

AGRICULTURAL CROP IDENTIFICATION USING SPOT AND LANDSAT IMAGES IN TASMANIA

Rachel BARRETT*, Paul CROWTHER**, Rowland LAURENCE***, Ross LINCOLNE****

University of Tasmania, Australia

* School of Agricultural Science

Rachel.Barrett@utas.edu.au

** Tasmanian Institute of Agricultural Research

Rowland.Laurence@utas.edu.au

*** School of Computing

P.Crowther@utas.edu.au

**** Central Science Laboratory

R.Lincolne@utas.edu.au

KEY WORDS: Agriculture, Remote sensing, Multi-spectral data, Image processing, Recognition

ABSTRACT

Recording precise crop location and paddock rotational history is becoming an increasingly important part of crop management and product quality assurance. Advances in satellite imagery now allow economies of scale in the calculation of paddock position, area measurement and mapping, when compared with aircraft-based photography. Crop recognition, growth and health monitoring are now also considered achievable goals using satellite imagery. This paper reports on the ability of current technology to recognise and differentiate between horticultural crops in Tasmania, Australia. The methodologies used for recognising crops, the testing methods used and the results of two years of interpretation are discussed.

1 INTRODUCTION

Tasmania is a source of many high value, commercial crops, including poppies for the pharmaceutical industry, pyrethrum for the production of natural insecticides, potatoes, onions and peas. The cool, temperate climate, winter dominant rainfall and variable soils enable a wide variety of intensive horticultural crops to be grown in areas from about 2 up to 20 hectares.

In 1996, the University of Tasmania initiated a project in response to a request from local industries. The aim was to find a simple and cost effective method of recognising crops in agricultural areas and record their exact location.

A pilot study indicated that distinction between some crops was possible with data from good sequential satellite imagery. In October 1997 funding was secured to investigate the ability of current technology to identify and differentiate between the many horticultural crops grown using commercially available, multi-spectral satellite images for the growing seasons of 1997-98 and 1998-99. The first stage of this project was to develop a methodology for the systematic recognition of individual crops. If this were achieved the establishment of growth and yield patterns could lead to large scale monitoring of agronomic influences. In addition the basis for a simple and cost-effective paddock inventory system would be provided.

1.1 Study Areas

Two representative but distinct study areas of intensive horticultural production within Tasmania, the North West Coast and the Coal River Valley were selected, since different crops grown in distinct climatic and edaphic environments may have dissimilar growth signatures (Nieuwenhuis *et al.*, 1996). This paper will report on the North West study area alone.

The North West Coast study area lies between 41° 07' 13" S and 41° 18' 45" S and from 146° 03' 58" E to 146° 34' 2" E (Figures 1& 2). Average annual rainfall in the study area ranges from 750mm to 1000mm, of which approximately two thirds occurs in winter. With mean daily maximum temperatures rarely exceeding 32 °C, summers are mild. The study area receives seven to eight hours of direct sunlight in mid-summer and about four hours in mid-winter.

These meteorological conditions and irrigable krasnozem soils make the North West Coast of Tasmania, and the study area, particularly suited to the production of many crops including pharmaceutical poppies, pyrethrum, potatoes, peas, beans, carrots and onions for both processing and fresh market. Other vegetable crops include pumpkin, broccoli, cauliflower, and corn for the fresh domestic market and seed production. Cereals such as barley, oats, wheat and triticale are mainly grown for stock feed.

