Agron. Sustain. Dev. 27 (2007) 111–117

DOI: 10.1051/agro:2006034

Original article

Advanced satellite imagery to classify sugarcane crop characteristics

Y.L. EVERINGHAM[a]*, K.H. LOWE[b], D.A. DONALD[a], D.H. COOMANS[a], J. MARKLEY[c]

[a] Center for Science and Technology, Mitretek Systems, Fairfax, Virginia, 22042, USA, formerly James Cook University
[b] School of Maths, Physics and IT, James Cook University, Townsville, Queensland, 4811, Australia
[c] Mackay Sugar Co-Operative, Mackay, Queensland, 4740, Australia

(Accepted 7 December 2006)

Abstract – Techniques that provide a rapid and widespread assessment of crop properties equip industry decision makers with knowledge to improve their farming environment, both tactically and strategically. An interdisciplinary approach that links the fields of hyperspectral remote sensing, statistical data mining and sugarcane systems was undertaken to establish new relationships to determine variety type and crop age of sugarcane plants. In contrast to commonly used sensors such as those occupied by Landsat satellites, images captured by hyperspectral sensors can provide a more detailed assessment of crop status. Appropriate statistical analysis methods are needed to decode the multifaceted information recorded in these hyperspectral images. A range of statistical approaches have been applied for analysis of an EO-1 hyperion hyperspectral image from a major sugarcane growing region in Australia. Two relatively new classification methods – support vector machines and random forests - demonstrated superior performance in classifying sugarcane variety and crop cycle, e.g. the number of times that the plant has grown back after harvest, when compared against traditional statistical methods. Assignment results were further enhanced when classifications of pixels within sugarcane paddocks were aggregated to paddock classifications using paddock boundary information. Whilst the analysis methods of the hyperspectral data have been tested for the classification of variety and crop cycle, the potential application arenas for this type of imagery is both extensive and relatively unexplored. This type of data coupled with appropriate analysis methods will play a vital role in futuristic sustainable agriculture practices as this imagery becomes more accessible and as land managers and researchers become more aware of the types of decisions that hyperspectral remote sensing data can aid.

precision farming / linear discriminant analysis / penalised discriminant analysis / random forests / support vector machines / hyperspectral

1. INTRODUCTION

Satellite imagery can provide detailed spatial information about an agro-ecological environment. The amount of detail contained in the imagery is dependent on the attributes of the satellite. Advancements in satellite imagery have provided a mechanism for extracting complex information about crop characteristics across a spatial domain. Previously, this level of information could only be obtained from laboratory analyses for point specific locations. To fully exploit the information captured by these modern satellite apparatus, statistical analysis methods are needed to extract relevant information embedded within the more volumuous image data, so that agricultural managers can capitalise on this knowledge to improve sustainable agronomic practices.

Satellite remote sensing instruments can measure and record as a digital image, the intensity of the reflected electromagnetic radiation from agricultural vegetation like sugarcane. The digital image is made up of many smaller discrete picture elements called pixels. Remote sensing instruments are characterized by several attributes that influence the quality of information recorded for each pixel in an image. Besides the spatial resolution, another important attribute is the spectral resolution. The spectral resolution of a satellite is determined by the number of bands (narrow portion of the electromagnetic spectrum) and the spectral location and widths of these bands. Low resolution sensors are constrained by the small number of bands that measure reflected radiation. Hyperspectral sensors however, are high resolution because they measure reflected radiation from a large number of bands to provide a more continuous like spectral signature. Embedded within the spectral signature is detailed quantitative and qualitative information about the medium being examined e.g. crop vegetation (Lillesand and Kiefer, 2000; Thenkabail et al., 2002). Typically, analysis of spectral like data requires more sophisticated statistical approaches. To explore the potential of hyperspectral sensors and methods for analysing hyperspectral data, we consider an example relevant to the Australian sugar industry.

