Continental Monitoring: 34 Years of Land Cover Change Using Landsat Imagery

P. Caccetta*, S. Furby, J. O'Connell, J. Wallace, X. Wu

* Mathematical and Information Sciences, CSIRO, Perth, 6019, Western AUSTRALIA – {Peter.Caccetta, Suzanne.Furby, Jared.O'Connell, Jeremy.Wallace, Xiaoliang.Wu}@CSIRO.au

Abstract – In Australia, remotely sensed Landsat data is routinely used for mapping and monitoring the change in extent of woody perennial vegetation. Time-series remotely sensed satellite imagery and ground information is used to form multi-temporal classifications of presence/absence of woody cover, typically for state-based management and reporting of clearing. More recently, there has been increased interest from landuse changes associated with agriculture and forestry being a significant component in Australia's carbon budget. To quantify the amount, the capability for continental monitoring of land cover changes using Landsat data has been developed through a collaboration between the Australian Greenhouse Office (AGO), the CSIRO, and other partners. The project, called the National Carbon Accounting - Land Cover Change Project (NCAS-LCCP) uses some 5000 Landsat MSS, TM and ETM+ images to map the presence/absence of perennial vegetation at 25m resolution for fifteen time periods since 1972. The landsat based monitoring program is expanding to include monitoring of sparse woodland (5-20% canopy cover) and urban change to meet both carbon accounting and natural resource management needs.

In this paper we describe our methodology, with a focus on the remote sensing algorithms and techniques we employ, and present results of forest change along with estimates of accuracy.

Keywords: classification, hidden Markov model, time series, calibration.

1. INTRODUCTION

In Australia, continental mapping and monitoring of the extent and change in perennial vegetation using Landsat satellite imagery is routinely performed as part of the National Carbon Accounting System - Land Cover Change Project (NCAS-LCCP). The methods within the operational LCCP system were developed by the CSIRO Mathematical and Information Sciences division in collaboration with the Australian Greenhouse Office (AGO) and other partners. Under a framework of contracts and Quality Assurance (QA) procedures, commercial companies apply these methods to the growing archive of Landsat images to produce time-series continental coverages of the presence and absence of perennial vegetation cover at a pixel resolution of 25m. The raw data archive currently consists of approximately 5000 Landsat images having an approximate data volume of 2×1012 bytes (2 terabytes), which is transformed into information products having similar data volumes. Given the above operating environment, accuracy, interpretability (for outsourcing and QA), computational efficiency, the ability to incorporate "better" algorithms, and reliability when applied through space and time, are important aspects for consideration during methodology development. The current operational monitoring system includes the following components:

- registration of time-series Landsat data to a common spatial reference,
- calibration of Landsat data to a common spectral reference,
- (if required) processing of the calibrated data to adjust for viewing geometry effects including differential terrain illumination,
- stratification of the data into 'zones', where landcover types within a zone have similar spectral properties,
- processing of the calibrated data to remove 'corrupted' data, which include dropouts, data affected by fire, smoke and cloud,
- analysis of ground and satellite spectral data to determine, for each date, a classifier and its parameters,
- specification of a joint model for multi-temporal classification
- validation of the classifications to quantify their accuracy.

The demands for the spatial and temporal resolution used here arose from the development of the reporting rules (Marrakech Accord) guiding the implementation of accounting procedures for the Kyoto Protocol. However the remote sensing program has many additional benefits for bodies interested in monitoring landuse change generally. An overview of techniques used for the NCAS LCCP are described in the ensuing sections.

