FOREST AND FOREST CHANGEMAPPING USING SAR DATA AND A ROBUST CLASSIFICATION METHODOLOGY

Anthea L. Mitchell[1*], Mark L. Williams[2], and Anthony K. Milne[1].
[1]Co-operative Research Centre for Spatial Information (CRC-SI), School of Biological, Earth & Environmental Sciences, The University of New South Wales, Kensington NSW 2052, Australia; Tel: +61 2 9385 1622; a.mitchell@unsw.edu.au, t.milne@unsw.edu.au

[2]Co-operative Research Centre for Spatial Information (CRC-SI), 204 Lygon Street, Carlton VIC 3053, Australia; Tel: +61 3 8344 0849; mark.williams@physics.org

KEY WORDS: Forests, Monitoring, ALOS PALSAR, RADARSAT-2

ABSTRACT: Evaluation of the changing extent of forest cover using wall-to-wall mosaics derived from imaging radar requires a robust, yet adaptable approach to classification. In this paper, the unique spatial, polarimetric, spectral and textural properties of dual polarisation ALOS PALSAR data acquired over northeast Tasmania in October 2008 are captured in representative training data and used as input to a classification of forest/non-forest cover using both a manual decision tree approach and the Support Vector Machine (SVM) technique. PALSAR data are multi-looked, ortho-rectified using an available state-wide Digital Elevation Model (DEM) and radiometrically calibrated. Data are then segmented using Definiens Developer software to create homogeneous clusters from which relevant statistical measures are retrieved. In the manual approach, we use the mean backscatter within the HH and HV segments, DEM height, segment shape, area and contextual information in the classification. An overall accuracy of 96.1 % for forest/non-forest was achieved. In the SVM approach, additional input layers are generated through polarimetric correlation and decomposition and texture metrics. Local Incidence Angle (LIA) and DEM derivatives (slope and aspect) are also used as inputs in the SVM classification. Training samples were identified using an existing ground-based classification and field data on a layer stack comprising all data layers, with 70% of training samples used in the initial training phase and 30% used in the testing phase. The libsvm SVM toolset was used to generate a classifier and yielded an overall classification accuracy of greater than 98% for forest/non-forest. Change analysis was applied to detect the difference in distribution of forest/non-forest between the two classifications. The high classification accuracy achieved using both the manual decision tree and SVM approaches confirmed the suitability of ALOS PALSAR data for forest mapping. In terms of computational efficiency however, the SVM approach was deemed superior, although highly dependent on the input training data. Such an approach could form part of an operational monitoring system, fully adaptable to new data as it becomes available, for national scale mapping and monitoring of forest cover.

1. INTRODUCTION

With increased awareness of the contribution of deforestation to greenhouse gas emissions and a leaning toward national accountability of forest and carbon stocks, robust and repeatable methods of satellite based estimation of changing forest cover are required. Consistent estimation of forest cover at national scale requires wall-to-wall time-series measurements of forest and land cover change at a scale appropriate to adaptive management. Up until recently, this has been achieved sporadically in different countries using time-series optical data acquired at moderate (e.g., by the Landsat series) to coarse resolutions (e.g., MODIS). The Landsat archives extend back to the 1970s and provide extensive data for the establishment of historical baselines of forest extent from which annual rates of deforestation can be determined (DeFries *et al.,* 2007). With declining availability of moderate resolution optical data however, data acquired by Synthetic Aperture Radar (SAR) presents a viable alternative to observing change in regional forest cover. SAR also has its advantages in the tropics, where cloud cover and haze are often limiting to detection of continuous land cover.

The Japanese Space Agency (JAXA) launched the ALOS PALSAR in December, 2006, and has since been committed to a systematic earth observation strategy that includes the acquisition of global scale multi-annual fine resolution data, regional scale polarimetry every 2 years, and coarse resolution global coverage annually and multi-annually at selected sites. Despite the unavoidable loss of the ALOS PALSAR instrument in April 2011, there is an extensive archive available from 2007 through to early 2011, providing many opportunities for time-series studies of the world's forested ecosystems. PALSAR-2 is due for launch in late 2012, and coupled with data acquired at shorter wavelengths such as RADARSAT-2 and TerraSAR-X, the continuity of SAR data for regional to global scale monitoring seems ensured.

