their natural habitat.

This study demonstrated the feasible use of hyperspectral data in discriminating woodland species to help improve the mapping of bulloak jewel Butterfly habitat. The next step is to conduct research on the use of hyperspectral imaging sensor (rather than the nonimaging technique as conducted here) for habitat mapping.

Acknowledgments

The authors would like to thank the University of Southern Queensland (USQ) for providing the research facilities, and the Ministry of Higher Education, Malaysia, and the University Putra Malaysia (UPM) for the financial support to Wan Nor Zanariah Zainol Abdullah. Mohd Nor Saifuddin Jamaludin also provided assistance in the field for which we are grateful.

References

1. S. C. Trombulak and R. F. Baldwin, *Landscape-Scale Conservation Planning*, Springer, Netherlands, New York (2010).
2. J. Nackoney and D. Williams, "Conservation prioritization and planning with limited wildlife data in a Congo Basin forest landscape: assessing human threats and vulnerability to land use change," *J. Conserv. Plann.* **8**, 25–44 (2012).
3. J. R. Mawdsley, R. O'Malley, and D. S. Ojima, "A review of climate-change adaptation strategies for wildlife management and biodiversity conservation [Una Revisión de las Estrategias de Adaptación al Cambio Climático para el Manejo de Vida Silvestre y Conservación de la Biodiversidad]," *Conserv. Biol.* **23**(5), 1080–1089 (2009).
4. A. O. Onojeghuo and G. A. Blackburn, "Mapping reedbed habitats using texture-based classification of QuickBird imagery," *Int. J. Remote Sens.* **32**(23), 8121–8138 (2011).
5. S. Weiers et al., "Mapping and indicator approaches for the assessment of habitats at different scales using remote sensing and GIS methods," *Landscape Urban Plann.* **67**(1–4), 43–65 (2004).
6. P. Andrzej and M. Pelechaty, "Spatial structure of vegetation in a small charophyte dominated lake," *Biodiversity Res. Conserv.* **29**(1), 97–104 (2013).
7. Y. Dian et al., "Comparison of the different classifiers in vegetation species discrimination using hyperspectral reflectance data," *J. Indian Soc. Remote Sens.* **42**(1), 1–12 (2013).
8. Y. Xie, Z. Sha, and M. Yu, "Remote sensing imagery in vegetation mapping: a review," *J. Plant Ecol.* **1**(1), 9–23 (2008).
9. B. A. Bradley et al., "Species detection vs. habitat suitability: are we biasing habitat suitability models with remotely sensed data?," *Ecol. Modell.* **244**, 57–64 (2012).
10. D. S. Boyd and G. M. Foody, "An overview of recent remote sensing and GIS based research in ecological informatics," *Ecol. Inf.* **6**(1), 25–36 (2011).
11. J. Thorner, L. Kumar, and S. D. A. Smith, "Fine-scale three-dimensional habitat mapping as a biodiversity conservation tool for intertidal rocky reefs," *J. Coastal Res.* **29**(5), 1184–1190 (2013).
12. K. L. Dunn and R. L. Kitching, "Distribution, status and management of the *piceatus jewel* butterfly on the Darling Downs, Queensland," A report to the Queensland Department of Environment and Heritage, Conservation Strategy Branch (1994).
13. P. Greenslade, "Leyburn's bull oak roadside: the only current locality for Australia's rarest butterfly, the Darling Downs jewel," nomination for the register of the National Estate to the Australian Heritage Commission (1995).
14. M. F. Braby, *The Butterfly of Australia: Their Identification, Biology, and Distribution*, CSIRO Publishing, Victoria, Australia (2000).
15. G. Lundy-Jenkins and A. Payne, *Recovery Plan for the Bulloak Jewel Butterfly (Hypochrysops piceata) 1999–2003*, Queensland Parks and Wildlife Service, Brisbane (2000).
16. A. Brosinsky et al., "Analysis of spectral vegetation signal characteristics as a function of soil moisture conditions using hyperspectral remote sensing," *J. Indian Soc. Remote Sens.* **42**(2), 311–324 (2014).
17. F. Tsai, E. K. Lin, and K. Yoshino, "Spectrally segmented principal component analysis of hyperspectral imagery for mapping invasive plant species," *Int. J. Remote Sens.* **28**(5), 1023–1039 (2007).
18. L. Kumar, A. K. Skidmore, and O. Mutanga, "Leaf level experiments to discriminate between eucalyptus species using high spectral resolution reflectance data: use of derivatives, ratios and vegetation indices," *Geocarto Int.* **25**(4), 327–344 (2010).
19. C. Zhang et al., "Spectral response to varying levels of leaf pigments collected from a degraded mangrove forest," *J. Appl. Remote Sens.* **6**(1), 063501 (2012).