2. MATERIALS AND METHODS

We use:

- Long-term sequences of orthorectified (section 2.1) and calibrated (section 2.2) Landsat MSS, TM, ETM+ satellite data to provide observations relating to land cover;
- Discriminant analysis techniques to (spectrally) separate classes of interest; in this case forest and non-forest. Sparse systems and urban areas are currently being investigated.
- (section 2.2) Supervised and automated approaches to specify / estimate classifier parameters; and
- (section 2.3) Spatial/temporal models to reduce errors

2.1 Landsat Rectification

The two steps in establishing a rectified sequence of Landsat imagery are: 1) establish a common ortho-rectified base mosaic of Landsat data; and 2) ortho-rectify temporal sequences of images to the common base. Accurate ortho-rectification was achieved using a rigorous earth-orbital model. PCI OrthoEngine software was used for this purpose. Once the ortho-rectified base was established, ground control points (GCP) were automatically matched using a cross-correlation technique. This approach improves efficiency and accuracy of the results. For quality assurance, visual inspection and numerical summaries based on cross-correlation feature matching are used to assess the accuracy of ortho-rectification of the time series images.

2.2 Image Calibration

Radiometrically calibrated images allow for comparisons between image scenes, and the possibility of better extrapolation of a chosen classifier. We convert raw digital counts to be consistent with a chosen reference image. Three calibration steps are applied in the radiometric correction procedure for Landsat imagery: 1) Top-Of-Atmosphere (TOA) reflectance calibration (also called sun angle and distance correction); 2) Bi-directional Reflectance Distribution Function (BRDF) calibration; and 3) terrain illumination correction. Each step is briefly described in the following subsections.

2.2.1 Top-Of-Atmosphere (TOA) reflectance calibration: The TOA calibration is to correct the reflectance differences caused by the solar distance and angle. The sun zenith and azimuth angles for each pixel and the distance from the scene centre to the sun are calculated and the reflectance correction is then calculated for each band as described in Vermote *et al.* (1994).

2.2.2 BRDF Calibration: Angular effects across the Landsat image result in BRDF effects which are relatively small, but significant in the context of broad-scale monitoring. The BRDF correction is a simple linear function of scan angle which is applied to each band. A two-kernel empirical BRDF model was used to correct the remaining scene-to-scene differences. Simple variations of Walthall's model, as described by Danaher *et al.* (2001), were used in the BRDF calibration approach. The model is a three-parameter model (see Equation 3 in (Danaher *et al.*, 2001)), where the three parameters were calculated by solving equations based on the image overlap areas, and the same parameters were applied to all scenes (Wu *et al.*, 2001).

2.2.3 Terrain Illumination Correction: This third step is required where there are significant terrain illumination effects, resulting in bright and dark sides of hills and mountains. This is particularly important for time series imagery where terrain effects vary with different dates. In practice it also resulted in reduced stratification for classification (section 2.2). The details of the terrain illumination correction used can be found in (Wu *et al.*, 2004), which is based on the C-correction (Teillet *et al.*, 1982) and incorporates a ray tracing algorithm for identifying true shadow. A high-resolution digital elevation model (DEM) is required to achieve adequate removal of terrain effects.

2.2 Classification of remotely sensed and other data

Given a time series of satellite and ground data, we wish to form a classification time series of landcover. One approach is simply to form the classification time series from independent classifications (discussed in section 2.2.1) of each of the images in the sequence. Previous work by Caccetta (1997) and Kiiveri and Caccetta (1998) have demonstrated that greater classification accuracies may be achieved by the use of spatial-temporal models which jointly estimate the time series of labels (discussed in section 2.2.2. Thus we use a two step approach comprised of forming the best single date classifications, expressed as posterior probabilities, as input to the second step which is a spatial temporal model.

2.2.1 Single Date Classification

Perhaps the most popular classifier for optical data is the Maximum Likelihood Classifier (MLC) (Rao, 1966). These classifiers generally assume that spectral descriptions for classes can be modelled using multivariate Gaussian densities.

Here we apply band reduction routines related to canonical variate analysis (CVA) to identify the important spectral band combinations for the questions of interest, and related routines (McKay & Campbell 1982) to smooth the band coefficients to derive simplified spectral indices. Instead of modelling the spectral data using multivariate Gaussian distributions, we derive index thresholds (for one or more indices) to produce single date classifications using intervals between thresholds to produce ‘soft’classifications (initial interpretations).