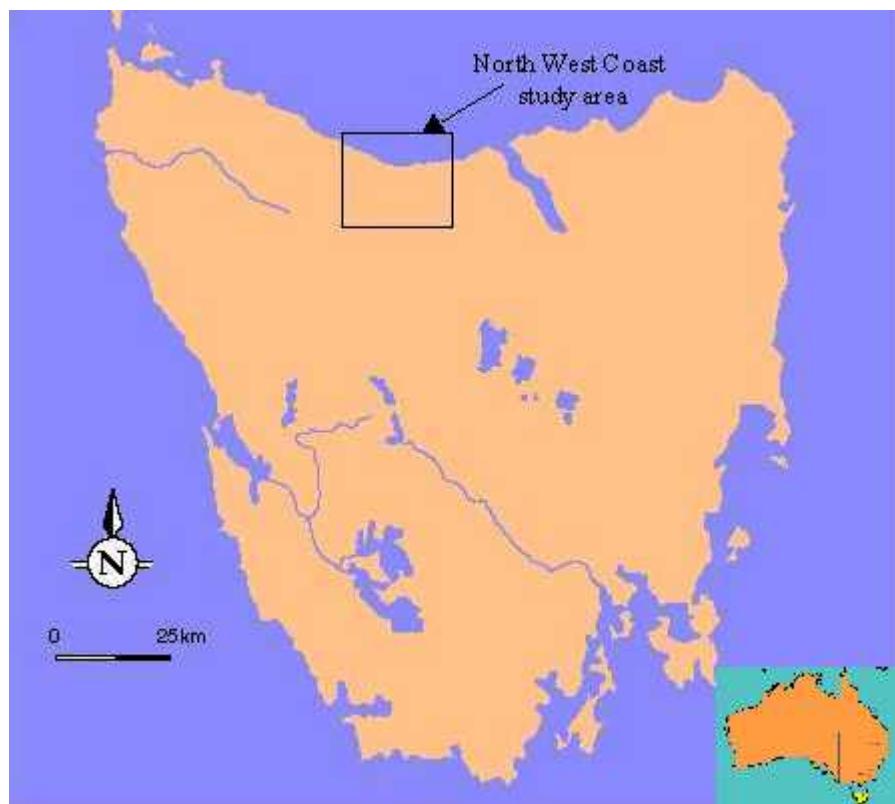


Figure 1. Location of the North West Coast study area in Tasmania, Australia.