With such large data volumes available and often a need for integration of multiple data sources, there is increased interest in machine learning techniques such as Artificial Neural Networks (ANN) and Support Vector Machines (SVM) to efficiently solve data analysis and image based classification problems (Kanevski *et al.*, 2009). Machine learning algorithms derive mathematical relationships between input variables for the purpose of predictive learning. The inputs can be numerous and variable, and include raw image data and derivatives. The selection of training data is typically guided by *a priori* knowledge and should foremost be representative of the desired output classes. In this paper, we compare a manual decision tree approach implemented in Definiens Developer and use of the libsvm SVM toolset to map the extent of forest/non-forest in north east Tasmania using data acquired by the ALOS PALSAR. In the manual approach, hierarchical decision trees are constructed using mean HH and HV backscatter, DEM height and spatial and contextual information to classify forest/non-forest. In the SVM approach, training samples are identified on a layer stack comprising PALSAR derived spatial, polarimetric, spectral and textural features. Classification accuracy and computational efficiency are evaluated in the process. Ultimately, and what is required, is a consistent, robust and repeatable method, suitable for regional scale forest monitoring.

Section 2 introduces the study area, typical vegetation and land use. In Section 3 we describe the data processing, Decision tree and SVM approaches to classification, and compare the results. Concluding remarks and opportunities for further work are outlined in Section 4.

2. STUDY SITE

A ~25x25 km subset was extracted from a single ALOS PALSAR scene acquired in north east Tasmania near the township of Mathinna (Figure 1). The landscape in this region is highly dissected with up to 1500 m relief. Forest and land cover is highly variable, with the majority of forested areas comprising Eucalyptus species (Harris and Kitchener, 2005). Highland treeless vegetation inhabits the higher elevations, in particular, Ben Lomond, the prominent rocky plateau in the southern part of the scene. Scrub and Buttongrass moorland are scattered throughout the area. Land use is dominated by commercial forestry and agriculture. Native hardwood (eucalypt species) plantation pre-dominates, with some large areas of softwood (pine) plantation in the north.

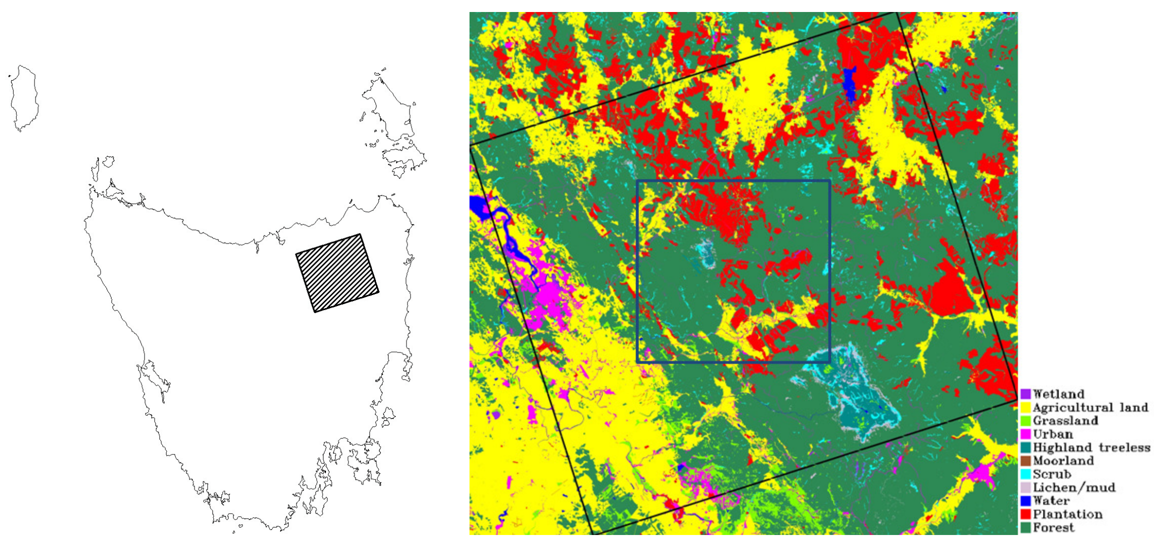

Figure 1. Tasmania outline map with location of the study site, and subset TASVEG state-wide vegetation mapping with PALSAR image extent (black outline) and extracted subset (blue outline) overlain.