Thus the classifier is constructed in two steps:
1. the derivation of linear combinations of spectral bands, ‘indices’, that give the “best” separation of classes using Canonical Variate Analysis (Campbell and Atchley, 1981). Typically two or three indices are used.
2. specification of decision boundaries (thresholds) on these indices to map the spectral space into classes. Here we consider two classes; *Forest*, and *non-forest*. For each decision boundary, we use two thresholds to produce a soft boundary. For each index, we have one or two decision boundaries.

Typically a set of spectral indices is chosen once for each zone in an image, and the thresholds are adjusted year to year to account for seasonal variation and limitations in the accuracy of calibration. In all, there are approximately 1000 zones defined for the continent.

There were several reasons why this relatively simple classifier was chosen for operational work;

• Initial pilot studies demonstrated that forest/non-forest classes were typically separable using two or three linear discriminant functions (Furby and Woodgate, 2002).
• The nature of the classifier allows for manual intervention in the fitting of models (via image processing software).
• The simplicity of the classifier meant we were confident in its ability to extrapolate well.
• Computational issues, a typical training region may be of the order of 2000×2000 = 4000000 observations (pixels) and it is useful in an operational setting to perform classification on desktop PC’s.

Initially the method was applied using operators to set the thresholds. An improvement in consistency and subsequent accuracy was achieved by incorporating an automated method which we will call “matching” in the following, leading to a relatively automated but still supervised approach, which we will now describe.

Following (Caccetta and Bryant, 2002), we assume a reference or ‘base’ image is formed by an experienced operator which is then to be used as a response variable. Given a new image to classify, we estimate thresholds by minimising over m pixels the objective function,

$$\sum_{k=0}^{m} | \hat{p}_k - p_k | \quad (1)$$

where p_k are the class probabilities for pixel k from the base image and $\hat{p}_k$ are the estimated probabilities for the new image. This function is minimised using the Simplex algorithm (Nelder and Mead, 1965). The nature of this model is conducive to manual fine tuning, a user can simply adjust the thresholds in the case of a poor fitting model, which do occur operationally with some frequency. We also investigated updating the index coefficients while estimating the thresholds, though this produced insignificant improvements.

We also compared the Matching (M) approach (see O'Connell and Caccetta, 2006) to several popular techniques; Random Forests (RF), CART (Breiman et al., 1984), Linear Discriminant Analysis (LDA) (Ripley, 1996) and logistic regression (LR) (McCullagh and Nelder, 1983). R (R Development Core Team, 2004) implementations of each classifier (excluding Matching) were used. Random Forests is considered a "state of the art" classifier, on a par with other techniques such as Neural Networks and Support Vector Machines (SVM) (Breiman, 2001). We view the classification accuracy obtained from Random Forests as an upper bound on the accuracies possible for single date classification. We consider the difference between this upper bound and our current technique as a guide for possible benefit (improved classification accuracy) versus costs (computational requirements, interpretability) associated with more sophisticated classification methods. For an area of no change, table A summarises the residual error as described by (1) resulting from the various classifiers. As a tradeoff between accuracy, interpretability, and computation, we concluded that the matching approach was adequate.

2.3 Spatial temporal models for classification

The time series is composed of sources of data of varying quality and spectral discrimination. To improve classification accuracy, we use joint models which incorporate error rates of the initial interpretations as well as temporal and spatial rules.

For pixel k=1..m and time slice i=1..n we write y_{ik}, y'_{ik}, l_{ik}, l'_{ik} and z_{ik} to represent landsat coverage, index image, initial interpretation, `true' class label and stratification zone respectively. We use the notation r(k) to denote the 8 pixels adjacent to k. The joint model for the observed and unobserved images can be written as

$$p(y'_{1k} .. y'_{nk}, l'_{1k} .. l'_{nk}, l_{1k} .. l_{nk}, l'_{1r(k)} .. l'_{nr(k)}, z_{1k} .. z_{nk}) = \Pi_{i=1..n}\, p(y'_{ik} \mid z_{ik}, l_{ik})\, p(l_{ik} \mid z_{ik}, l'_{ik})\, p(l'_{ik} \mid l'_{i-1,k}, l'_{ir(k)})\, p(z_{ik}) \quad (2)$$

where $p(l'_{1k} \mid l'_{0k}, l'_{ir(k)})$ is defined as $p(l'_{1k} \mid l'_{ir(k)})$.