20. N. M. Nawi et al., "Prediction and classification of sugar content of sugarcane based on skin scanning using visible and shortwave near infrared," *Biosyst. Eng.* **115**(2), 154–161 (2013).
21. S. Taylor et al., "Optimal band selection from hyperspectral data for *Lantana camara* discrimination," *Int. J. Remote Sens.* **33**(17), 5418–5437 (2012).
22. T. R. New and D. P. A. Sands, "Management of threatened insect species in Australia, with particular reference to butterflies," *Aust. J. Entomol.* **43**(3), 258–270 (2004).
23. M. Maron and A. Lill, "Discrimination among potential buloke (*Allocasuarina luehmannii*) feeding trees by the endangered south-eastern red-tailed black-cockatoo (*Calyptorhynchus banksii graptogyne*)," *Wildl. Res.* **31**(3), 311–317 (2004).
24. "PlantNET—the plant information network system of the royal botanic gardens and domain trust," <http://plantnet.rbg.syd.nsw.gov.au/> (10 February 2014).
25. "Mapping and data," Department of Natural Resources and Mines, 25 November 2013, <http://www.dnrm.qld.gov.au/mapping-data/maps> (03 March 2014).
26. "Regional ecosystem maps, vegetation maps, PMAVS and regional ecosystem data" Department of Environment and Heritage Protection, 19 February 2014, <http://www.ehp.qld.gov.au/ecosystems/biodiversity/regional-ecosystems/maps/index.php> (03 March 2014).
27. Santos, *Field Guide to Trees and Shrubs of Eastern Queensland Oil and Gas Fields*, Santos Ltd., Adelaide, South Australia (2000).
28. Analytical Spectral Devices (ASD), *Spectroradiometer Users Guide*, ASD, Boulder, CO (2008).
29. CAMO Inc., *The Unscrambler 9.2*, CAMO Inc., Corvallis, OR (2004).
30. G. K. Naganathan et al., "Partial least squares analysis of near-infrared hyperspectral images for beef tenderness prediction," *Sens. Instrum. Food Qual. Saf.* **2**(3), 178–188 (2008).
31. J. F. Huang and A. Apan, "Detection of *Sclerotinia* rot disease on celery using hyperspectral data and partial least squares regression," *J. Spat. Sci.* **51**(2), 129–142 (2006).
32. K. H. Esbensen, *Multivariate Data Analysis—In Practise*, 5th ed., CAMO Inc., Corvallis, OR (2002).
33. M. A. Cho et al., "Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression," *Int. J. Appl. Earth* **9**(4), 414–424 (2007).
34. R. Darvishzadeh et al., "Estimation of vegetation LAI from hyperspectral reflectance data: effects of soil type and plant architecture," *Int. J. Appl. Earth Obs. Geoinf.*, **10**(3), 358–373 (2008).
35. N. M. Nawi, G. Chen, and T. Jensen, "Visible and shortwave near infrared spectroscopy for predicting sugar content of sugarcane based on a cross sectional scanning method," *J. Near Infrared Spectrosc.* **21**(4), 289–297 (2013).
36. T. Selige, J. Böhner, and U. Schmidhalter, "High resolution topsoil mapping using hyperspectral image and field data in multivariate regression modeling procedures," *Geoderma* **136**(1–2), 235–244 (2006).
37. C. Gomez, R. A. V. Rossel, and A. B. McBratney, "Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: an Australian case study," *Geoderma* **146**(3–4), 403–411 (2008).
38. N. M. Nawi, T. Jensen, and G. Chen, "The application of spectroscopic methods to predict sugarcane quality based on stalk cross-sectional scanning," *Am. Soc. Sugar Cane Technol. J.* **32**, 16–27 (2012).
39. E. Adam, O. Mutanga, and D. Rugege, "Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review," *Wetlands Ecol. Manage.* **18**(3), 281–296 (2010).
40. M. A. Cochrane, "Using vegetation reflectance variability for species level classification of hyperspectral data," *Int. J. Remote Sens.* **21**(10), 2075–2087 (2000).
41. M. M. Lewis, "Discriminating vegetation with hyperspectral imagery—what is possible?," in *Proc. Int. Geoscience and Remote Sensing Symposium*, Vol. 6, pp. 2899–2901, IEEE, Sydney, Australia (2001).
42. J. C. Price, "How unique are spectral signature?" *Remote Sens. Environ.* **49**(3), 181–186 (1994).

Wan Nor Zanariah Zainol Abdullah is a postgraduate student at the University of Southern Queensland, Australia, as well as a tutor at the University Putra Malaysia, Malaysia. She received her BS degree from University of Technology Malaysia, Malaysia, in 2006. She was awarded her MS degrees in remote sensing and GIS from the University Putra Malaysia, Malaysia, in 2010. Currently, she is doing her PhD degree in fine-scale habitat mapping and modeling of an endangered butterfly species and climate change impact analysis.

Armando A. Apan is currently an associate professor at the University of Southern Queensland. His current research area focuses on the application of GIS and remote sensing to observe terrestrial ecosystems and their responses to environmental change; vegetation and habitat mapping; species distribution modelling; land use/cover change analysis; and monitoring of agricultural crops. He has over 120 papers published in international refereed journals, conference proceedings and book chapters.

Tek N. Maraseni has over 20 years work experience in carbon and forest related areas in Nepal, Thailand, China, and Australia. Currently, he is the VC's senior research fellow with the University of Southern Queensland, Australia, and also a senior international scientist with the Chinese Academy of Sciences, China. He has produced over 90 publications, including two books in the last seven years. His published work has been recognized by several national and international fellowships/grants/awards.

Andrew F. Le Brocque is a vegetation ecologist and has researched and published in the areas of plant-environment relationships, biodiversity and ecosystem services in agricultural landscapes, and plant-fungal relationships for over 20 years. He is currently senior lecturer in ecology and sustainability at the University of Southern Queensland and has taught ecology, conservation and sustainability at USQ since 1996 and currently coordinates and teaches courses in ecology, conservation biology, environmental science, and sustainability.