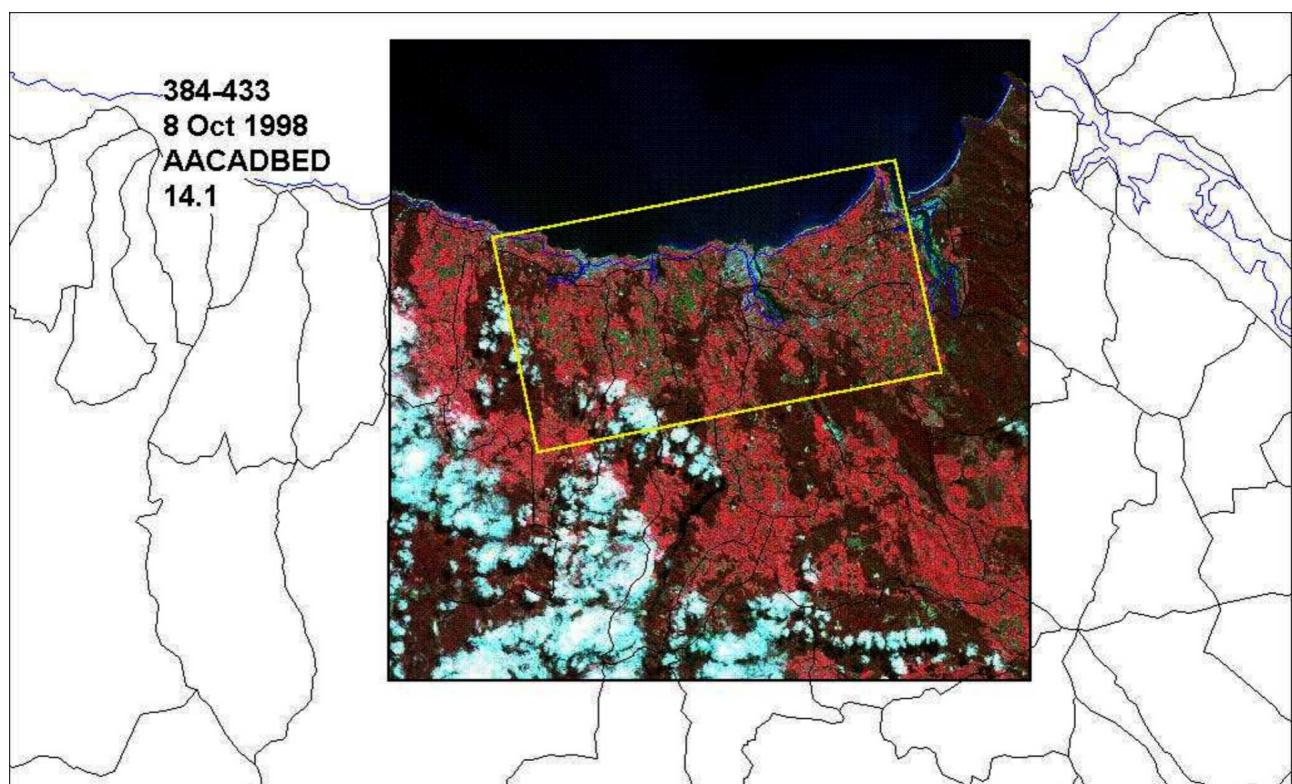


Figure 2. SPOT XS image of the North West Coast study area.

2 MATERIALS AND METHODS

Digital data from SPOT-XS, SPOT-XI and Landsat TM were acquired over the study area during the spring/summer growing seasons of 1997/98 and 1998/99. These data were processed, classified and compared to ground-truthed data of crop cover as described below.

2.1 Collection of Ground Based Data.

Ground-based data were collected at fortnightly intervals during these growing seasons. The 17 crop types selected for recording based on commercial significance, were oats (OA), triticale (TR), wheat (WH), barley (BA), onions (ON), carrots (CS), peas (PE), beans (BE), potatoes (PO), pumpkin (PU), broccoli (BR), poppies (PP), pyrethrum (PY), cauliflower (CA), sweet corn, lucerne and pasture. The collected data assessed and recorded crop type, growth stage and ground cover from randomly selected field areas representing approximately 10% of the total area under cultivation. Crop boundaries were recorded utilising 1:5000 orthophoto map series.

2.2 Data Acquisition and Initial Processing.

Digital data from SPOT-XS, SPOT-XI and Landsat TM systems were acquired, for several reasons. As stated by Reid *et al* (1993), the high incidence of cloud over Tasmania makes the acquisition of cloud-free images via the Landsat TM system alone somewhat difficult. These limitations occur elsewhere (Stoney, 1996), (Wiegand *et al*, 1996). The SPOT system increases the ability to acquire cloud-free images by the use of off-nadir viewing. The use of both Landsat TM and SPOT increases the frequency of target overpass compared with the use of one system only.

Subsequent to image acquisition, and using ERDAS Imagine Version (8.3) the images were resampled to produce uniform 20 metre square "pixels", and to accurately align the images to one another. The images were then "stacked" to make one image. This made it possible to analyse all the information for the growing season at one time. Crop boundaries, which had been manually digitised from 1:5000 orthophoto map series were converted to digital form and then incorporated as a mask, so that information outside of the crop boundaries (areas of interest, AOI) were masked out to black.

For the 1997/98 growing season, six images were acquired, between 02 July 1997 and 18 April 1998. The sequence was as follows; TM 02 July 1997, TM 06 October 1997, TM 23 November 1997, TM 10 January 1998, XS 23 February 1998 and XS 18 April 1998, providing a total of 30 measurements in each "pixel" of the AOI. For the 1998/99 growing season, seven images were acquired, between 21 July 1998 and 03 March 1999. The sequence was as follows; TM 21 July 1998, TM 27 September 1998, XS 08 October 1998, XI 31 December 1998, TM 29 January 1999, TM 03 March 1999 and XI 29 March 1999, providing a total of 35 measurements in each "pixel" of the AOI.

2.3 Image Classification

The software package chosen was ERDAS Imagine Version (8.3). Figure 3, represents diagrammatically the image classification process. The first step in the analysis procedure was a Principal Components Analysis (PCA) executed on the "stacked", masked images incorporating all the season's image data. This procedure is a standard procedure to concentrate the significant information into a smaller number of bands. About a third of the bands were discarded after this procedure because they mostly contained noise.