Sugar is Australia's second largest agricultural export commodity that contributes between 1 to 2 billion Australian dollars to the nation's economy. Sugarcane starts as a plant crop and when mature is harvested (Fig. 1). The plant is then allowed to regrow (or ratoon) for harvesting approximately 12 months later, depending on the region. This *cycle* typically happens 4 to 5 times before the crop is completely ploughed out and replanted as a new plant crop. A large proportion of sugar produced by the sugarcane is exported to countries overseas. Therefore, advance knowledge of crop size is needed by marketers to improve forward selling strategies of the commodity. Crop size is heavily dependent on variety and crop

* Corresponding author: yvette.everingham@jcu.edu.au

Figure 1. A picture of sugarcane variety 136 grown in the Mackay region.

cycle. The ability to better predict these characteristics offers the potential to improve yield forecasting models. It is also important to recognise that some varieties are more susceptible to disease. The ability to accurately classify varieties can help enormously to reduce negative impacts associated with disease spread.

The application of hyperspectral imagery in sugar industries world wide is limited. The only use of hyperspectral sensor data for the Australian sugarcane industry was conducted by Apan et al. (2003) who used discriminant function analysis to monitor the disruption of a sugarcane disease called orange rust. For the Brazilian sugarcane industry Galvão et al. (2005, 2006) reported success in using the E0-1 hyperion to classify sugarcane varieties. Our investigation has extended the application of classifying sugarcane varieties using the E0-1 hyperion to the Australian continent and to identify if this imagery can also be used to classify the crop cycle (i.e. number of times the crop has ratooned or grown back after harvesting). Variety and crop cycle are classified at two levels. The first level of classification is at a pixel level and the second level is at the paddock scale. Since several pixels will occupy a sugarcane paddock where crop variety and crop cycle are typically uniform, the classification approach is applied at a paddock level.

Spectral signatures, like those obtained from hyperspectral imagery can be quite similar for different sugarcane varieties and crop cycles. Thus, modern statistical data mining techniques offer the potential to better discriminate spectral characteristics. The term "high (low) dimensional classifiers" is typically reserved for classifiers suited to datasets with highly correlated variables where the ratio of observations to variables is small (large). The ill-posed problem caused by limited observations relative to the number of variables has been referred to as the Hughes phenomena (Hughes, 1968) and is frequently encountered with spectral data. To overcome the low observation-to-variable ratio and strong correlations, traditional methods that involve maximum likelihood and linear discriminant analysis methods have tended to use vegetation indices and feature extraction steps such as a stepwise band selection, principal components and wavelet transforms (Hsu and Tseng, 1999) to incorporate a small number of influential features into the model. The performance of low and high dimensional statistical classifiers for the hyperspectral imagery was considered in this investigation. These included – support vector machines, random forests, penalised discriminant analysis and linear discriminant analysis.

Linear discriminant analysis is more recognised as a low dimensional classifier in contrast to penalised discriminant analysis which is considered a high dimensional classifier. Applications of linear and penalised discriminant analysis methods have been applied to hyperspectral imagery with varying degrees of success (see for example Yu et al., 1999 and Gong et al., 1998). Applications of random forests to hyperspectral data are however more limited. Random forests (Breiman, 2001) is modern method capable of classifying (or regressing) objects from data sets with a relatively large number of variables. The random forest method does not use all the variables at the same time. Integrated within this technique is an "in-house" variable selection method. Compared to random forests, literature describing applications of support vector machines for classifying hyperspectral data is more abundant. Support vector machines are a high dimensional classifier. Compared to traditional classifiers, support vector machines offer a distinct advantage for classifying hyperspectral data of different land cover types (Gualtieri and Chettri, 2000; Camps-Valls et al., 2003, 2004; Huang et al., 2002; Lennon et al., 2002; Mercier and Lennon 2003; Shah et al., 2003; Wilson et al., 2004). The success of this method can be attributed to the ability of support vector machines to suppress information in the spectral signature that is not useful (noisy or redundant) for discriminatory purposes (Camps-Valls et al., 2004). The next section of this paper explains these statistical methods in more detail.