3. REMOTE SENSING DATA AND PROCESSING

The PALSAR data used in this study are part of a time-series wall-to-wall archive acquired over Tasmania for the Group on Earth Observations (GEO) Forest Carbon Tracking (FCT) task. A modified version of the processing stream as applied to the entire archive was implemented for the SVM classification.

We have begun to test the SVM technique against the approach previously adopted in this exercise which has been to manually generate hierarchical decision trees using Definiens Developer software. A single scene of Fine Beam Dual (FBD) polarisation ALOS PALSAR data acquired on 4 October 2008 was used. Both the decision tree and SVM approaches start with the L1.1 PALSAR product which is multi-looked, radiometrically scaled and ortho-rectified. Using both approaches, we generate a forest/non-forest classification of the type required for carbon accounting reports.

3.1 Decision tree classification

Using ENVI SARscape, Single Look Complex (SLC) data were multi-looked using 4 looks in azimuth and 1 look in range to produce quasi-square pixels. Data were then ortho-rectified using an available 25 m Digital Elevation Model (DEM), resampled to 12.5 m spatial resolution and radiometrically calibrated and normalised.

Baseline land cover was first mapped using 2007 PALSAR data. An object based approach to classification was implemented using Definiens Developer software. Multi-resolution segmentation was applied, a process whereby pixels are iteratively merged into objects of maximum homogeneity. Segmentation parameters were set using a scale factor of 10 and values of 0.1 and 0.5 for shape and compactness respectively. Cover classes were identified using field knowledge and TASVEG mapping as reference. Rule sets were established for each class using membership functions and/or thresholds relying on spectral (mean HH and HV backscatter) and spatial (context) features. Region growing was applied, where strict thresholds are used to classify segments of definite membership, after which the process runs in a loop, iteratively classifying neighbouring segments that satisfy a given criterion.

Following generation of the land cover map, the classes were merged into ‘forest’ and ‘non-forest’ super classes. The land cover map for 2008 was generated using the 2007 baseline land cover and change detection results. The change maps are used to identify changes in brightness between dates which can, in some instances, be related to a change in land cover. Mean 2007 and 2008 HV backscatter data were used in the change analysis. The relevance of the change, e.g., a decrease in backscatter between dates due to clearing of forest, was confirmed by simultaneous observation of the two image dates and knowledge of on-ground practices. Masks were generated that showed instances of where change had occurred in a particular cover class, and used to update the land cover classification for 2008. A forest/non-forest map was produced as for 2007 above. A subset of the 2008 forest/non-forest classification over the area of interest in this study is shown in Figure 2. Compared to field data and TASVEG, an overall forest/non-forest mapping accuracy of 96.1 % was achieved.