An unsupervised classification process (maximum likelihood), then noted which pixels were similar and formed classes for similar pixels. Five hundred classes were utilised in the 1997/98 season and 400 in the 1998/99 season. Classes were used to "classify" the whole-unmasked area. After this classification process, a file was produced which contained the details of the average values for each of the classes. All pixels for each paddock could then be represented by a class, to which the pixel belonged.

The data was then prepared for entry into a custom computer program ("PaddockId", compiled in C Program language). The PaddockId program had three input files. The first was the information from an unsupervised classification, containing the average values of the classes. The second file listed, for each paddock, the number of pixels in each of these classes. The program also required a training dataset. About one third of the field areas for each crop type recorded in the ground-truthing was used as "training data", which was used to "learn" the characteristics of the crops and their variation with time. These were called the "known field areas". The remainder, were the "unknowns", and analysed by the procedure and used to score the accuracy of the process. The third file listed the class types expected, the identification numbers of the "known" field areas and their crop types, the paddock identification numbers of any field areas affected by cloud, and some probability limits used by the program to decide when field areas could be considered to be similar.

The program used information extracted from the "stacked", masked image. Several runs of this program were conducted, using all of the bands of information for one year. The amalgamation of the results from the various

temporal analyses produced a report of the various crops grown in that season. The results of the different iterations of the program were amalgamated to identify the crop or crops grown in each field area during the season.