2. STATISTICAL BACKGROUND

2.1. Fisher's Linear Discriminant Analysis (LDA)

Fisher's linear discriminant analysis (Fisher, 1936) is a classical statistical approach for classifying samples whose group membership is unknown, based on training samples from g known classes (supervised learning). Fisher's linear discriminant analysis searches for a linear combination(s)

$$z = a_1x_1 + a_2x_2 + ... + a_px_p$$

of the p independent variables $x_1, x_2, ..., x_p$, that maximise the ratio of the between-group sum-of-squares to the within-groups sum-of squares in the discriminant space. The weights in the linear combination(s) are the eigenvectors of $W^{-1}B$, where W and B are the within- and between-groups covariance matrices, respectively. The eigenvector associated with largest eigenvalue has the largest discriminatory power, the eigenvector associated with the next largest eigenvalue, produces the next largest disciminant criterion and is uncorrelated with the first discriminant function, and so on. The maximum number of discriminant functions is equal to $\min(g\text{-}1, p)$. A new observation is assigned to the group that it lies nearest to in the

discriminant space. Refer to Johnson and Wichern (1988), for a more comprehensive account of Fisher's linear discriminant analysis.

2.2. Penalised Discriminant Analysis (PDA)

Penalised discriminant analysis (Hastie et al., 1995) is similar to Fisher's linear discriminant analysis in that linear combinations of predictor variables are produced that best discriminate the different classes. Penalised discriminant analysis differs from Fisher's linear discriminant method by attempting to overcome problems associated with high dimensional data. These include the high correlation between variables and the low observation to variable ratio which both contribute to instabilities in estimating W^{-1}. Hastie et al., (1995) added the penalty term (Ω) to the within-group covariance matrix W to improve the performance of Fisher's linear discriminant analysis by reducing instabilities associated with calculating W^{-1}. Thus, W is replaced by $W_{pda} = W + \Omega$ in penalised discriminant analysis.

2.3. Random Forests (RF)

Random forests (Breiman, 2001) extend upon the principles of CART – classification and regression trees (Breiman et al., 1984). A classification tree recursively partitions the data into homogenous subsets. Terminating nodes that contain observational units are produced when a rule that ceases further partitioning has been activated. The class of the terminal node is determined by a majority vote of class membership of the observations within that terminal node. To improve this approach, the random forest method incorporates bagging (bootstrap aggregating of randomly sampled cases) and random variable selection to improve predictive accuracy (Breiman, 2001). The random forest method can perform well when the number of variables is much larger than the number of samples (Breiman, 2001), a frequent occurrence in hyperspectral data. One output provided by the random forest method is a variable importance list. The random forest method can provide information about the influence that variables have on model performance (Truong et al., 2004). This can be helpful if a researcher is interested in gaining a better understanding about which variables, or parts of the spectral signature are most useful for classification. Our study was primarily interested in raising awareness of the potential hyperspectral imagery and outlining different methods for analysing this type of data, consequently, the variable importance list was not consulted.

2.4. Support Vector Machine (SVM)

The support vector machine is a statistical learning method for solving supervised classification and nonlinear regression problems. Like Fisher's linear discriminant analysis, penalised discriminant analysis and random forests, support vector machines require no assumption about the population distribution from which the sampled data were drawn. Support vector machines seek an optimal separating surface between groups of data. When two classes ($g = 2$) are present, support vector machines partition the two groups using a linear separating hyperplane. The optimal hyperplane is located where the margin between the two classes is maximised and the error is minimized. When linear hyperplanes cannot be found to separate the classes without misclassifications, support vector machines can be generalised to calculate nonlinear decision surfaces. For multiple class classification problems a binary classifier is constructed for each pair of classes. The appropriate class is then found by a voting scheme within the system. For more details about support vector machines the reader is referred to Boser et al. (1992), Burgess (1998), Bruzzone and Melgani (2002) and Vapnik (1998).

2.5. Accuracy Assessment

There exists many ways that the performance of different models can be validated. These techniques range from the use of independent training and testing data to cross-validation techniques (Kohavi, 1995). In this paper, the v-fold cross-validation method with $v = 10$ has been used to assess the classification accuracy of each statistical method. In v-fold cross-validation, the data are divided into v subsets of approximately equal size. The model is trained and fitted v times, each time leaving out one of the subsets from training. Only the omitted subset is predicted to derive a correct classification rate (CCR). Ten-fold ($v = 10$) cross-validation has been shown to have a better combination of low bias and low variance over leave-one-out cross-validation (Kohavi, 1995). Ten-fold cross-validation offers advantages when small data sets do not permit separate training and testing data and when the data consists of highly unbalanced classes (both cases in this research).