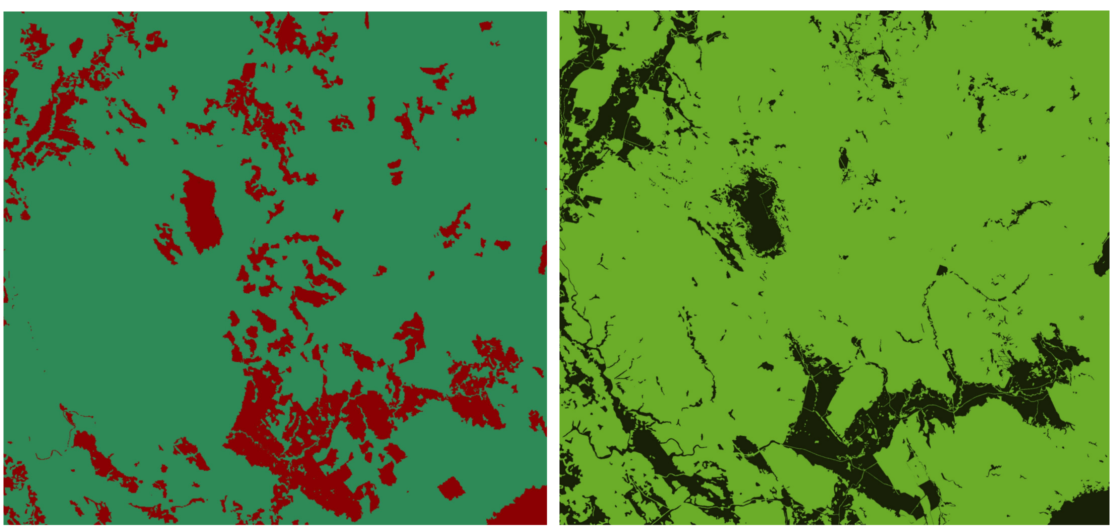

Figure 2. Left) Forest/non-forest classification produced using decision tree approach for the area of interest, and Right) TASVEG forest/non-forest estimate.

3.2 Support Vector Machine (SVM) classification

A Support Vector Machine (SVM) solves for the optimal decision hyperplane based on a trained set of input parameters to maximise the separation between two classes, and multiple classes by extension. Such an approach is attractive since it reduces the classification problem to the generation of correct training data, and relieves the operator from the process of manually determining a decision tree that most likely will be sub-optimal.

In the SVM approach we have in addition sought to exploit textural information in the SAR imagery. In order to preserve this information as much as possible during the pre-processing stages we have chosen to multi-look only twice in azimuth, and ortho-rectify at 6.25m pixel spacing using the high-resolution DEM available in Tasmania.

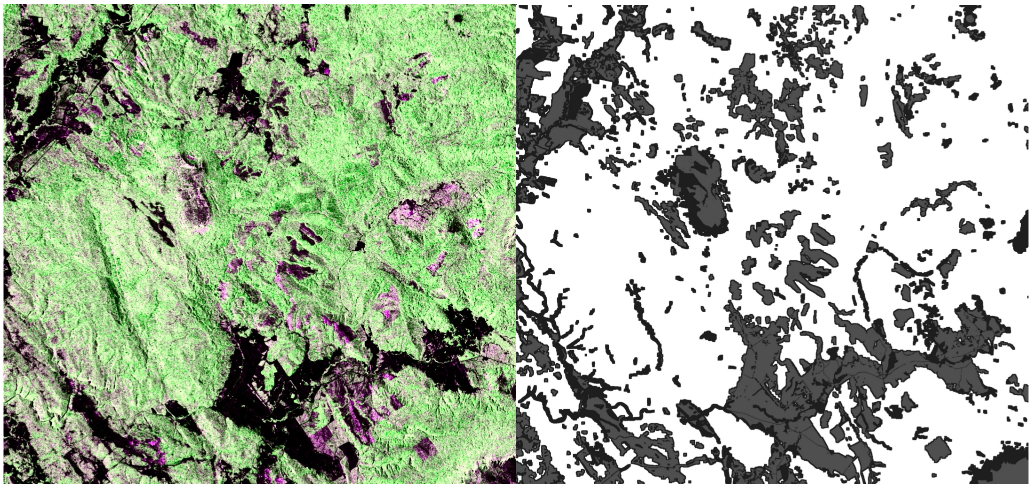

Figure 3. Left) False colour composite formed from L-HH and L-HV orthorectified PALSAR data, and Right) training data class map used for forest (white)/non-forest (grey) classification derived from the TASVEG land cover map.

Training samples for 'forest' and 'non-forest' classes were identified using a reduction of an existing vegetation layer and field knowledge (Figure 3). The PALSAR data were segmented using eCognition software. Training class data were then examined within segments, and those segments having at least 95% of their constituent pixels with the same class were chosen to train the SVM classifier. The total population of such "predominant-class" segments was divided into sets of 70% for training and 30% for testing on a random basis.

Features were then calculated from data layers on a segment basis (Figure 4). Point statistics (mean and standard deviation) of feature layers within segments were calculated for HH and HV magnitude, as well as DEM height, DEM gradient and local incidence angle. The latter helps the classifier isolate residual terrain dependence of backscatter within a single class. Texture features were calculated using the HH and HV magnitude layers. Grey level co-occurrence matrix measures were generated using segment pixels and histogram-equalized 32-level grey images of the HH and HV data. Grey level texture measures included contrast, correlation, energy, entropy, sum average, inverse difference moment, dissimilarity, inverse difference and maximum probability (Haralick, 1979).