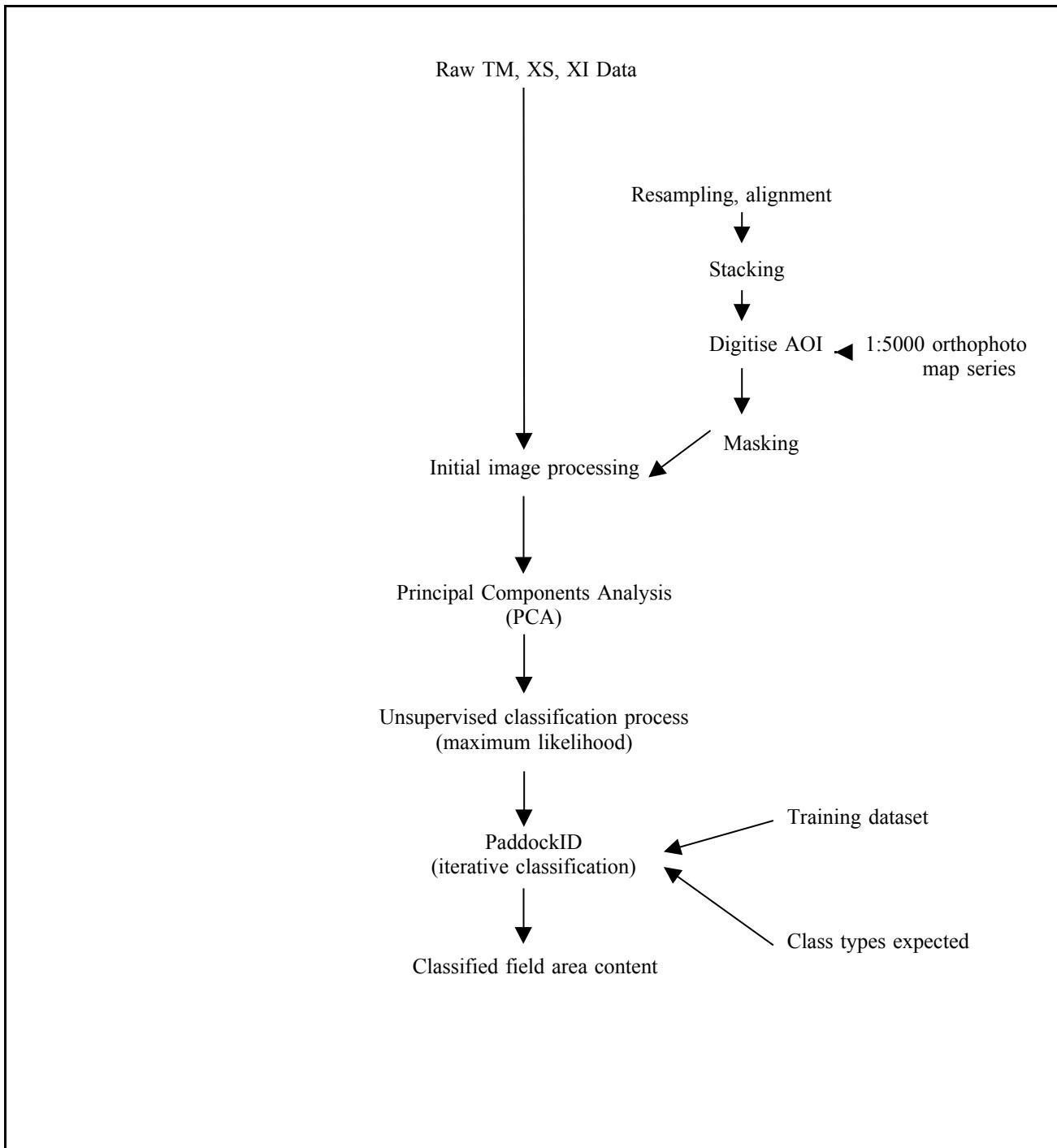


Figure 3. Diagrammatic representation of the image classification process.

Different operators scored the results, so that the identity of the "unknown" field areas was not available to those performing the analysis.

2.4 Accuracy Assessment

The prediction results achieved were entered in a matrix table and errors of omission, commission, and normalised overall accuracy were calculated as described by Congalton (1991), with actual crop identity displayed as y coordinate (vertical axis) and the prediction displayed as x coordinate (horizontal axis). This representation of the data therefore reported not only the predictive accuracy of individual crop types but also exemplifies the outcomes of unsuccessful predictions.

3 RESULTS AND DISCUSSION

3.1 North West Coast 1997/98 Growing Season

Table 1 shows a range of predictive accuracy for the subject crops from 0% for triticale to 100% for oats. The average predictive accuracy over all crops obtained from this assessment is 70%, with major crops such as potatoes, poppies, pyrethrum, carrots, onions and peas approximately 76%. These data indicate that predictive accuracies should be greatly improved by specifically tailoring interpretative images to cropping sequences.

		Predicted															
		BA	OA	TR	WH	ON	CS	PE	BE	PO	PU	BR	CA	PP	PY		
Field	BA	5	0	0	1	0	0	0	0	0	1	0	0	0	0	7	71.4
	OA	0	1	1	0	0	0	1	1	0	0	0	0	1	0	5	20.0
	TR	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0.0
	WH	0	0	1	4	0	0	0	0	0	0	0	0	0	0	5	80.0
	ON	1	0	0	0	16	0	0	1	0	0	0	0	0	0	18	88.9
	CS	0	0	0	0	3	8	2	1	1	1	0	0	0	0	16	50.0
	PE	1	0	0	0	0	0	16	0	3	0	0	0	0	2	22	72.7
	BE	0	0	0	0	0	0	0	3	1	0	0	0	0	0	4	75.0
	PO	3	0	0	0	1	1	3	1	23	2	0	1	3	1	39	59.0
	PU	0	0	0	0	1	0	0	1	0	3	0	0	1	1	7	42.9
	BR	0	0	0	0	0	0	5	1	0	0	2	0	2	0	10	20.0
	CA	0	0	0	0	0	0	0	0	0	0	0	4	0	1	5	80.0
	PP	0	0	0	0	0	0	2	1	1	0	0	0	34	2	40	85.0
	PY	0	0	0	0	0	0	1	0	2	0	1	0	1	34	39	87.2
		10	1	2	5	21	9	30	10	31	7	3	5	43	41	153	
		50.0	100.0	0.0	80.0	76.2	88.9	53.3	30.0	74.2	42.9	66.7	80.0	79.1	82.9	218	218
Producers Accuracy (%)																218	70.2%

Table 1. Omission/comission matrix for North West Coast 1997/98 growing season

Crops differ not only in phenology, but also in morphology and physiology, which affect the leaf area index and in turn, reflectance (Nieuwenhuis *et al*, 1996). This point is particularly emphasised by the low accuracy with respect to the identification of bean crops, which was 30%. Beans have a short growth cycle (approximately 8 to 10 weeks) but conversely, a specific canopy closure period. As a result, only two images were of significant benefit in identification of this crop, with the remaining 4 images, which were included, probably creating distortion within the programming algorithms.