3. DATA AND PRE-PROCESSING PROCEDURES

The hyperspectral data was acquired from the NASA Earth Observing 1 (EO-1) satellite's Hyperion imaging spectrometer (NASA, 2003). The Hyperion sensor has a spatial resolution of 30 m^2 and a spectral resolution of 400 nm to 2500 nm in 242 near continuous 10 nm bandwidths. The study area represents a portion of the Mackay sugarcane growing region in Queensland, Australia. The image, approximately centered on 149° 04' E and 21° 15' S was captured on April 2nd, 2002 contains some 400 000 pixels covering approximately 7.5 km × 45.6 km. The image was originally captured for a study on discriminating "orange rust" disease in sugarcane from hyperspectral data (Apan et al., 2004). This research makes use of the same processed image from Apan et al. (2004). The 242 bands were reduced to 150 bands by eliminating bands that displayed non-uniformities after formal remote sensing calibration procedures were implemented. The 150 bands were given variable names of $x_1, x_2, ..., x_{150}$.

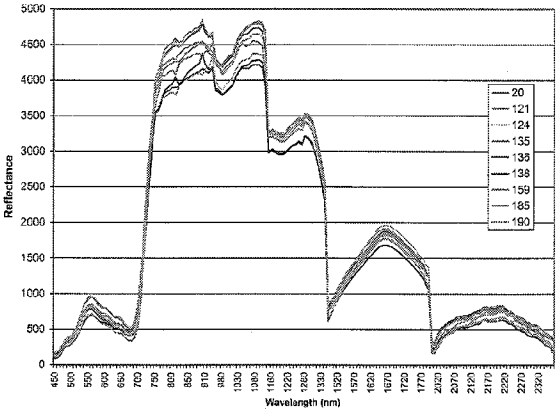

Figure 2. Spectral signatures of nine different sugarcane varieties.

Table I. Distribution of the 2402 pixels and 84 paddocks by variety and crop cycle (ratoon).

Variety			Cycle		
	Paddocks	pixels		Paddocks	pixels
20	5	101	1	25	637
121	2	36	2	20	615
124	35	889	3	11	248
135	13	322	4	5	98
136	10	578	5	12	345
138	7	133	6	5	163
159	1	13	7	1	36
185	7	185	8	3	209
190	4	145	9	2	51

Nine varieties of sugarcane across 84 paddocks have been recorded. These 84 paddocks contained 2402 pixels. Figure 2 shows the spectral graphs of the different sugarcane varieties (randomly selected pixel observations) and the associated band numbers. The varieties are referred to as variety – 20, 121, 124, 135, 136, 138, 159, 185 and 190. Figure 1 shows variety 136 sugarcane stalks grown in the Mackay region. Variety 20 represents mixed varieties. This occurs for example when growers gradually introduce new varieties released by plant breeding authorities into their paddocks. Additional information about the crop cycle had been recorded for the 84 paddocks. The crop cycles have been labeled with values of 1 to 9 (1 = plant cane, i.e. cane that was planted for the first time following the previous harvest, 2 = 1st ratoon, 3 = 2nd ratoon, ..., 9 = 8th ratoon). Table I details the distribution of pixels and paddocks for variety and crop cycle.

Hyperion pixel spectral data were linked with GIS (geographical information system) vector data for each paddock in the farm survey. A mask was created for each paddock that deleted pixels that extended 5 m or more past paddock boundaries. This was done to reduce and/or remove edge effects. Paddocks were then selected if they contained 10 or more pixels per paddock. The decision of 10 or more pixels was an iterative process based on classification accuracies.

Table II. Overall correct classification rates for the (a) 2402 pixels and (b) 84 paddocks.

(a)

Method	Per pixel classification accuracy (%)	
	Variety	Cycle
LDA Stepwise	76.4	57.4
LDA	76.8	57.0
PDA	79.7	62.3
RF	87.5	80.4
SVM	90.0	83.9

(b)

Method	Per paddock classification accuracy (%)	
	Variety	Class
LDA Stepwise	82.7	63.9
LDA	86.9	68.7
PDA	85.7	77.1
RF	100.0	97.6
SVM	98.8	100.0

4. METHODS

Stepwise Fisher's linear discriminant analysis, Fisher's linear discriminant analysis using all the variables, penalised discriminant analysis, random forests and support vector machines were used to predict the categorical responses of sugarcane variety and cycle. The statistical programming package R was used to implement each of these techniques. The methods were compared by a correct classification rate derived from 10-fold cross-validation.