Using these features a classifier was generated using the libsvm SVM toolset (http://www.csie.ntu.edu.tw/~cjlin/libsvm) and 70% of the "predominant-class" segments. The classifier was subsequently tested on data from the remaining 30% of "predominant-class" segments and found to yield an overall accuracy of just under 98% for forest/non-forest classes. This level of performance is very high and not only confirms the information content of the PALSAR data of Tasmania for forest reporting, but also the efficacy of the SVM technique.

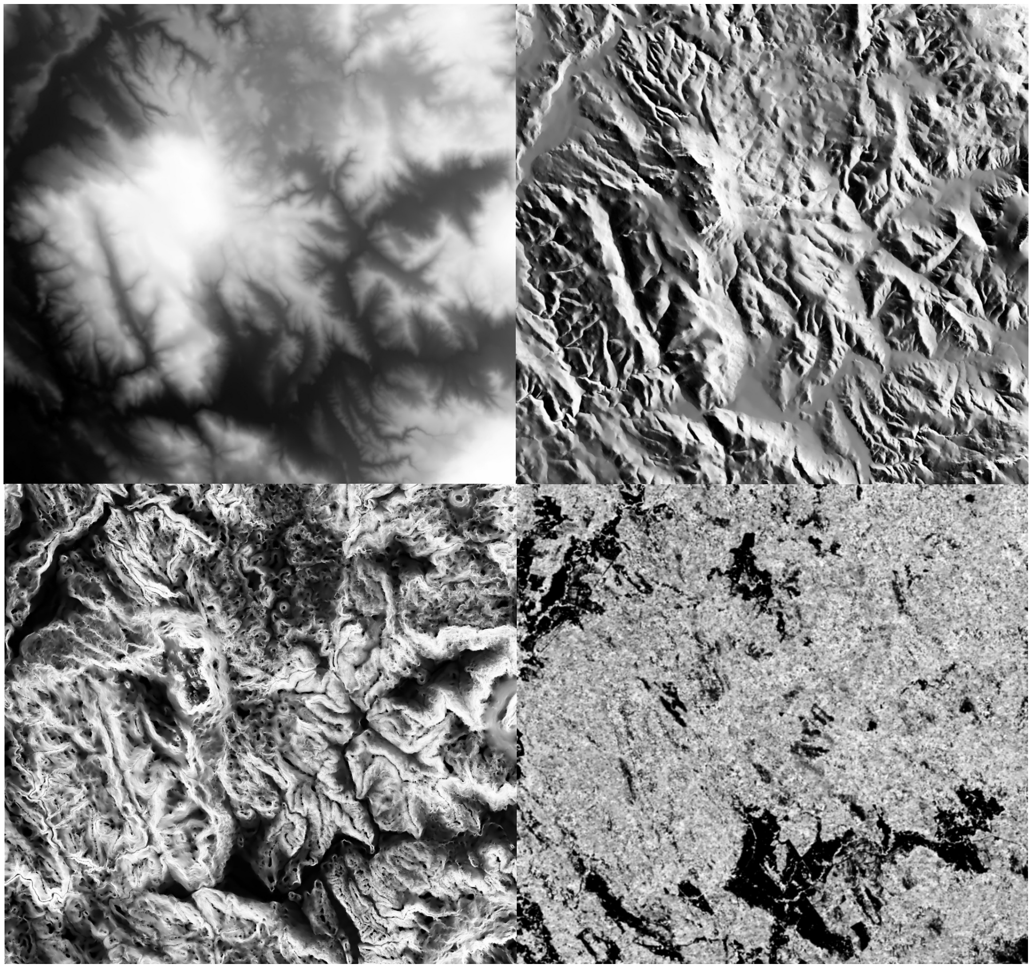

Figure 4. Feature layers used in the SVM classification process. Top left) DEM height, Top right) local incidence angle calculated using SARscape software, Lower left) DEM gradient, and Lower right) GLCM contrast calculated from HV data within segments derived using eCognition.