3.2 North West Coast 1998/99 Growing Season.

The methodology utilised in this project has resulted in some crops being identified with better than 90% accuracy (Table 2). Favourable predictive outcomes have been achieved for crop types for which the greatest quantity of both training and test data were acquired (poppies, 65.8%, pyrethrum, 86.4%, potatoes, 85.0%, peas, 78.0% and onions 69%). This reiterating that the choice of a good training set can also have significant influence on the success of a classification approach, (Shine & Wakefield, 1999).

The overall accuracy for the 1998/99 growing season (69.8%) has declined from that recorded in the 1997/98 growing season. Given that there was an increased number of observations made over the majority of crop types, this decline in performance may be attributed to other factors.

The combination of too many variables and few observations (i.e. sample sizes) usually causes unreliable estimates (Cruz-Castillo *et al*, 1994). It is therefore imperative that, if increases in accuracy of prediction are to be achieved, sufficient calibration data (ground-truthing) must be obtained. There is however, a critical level where any increases in the quantity of calibration data will no longer result in increases in accuracy.

Another possible contributing factor in the decline of the overall accuracy is the unsupervised classifier choice of maximum likelihood. It is generally acknowledged that human photointerpreters use a considerable variety of contextual information and common-sense experience in interpreting aerial imagery (Wilkinson & Burrill, 1991), (Kontoes *et al*, 1993). Given that many classifier types are used in remote sensing (maximum likelihood, two-class decision, parallelepiped, knowledge based, semi-variogram, fuzzy, kriging, etc.), much of this common-sense experience comes into play with regard to selection of a classifier

		Predicted															
		BA	OA	TR	WH	ON	CS	PE	BE	PO	PU	BR	CA	PP	PY		
Field	BA	1	0	0	0	2	0	0	0	0	0	0	0	0	0	3	33.3
	OA	0	2	0	0	0	0	0	0	1	0	0	0	0	0	3	66.7
	TR	0	1	3	6	0	0	0	0	0	0	0	0	0	2	12	25.0
	WH	0	0	1	2	0	0	0	0	0	0	0	0	1	0	4	50.0
	ON	0	0	0	0	20	1	0	0	0	0	0	0	8	0	29	69.0
	CS	0	0	0	0	3	0	1	0	1	0	0	0	2	1	8	0.0
	PE	0	0	0	0	0	1	39	1	1	1	0	1	4	0	48	81.3
	BE	0	0	0	0	0	2	1	2	0	0	0	0	2	0	7	28.6
	PO	0	0	1	0	0	0	3	0	17	0	0	1	5	1	28	60.7
	PU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0
	BR	0	0	0	0	1	0	3	0	0	0	0	0	3	0	7	0.0
	CA	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0.0
	PP	0	0	0	0	3	2	0	0	0	0	0	0	52	2	59	88.1
	PY	0	0	0	0	0	1	2	0	0	0	0	0	2	38	43	88.4
		1	3	5	8	29	7	50	3	20	1	0	2	79	44	176	
		100.0	66.7	60.0	25.0	69.	0.0	78.0	66.7	85.0	0.0	0.0	0.0	65.8	86.4	252	252
		Producers Accuracy (%)														252	69.8%

Table 2. Omission/comission matrix for North West Coast 1998/99 growing season

4 FUTURE WORK

The objective of this project, to develop a methodology for the systematic recognition of individual crops, has been partially achieved. In order to better utilise the technology in a commercial environment, the methodology will be tailored individually to five commercially significant horticultural crops. Poppies, pyrethrum, potatoes, peas and onions have been selected, not only for their commercial significance but also because, in the initial project, the predictive accuracy of these particular crops was the highest of 17 different crops investigated.