Predictions were initially made on a per-pixel basis. The initial factor attributes of the paddocks were assigned to the individual pixels. Predictions were then re-evaluated to produce a per-paddock prediction. This was feasible since paddocks usually contain only one sugarcane variety of a particular cycle. Predicting on a per paddock basis incorporates an object-based classifier approach (de Wit and Clevers, 2004) using paddock vector boundaries supplied by Mackay Sugar Limited. For variety and cycle, a majority or modal statistic was applied to the predicted pixels within a specific paddock. Accuracies were then assessed by comparing the predicted paddock response with the actual response. Predicting on a per-paddock basis minimizes the effect of spectral variability that occurs within any crop paddock. Pixel spectral variability in any paddock is related to, among others, variations in soil moisture condition, nutrient limitations and/or disease (de Wit and Clevers, 2004). As expected, studies by Turker and Arikan (2004), and de Wit and Clevers (2004) have shown the per-paddock classification method significantly improves the overall classification accuracies when considering a homogenous land cover type such as agricultural crops.

5. RESULTS AND DISCUSSION

Table II shows the overall 10-fold correct classification rate for classifying crop variety and cycle using stepwise Fisher's linear discriminant analysis, Fishers linear discriminant analysis, penalised discriminant analysis, random forests

Table III. Cross-validated varietal predictions for the (a) 2402 pixels and (b) 84 paddocks using the support vector machine procedure.

(a)

Actual	Predicted varietal group using the 2402 pixels data set									
Variety	20	121	124	135	136	138	159	185	190	% Correct
20	72	0	1	12	5	1	0	8	2	71.3
121	0	33	2	0	0	0	0	0	1	91.7
124	3	0	862	12	3	3	0	2	4	97
135	2	0	16	256	17	11	0	12	8	79.5
136	2	0	2	17	550	4	0	1	2	95.1
138	5	0	4	10	4	107	0	2	1	80.5
159	0	0	0	0	0	0	12	1	0	92.3
185	6	0	1	9	8	3	1	152	5	82.2
190	0	1	4	12	8	3	0	1	116	80

(b)

Actual	Predicted varietal group using the 84 paddocks data set									
Variety	20	121	124	135	136	138	159	185	190	% Correct
20	5	0	0	0	0	0	0	0	0	100
121	0	2	0	0	0	0	0	0	0	100
124	0	0	34	0	0	0	0	0	0	100
135	0	0	0	13	0	0	0	0	0	100
136	0	0	0	0	10	0	0	0	0	100
138	0	0	0	0	0	7	0	0	0	100
159	0	0	0	0	0	0	1	0	0	100
185	0	0	0	0	0	0	0	7	0	87.5
190	0	0	0	0	0	0	0	1	4	100

and support vector machines derived from the (a) per-pixel and (b) per-paddock prediction approach. Support vector machines and random forests produced the highest correct classification rates of the techniques considered. With the exception of paddock variety classification, support vector machines tended to outperform the random forest technique. Since support vector machines performed well, Tables III and IV show the confusion matrices resulting from the support vector machine variety and crop cycle classifications, respectively.

Table IIIa shows that variety 20 has the worst classification rate which was partially compounded from the mixed varieties in paddocks. The distribution of misclassification rates show that varieties 135, 138, 185 and 190 are the most spectrally confused with other varieties. However, much of this confusion diminishes when the modal statistic is applied to the pixels within the paddocks (Tab. IIIb).

Table IVa shows the class confusion matrix using the support vector machine method for the 2402 pixels where the different class levels are ordinal. Cycle 1 is the newly planted crop, cycle 2 is the first ratoon of the crop, etc. A majority of misclassifications occur just off the down-diagonal elements. An example is between the plant crop and the first ratoon. It is understandable that a first ratooning crop would be more spectrally associated with a plant crop as opposed to an 8th ratooning crop. Table 4b shows the class confusion matrix using the support vector machine for the 84 paddocks from the 2402 pixel data set. Again, the modal statistic removes extreme pixel predictions to more accurately predict paddock cycle.