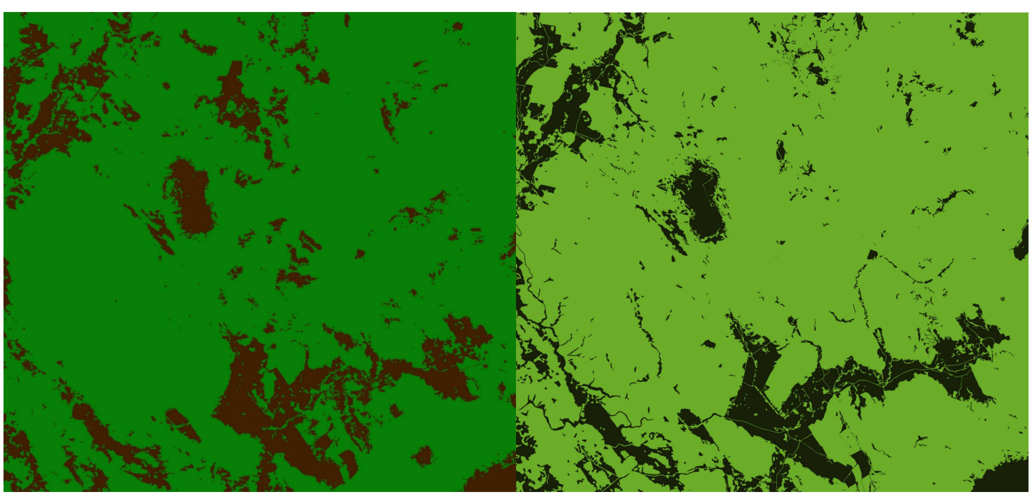

Figure 5. Left) SVM derived forest/non-forest classification for the area of interest, and Right) original TASVEG forest/non-forest estimate.

A good correspondence in forest/non-forest cover was obtained between the SVM derived forest/non-forest classification and that derived using the manually obtained decision tree (Table 1). For the area of interest, forest cover was estimated as 83.9 % in the decision tree classification, 83.8 % in the SVM and 84.2 % in the TASVEG reference layer. While overall forest extent was similar, subtle localised differences in the spatial distribution of forest cover were observed. The differences were due largely to the varying capacity of the two classifiers to discriminate between bare/cleared (considered as non-forest) and regenerating/mature plantation (considered as forest). It is our intent to apply the SVM technique to the question of detecting change, which will be addressed in future work.

Table 1. Forest/non-forest estimates compared for the 3 different layers: Decision tree classification, SVM classification and TASVEG reference.

Layer	**% Forest**	**% Non-forest**
Decision tree	83.9	16.1
SVM	83.8	16.2
TASVEG	84.2	15.8

4. CONCLUDING SUMMARY

The study has reaffirmed the high information content of ALOS PALSAR imagery and suitability for forest cover mapping. Out of the two approaches tested, the strength of the SVM technique, which is reliant upon standard product generation, is in the flexibility of inputs for training the classifier, which are not restricted to PALSAR or other radar data, and the facility to classify large datasets. Future work will expand on the selection of training data in attempts to classify Tasmania-wide forest/non-forest, and incorporate other remote sensing data sources such as RADARSAT-2, TerraSAR-X and Tandem-X data into the classification. It is suggested that such an approach could form part of an operational monitoring system, readily adaptable to newly acquired data, for national scale mapping and monitoring of forest cover.

5. REFERENCES

DeFries, R., Achard, F., Brown, S., Herold, M., Murdiyarso, D., Schlamadinger, B. and de Souza, C. Jr. 2007. Earth observations for estimating greenhouse gas emissions from deforestation in tropical countries. Environmental Science and Policy, 10, pp. 385-394.

Haralick, R.M. 1979. Statistical and Structural Approaches to Texture. Proc. IEEE, vol. 67, pp. 786-804.

Harris, S. and Kitchener, A. 2005. From forest to fjaeldmark: descriptions of Tasmania's vegetation. Department of Primary Industries, Parks, Water and Environment, Printing Authority of Tasmania, Hobart.

Kanevski, M., Pozdnoukhov, A. and Timonin, V. 2009. Machine Learning For Spatial Environmental Data. Theory, Applications And Software. EPFL Press, Switzerland.