Investigation of the timing of image acquisition, and the number of images required for specific crop growth cycles will theoretically enable accuracy to be increased.

Determination of the most appropriate classifier type for individual crops will enable the most accurate and specific prediction methodologies to be documented.

Given the differences between the SPOT and Landsat platform sensors, resolution and bandwidths, examination of the most appropriate combination of SPOT and/or Landsat TM images will determine if a particular platform or combination is more suited to a particular crop type. Again, this will enable the use of remote sensing data for crop recognition to become more cost effective and accurate.

5 CONCLUSION

Favourable predictive outcomes have been achieved for crop types for which the greatest quantity of both training and test data were acquired. These were poppies, pyrethrum and potatoes. This is a logical conclusion from both an agronomic and computational point of view, in that the probability of achieving a correct outcome will be enhanced if a representative range of horticultural and environmental conditions have been used in both the training and test data. The predictive accuracy should be greatly improved by specifically tailoring interpretative images to cropping sequences. A point reinforced by the low predictive accuracy with respect to bean crops.

ACKNOWLEDGEMENTS

The authors wish to thank the Horticultural Research and Development Corporation (HRDC), Australia and Botanical Resources Australia (BRA) for funding this research.

REFERENCES

- Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37(1): 35-46.
- Cruz-Castillo, J.G., Ganeshanandam, S., MacKay, B.R., Lawes, G.S., Lawoko, C.R.O., Woolley, D.J., 1994. Applications of canonical discriminant analysis in horticultural research. *HortScience*, 29(10): 1115-1995.
- Kontoes, C., Wilkinson, G.G., Burrill, A., Goffredo, S., Megier, J., 1993. An experimental system for the integration of GIS data in knowledge-based image analysis for remote sensing of agriculture. *International Journal of Geographical Information Systems*, 7(3): 247 - 262.
- Nieuwenhuis, G.J.A., de Wit, A.J.W., van der Wal, T., Müncher, C.A., van Leeuwen, H.J.C., 1996. Integrated use of high and low resolution satellite data and crop growth models. Report 127. DLO Winand Staring Centre, Wageningen, The Netherlands.
- Reid, R.N.D., Vickery, P.J., Hedges, D.A., & Williams P.M., 1993. Measuring the response of pasture to superphosphate using aircraft and satellite remote sensing. *Australian Journal of Experimental Agriculture*, 33(5): 597-600.
- Shine, J.A., Wakefield, G.I., 1999. A comparison of supervised imagery classification using analyst-chosen and geostatistically-chosen training sets. In: U.S.A.T.E. Centre (Editor). U.S. Army Topographic Engineering Centre, Alexandria, VA, USA.
- Stoney, W.E., 1996. Land observation satellites in the next ten years. InAgriculture in the 21st Century; Proceedings of the Workshop on Remote Sensing at University of California, Davis, California.. Lecture Outline/Papers from the Workshop on Remote Sensing for Agriculture in the 21st Century.
<http://cstars.ucdavis.edu/proj/ag-21/lect0912.pdf>
- Wiegand, C., Anderson, G., Lingle, S., Escobar, D., 1996. Soil salinity effects on crop growth and yield - illustration of an analysis and mapping methodology for sugarcane. *Journal-of-Plant-Physiology* Special issue: vegetation stress II. First international symposium on vegetation stress, Munich, Germany, 19-21 June 1995., 145(3-4): 418 - 424.
- Wilkinson, G.G., Burrill, A., 1991. Integration of GIS-derived spatial data with satellite imagery for improved land cover mapping using an expert system. In: H.F.L.O. Janjaap Harts, Henk J. Scholten. (Editor), European Conference on Geographical Information Systems. EGIS '91 : proceedings : Utrecht, Netherlands : EGIS Foundation, Brussels, Belgium, pp. 1241-1249.