To investigate the effects of class sample size, Figure 3 details the overall classification accuracies of the different classifiers for the variety factor along with the number of pixels per variety used in the models. The lower performance of Fisher's linear discriminant analysis can be linked to the sample size, which in turn directly effects the calculation of the inverse of the within covariance matrices this procedure relies on. The support vector machines and random forest procedures were to a large degree unaffected by sample size, with high accuracies achieved across all varieties. Both Fisher's linear discriminant analysis and penalised discriminant analysis methods have large variations in accuracies with respect to sample sizes. Variety 121 had the highest accuracies for a relatively small sample size. In general, Fishers linear discriminant analysis and penalised discriminant analysis methods displayed increased accuracy when the sample size increased.

6. CONCLUSION

Motivated, by limited research investigating hyperspectral imagery for agricultural systems, the performance of several discriminant methods was assessed for the purpose of discriminating sugarcane crop variety and cycle. Despite the close spectral resemblances, the random forest and support vector machines methods outperformed traditional methods and were shown to be effective on small and unbalanced class sample sizes.

Previous literature has demonstrated support vector machines to be a robust and accurate classification tool for broad vegetation class discrimination in other hyperspectral studies. This was reinforced in our investigation. Random forests also performed well and, if required, has the added benefit of obtaining a variable of importance measure that identifies which bands of the spectral image are most responsible for discriminating between different responses of the sugarcane crop. Outcomes from this research support further investigations of the

Table IV. Cross-validated sugarcane crop cycle predictions for the (a) 2402 pixels and (b) 84 paddocks using the support vector machine procedure.

(a)

Actual	Predicted crop cycle using the 2402 pixels data set									
crop cycle	1	2	3	4	5	6	7	8	9	% Correct
1	557	66	1	1	2	1	0	8	1	87.4
2	48	530	10	1	8	0	0	15	3	86.2
3	6	16	185	5	31	5	0	0	0	74.6
4	3	4	5	67	16	3	0	0	0	68.4
5	3	8	19	13	290	10	0	0	2	84.1
6	4	1	7	4	13	132	0	0	2	81
7	0	2	1	0	0	0	32	1	0	88.9
8	3	16	0	0	0	0	1	188	1	89.5
9	1	6	0	0	1	5	0	1	37	72.5

(b)

Actual	Predicted crop class using the 84 paddocks data set									
crop class	1	2	3	4	5	6	7	8	9	% Correct
1	24	0	0	0	0	0	0	0	0	100
2	0	20	0	0	0	0	0	0	0	100
3	0	0	11	0	0	0	0	0	0	100
4	0	0	0	5	0	0	0	0	0	100
5	0	0	0	0	12	0	0	0	0	100
6	0	0	0	0	0	5	0	0	0	100
7	0	0	0	0	0	0	1	0	0	100
8	0	0	0	0	0	0	0	3	0	100
9	0	0	0	0	0	0	0	0	2	100

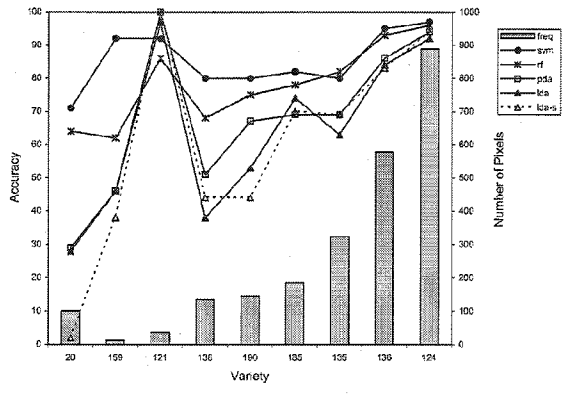

Figure 3. Accuracies of classification methods by number of pixels in each variety. The left vertical axis is the overall correct classification rate of the different classifiers and the right vertical axis is the number of pixels in each variety (bar graph). Accuracies have been derived for the 2402 pixel data set. The classifiers examined are support vector machines (svm), random forests (rf), penalised discriminant analysis (pda), Fisher's linear discriminant analysis (lda) and stepwise Fisher's linear discriminant analysis (lda-s).