The `true' class labels are obtained from the joint distribution of $l'_{1k} .. l'_{nk}$ given the observable images by a cyclic ascent algorithm as in (Caccetta, 1997 pp 187-192), (Kiiveri and Caccetta, 1998).

Useful properties of the approach include:

- propagation of *uncertainties* in inputs and calculation of uncertainties in outputs
- production of *hard* and *soft* maps
- handling of *missing data* by using all available information to make predictions
- existence of well-developed statistical tools for parameter estimation.

The effect of applying this step is demonstrated in Table A, where the columns identified with the * are the results after application of the model.

Table A. Typical results: summary of residual differences for SI56 SW, Zone 4.

	Method						
Year	LR	LDA	CART	M	RF	M*	RF*
1989	6.63	6.47	7.6	6.36	6.21	6.04	4.52
1991	6.23	6.04	7.33	6.11	5.78	4.95	4.10
1992	6.15	5.79	6.09	5.54	5.34	3.72	3.24
1995	3.44	3.58	3.68	2.8	2.87	1.56	1.52
2000	5.76	5.73	6.22	5.21	5.2	3.51	3.26
2002	6.33	6.64	6.75	6.29	5.84	5.00	4.24
2004	9.16	8.53	9.89	8.39	8.07	7.77	5.29
Mean	6.24	6.11	6.8	5.82	5.62	4.65	3.73

For the project as a whole, the accuracy of the final forest presence/absence classifications are independently validated. Validation involves comparison of classifications against "truth" obtained from aerial photo interpretation. Details can be found in Lowell *et al* (2003).

2.4 Discussion

Matching generally compared favourably with RF, with both classifiers achieving similar results for the majority of data tested. There were some notable exceptions, where RF could operate accurately for atypical data such as haze or early growth, and was inaccurate when a poor 'base image' was chosen while Matching was somewhat more robust. Indeed, it was found that the choice of base image was far more influential on estimates than the choice of classifier.

We face some difficult challenges with data of this nature. An obvious problem is size, few classification techniques are aimed at data sets where n > 1000000, this is largely an implementation issue. Another (more difficult) challenge is the nature of our response variable, when change occurs, we have mislabelled pixels in our response. The new growth was spectrally different to older growth and this caused problems for RF and Matching. However, Matching was more robust to this problem than RF and this speaks strongly for parsimony in modelling.
The cases where RF performed better were;
• the identification of plantations in early stages of growth (when a later base image is available)
• accurate classification in the presence of haze or smoke

The benefits of Matching were;
• computationally cheap (an order of magnitude less than RF)
• more robust to large amounts of change
• models can be fitted manually using standard image processing software (interpretability)

The application of the spatial –temporal model significantly reduces classification errors, but adds computational cost. In practice we use a cluster implementation to perform this step.

3. SOME FOREST CONVERSION STATISTICS

From the time series classification, forest conversion figures may be estimated for regions of interest. Table B presents the national figures, extracted from table 4 of the "national carbon accounting system Greenhouse Gas Emissions from Land Use Change in Australia: Results of the National Carbon Accounting System 1988-2003" (http://www.greenhouse.gov.au/ncas/reports/pubs/ncasresults.pdf). One observes a reduction in conversion since 1989. Landcover change results prior to 1989 are primarily used to give stand age estimates for modelling of carbon flux.

Table B. Rates of Forest Conversion and Reclearing (thousands of hectares).

	National	
Year	Conversion	Reclearing
1989	437.0	160.1
1991	334.7	175.2
1992	294.8	159.9
1995	221.6	162.9
2000	231.9	193.8
2002	154.1*	158.6*

ACKNOWLEDGEMENTS

This work has been supported by the Australian Government Department of Environment and Heritage, through management and coordination of the National Carbon Accounting System Landcover Change Project, details of which can be found at http://www.greenhouse.gov.au/ncas/.