advantages that hyperspectral imagery can bring to sustainable agricultural practices. For example the type of information embedded in hyperspectral signatures could greatly improve nitrogen, weed and irrigation management at the precision farming level in similar ways that this technology has been explored for corn (Waheed et al., 2006) and cotton (DeTar et al., 2006). Aided by global positioning systems installed on harvesting systems, the Australian sugar industry has seen an increase in spatially rich data in recent years. This environment would perfectly complement an extension to precision farming activities in order to take advantage of hyperspectral satellite sensors as this data becomes more accessible.

Limitations are recognised in this research approach. Discrimination of crop features focused on homogenous crop areas that are free from edge effects and mixed vegetation pixels. The context based classifier approach to improve accuracies is only applicable where detailed paddock vector data is available. In addition, this is only one season of data to draw conclusions on. Future research should focus on applying the support vector machines and random forests to other areas and multi-temporal images of sugarcane to check the consistency of our findings. Lastly, this particular area of sugarcane was well irrigated and the growing season did not suffer from drought or flooding.

Acknowledgements: The authors would like to thank Alex Held of CSIRO for providing the Hyperion data Mackay Sugar Limited for supplying the farmer survey data. Stuart Phinn of the University of Queensland and Armando Apan of the University of Southern Queensland for helpful discussions on this research. Resources were provided by James Cook University and CSIRO.

REFERENCES

Apan A., Held A., Phinn S., Markley J. (2003) Formulation and Assessment of Narrow-Band Vegetation Indices from EO-1 Hyperion Imagery for Discriminating Sugarcane Disease. in: Proc. of the Spatial Sciences Conference, 22–27 September 2003, Canberra.

Apan A., Held A., Phinn S., Markely J. (2004) Detecting sugarcane 'orange rust' disease using EO-1 Hyperion hyperspectral imagery, Int. J. Remote Sens. 25, 489–498.

Boser B., Guyon I., Vapnik V. (1992) A training algorithm for optimal margin classifiers, in: Proc. of the Fifth Annual Workshop on Computational Learning Theory, pp. 144–152.

Breiman L. (2001) Random forests, Mach. Learn. 45, 5–32.

Breiman L., Friedman J., Olshen R., Stone C. (1984) Classification and Regression Trees, Wadsworth, Belmont CA.

Bruzzone L., Melgani F. (2002) Support vector machines for classification of hyperspectral remote-sensing images, in: IEEE Int. Geosci. Remote Sens. Symp. (IGARSS), Toronto, Canada, June, Vol. 3, pp. 1670–1672.

Burges C. (1998) A Tutorial on Support Vector Machines for Pattern Recognition, Knowledge Discovery and Data Mining 2, 121–167.

Camps-Valls G., Gómez-Chova L., Calpe J., Soria E., Martín J., Moreno J. (2003) Support vector machines for crop classification using hyperspectral data, in: Lecture Notes in Computer Science 1st Iberian Conference on Pattern Recognition and Image Analysis, Springer-Verlag, pp. 134–141.

Camps-Valls G., Gómez-Chova L., Calpe J., Soria E., Martín J., Chorda L. Moreno J. (2004) Robust support vector method for hyperspectral data classification and knowledge discovery, IEEE Int. Remote Sens. 42, 1530–1542.

DeTar W.R., Penner J.V., Funk H.A. (2006) Airborne remote sensing to detect plant water stress in full canopy cotton, T. Asabe 49, 655–665.

De Wit A., Clevers J. (2004) Efficiency and accuracy of per-field classification for operational crop mapping, Int. J. Remote Sens. 22, 4091–4112.

Fisher R.A. (1936) The use of multiple measurements in taxonomic problems, Ann. Eugenics 7, 179–188.

Galvão L., Formaggio A., Tisot D. (2005) Discrimination of sugarcane varieties in southeastern Brazil with EO-1 hyperion data, Remote Sens. Environ. 94, 523–534.

Galvão L., Formaggio A., Tisot D. (2006) The influence of spectral resolution on discriminating Brazilian sugarcane varieties, Int. J. Remote Sens. 27, 769–777.

Gong P., Pu R., Yu B. (1998) Conifer species recognition with seasonal hyperspectral data, J. Remote Sens. 2, 211–217.

Gualtieri J., Chettri S. (2000) Support vector machines for classification of hyperspectral data, in: Proc. IEEE Int. Geosci. Remote Sens. Symp., (IGARSS), Vol. 2, pp. 813–815.

Hastie T., Buja A., Tibshirani R. (1995) Penalised discriminant analysis, Ann. Stat. 23, 73–102.

Hsu P., Tseng Y. (1999) Feature extraction for hyperspectral image, in: Proc. of the 20th Asian Conference on Remote Sensing, Vol. 1, pp. 405–410.

Huang C., Davis L., Townshend J. (2002) An assessment of support vector machines for land cover classification, Int. J. Remote Sens. 23, 725–749.

Hughes G. (1968) On the mean accuracy of statistical pattern recognizers, IEEE T. Inform. Theory 14, 55–63.

Johnson R.A., Wichern D.W. (1988) Applied multivariate statistical analysis, 2nd ed., Englewood Cliffs, Prentice Hall NJ.

Kohavi R. (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection, in: Int. Joint Conference on Artificial Intelligence (IJCAI), http://robotics.stanford.edu/users/ronnyk/.

Lennon M., Mercier G., Hubert-Moy L. (2002) Classification of hyperspectral images with nonlinear filtering and support vector machines, in: IEEE Int. Geosci. Remote Sens. Symp. (IGARSS), Toronto, Canada, June, Vol. 3, pp. 1670–1672.

Lillesand T., Kiefer R. (2000) Remote Sensing and Image Interpretations, Fourth Ed., Wiley and Sons, New York.

Mercier G., Lennon M. (2003) Support vector machines for hyperspectral image classification with spectral based kernels, in: Proc. IEEE International Geoscience and Remote Sensing Symp. (IGARSS), Toulouse, France, 21–25 july, CD-ROM.

NASA (2003) New Millennium Program Earth Observing-1 Brochure. NASA Goddard Space Flight Center, Greenbelt, Maryland, http://eo1.gsfc.nasa.gov.

Shah C.A., Watanachaturaporn P., Arora M.K., Varshney P.K. (2003) Some Recent Results on Hyperspectral Image Classification, in: IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data, October 27–28, 2003, NASA Goddard Spaceflight center, Greenbelt, MD, CD-ROM.

Thenkabail P., Smith R., Depauw E. (2002). Evaluation of narrowband and broadband vegetation indices for determining optimal hyperspectral wavebands for agricultural crop characterisation, Photogramm. Eng. Rem. S. 68, 607–621.

Truong Y., Lin X., Beecher C. (2004) Learning a complex metabolomic dataset using Random Forests and Support Vector Machines, Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining, August 22–25, Seattle, Washington.

Turker M., Arikan, M. (2004) Field-Based Crop Mapping through Sequential Masking Classification of Multi-temporal LANDSAT-7 ETM+ Images in Karacabey, Turkey, Int. Arch. Ph. R. S., Band XXXV, Part B7, Istanbul, pp. 192–197.

Vapnik V. (1998) Statistical Learning Theory, John Wiley & Sons, New York.

Waheed T., Bonnell R.B., Prasher S.O., Paulet E. (2006) Measuring performance in precision agriculture: CART – A decision tree approach, Agr. Water Manage. 84, 173–185.

Watanachaturaporn P., Arora M., Varshney P. (2004) Evaluation of factors affecting support vector machines for hyperspectral classification, in: American Society for Photogrammetry & Remote Sensing (ASPRS) 2004 Annual Conference, May 23–28, Denver, CO, CD-ROM.

Wilson M.D., Ustin S.L., Rocke D.M. (2004) Classification of contamination in salt marsh plants using hyperspectral reflectance, IEEE T. Geosci. Remote 42, 1088–1095.

Yu B., Ostland M., Gong P., Pu R. (1999) Penalized linear discriminant analysis of in situ hyperspectral data for conifer species recognition, IEEE T. Geosci. Remote 37, 2569–2577.