REFERENCES

Breiman, L., 2001. Random forests. Machine Learning 45(1), pp. 5–32.

Breiman, L., Friedman, J. H., Olshen, R. A. and Stone, C. J., 1984. Classification and regression trees. Wadsworth Statistics/ Probability Series, Wadsworth Advanced Books and Software, Belmont, CA.

Caccetta, P. A. (1997), *Remote Sensing, GIS and Bayesian Knowledge-based Methods for Monitoring Land Condition*, PhD thesis, School of Computing , Curtin University of Technology.

Caccetta, P. and Bryant, G., 2002. Notes on the automatic determination of index thresholds for classification of perennial vegetation change. Technical Report 247, CSIRO Mathematical & Information Sciences.

Campbell, N. A. and Atchley, W. R. (1981), 'The geometry of canonical variate analysis', *Syst. Zoology*, Vol. 30, No. 3, pp. 268-280.

Danaher, T., Wu X. and Campbell, N.A. " Bi-directional Reflectance Distribution Function Approaches to Radiometric Calibration of Landsat TM imagery," *Proceedings IGARSS 2001 conference,* Brisbane, Australia, July 2001.

Furby, S.L. and Woodgate, P. (Edited by) (2002), *Remote Sensing Analysis of Land Cover Change - Pilot Testing of Techniques*, National Carbon Accounting System Technical Report No. 16, October 2002.

Furby, S.L. (2002), *Land Cover Change: Specification for Remote Sensing Analysis,* National Carbon Accounting System Technical Report No. 9, November 2002.

Kiiveri, H. T. and Caccetta, P. A (1998), *Image fusion with conditional probability networks for monitoring salinisation of farmland, Digital Signal Processing*, Vol. 8, No.4, pp. 225-230.

Lowell, K., Woodgate, P., Jones, S. and Richards, G. (2003), Continuous Improvement of the national carbon accounting system land cover change mapping, National Carbon Accounting System technical report No. 39. (http://www.greenhouse.gov.au/ncas/reports/pubs/tr39final.pdf).

McCullagh, P. and Nelder, J. A., 1983. Generalized linear models. Monographs on Statistics and Applied Probability, Chapman & Hall, London.

MacKay R.J. and Campbell, N.A. 1982. Variable selection techniques in discriminant analysis 1. Description. British Journal of Mathematical and Statistical Psychology. 35:1-29.

Nelder, J. A. and Mead, R., 1965. A simplex algorithm for function minimization. Computer Journal (7), pp. 308–313.

O'Connell, J. and Caccetta, P. A. (2005), A Parallel Implementation of an Image Processing Algorithm. APAC conference and exhibition on advanced computing, grid applications and eResearch, Royal Pines Resort, Gold Coast, 26 - 30 September 2005

O'Connell, J. and Caccetta, P. A. (2006), Testing of Alternate Classification Procedures within an Operational, Satellite based, Forest Monitoring System, The 13th Australasian Remote Sensing and Photogrammetry Conference Proceedings, Canberra, Australia, 20-24 November, 2006.

R Development Core Team, 2004. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-00-3.

Ripley, B. D., 1996. Pattern recognition and neural networks. Cambridge University Press, Cambridge.

Teillet, P.M., Guindon, B., and Goodenough, D.G., (1982). *On the Slope-Aspect Correction of Multispectral Scanner Data. Can. J.Remote Sensing*, 8(2): 84 – 106.

Vermote, E., D. Tanré, J. L. Deuzé, M. Herman, and J.J. Morcrette, 1994 *Second Simulation of the Satellite Signal in the Solar Spectrum (6S), 6S User Guide Version 0*, April 1994.

Wu, X., Danaher, T., Wallace, J. F. and Campbell, N. A. (2001), *A BRDF-Corrected Landsat 7 Mosaic of the Australian Continent*. IGARSS 2001 conference proceedings, July 2001.

Wu, X., Furby, S. L. and Wallace, J. F. (2004), An Approach for Terrain Illumination Correction, The 12th Australasian Remote Sensing and Photogrammetry Conference Proceedings, Fremantle, Western Australia, 18-22 October.