Mapping candidate native non-woody extent and condition in the Central West-Lachlan CMA regions (winter 2008 – summer 2009)

Adrian Fisher, Graham Turner and Majid Mirbagheri
Remote Sensing Unit
Department of Environment, Climate Change and Water

First draft – 15/9/2009

1. Introduction

The main principle in mapping candidate native non-woody is to remove land that can be identified as having alternative land cover types such as woody or cropping. This can be done through a combination of supervised classification, NDVI analysis and using other complementary data. The mapping process involved three main steps: 1) segmenting the region into polygons according to landcover homogeneity; 2) attributing the polygons with data derived from relevant raster data, and; 3) classifying the polygons into landcover categories based on the attribute data.

2. Method

2.1 Object oriented segmentation

Paddock scale polygons were produced through image segmentation of 2.5 m resolution SPOT data in Definiens. This was performed on unmosaiced spot scenes within 250K map sheets, resulting in many edge effects. The polygon shapefiles were then subjected to the following editing in ArcGIS:

- Smoothed using the generalise tool.
- Clipped to the Central West-Lachlan CMA boundary.
- Polygons with area < 2500 m^2 were eliminated through being merged with the neighbouring polygon that shared the longest common border.

Some of the smaller shapefiles were merged resulting in the following 16 tiles:

- Balranald-Hay
- Bathurst-Sydney
- Booligal-Pooncarie
- Cargelligo-Narrandera
- Cobar-Bourke
- Cootamundra
- Dubbo-Singleton
- Forbes
- Gilgandra-Tamworth
- Goulburn
- Ivanhoe-Manara
- Narrabri
- Narromine
- Nymagee
- Nyngan
- Walgett

Each polygon in each shapefile was also attributed with:

- Id (a unique identifier for each polygon within each shapefile)
- Perimeter (metres)
- Area (metres squared)

2.2 Supervised classification

A supervised classification of winter 2008 and summer 09 Landsat TM imagery was performed, using field sites as training areas. Initial classifications attempted to separate many different classes, however it was soon realised that spectral variation within each class was too great. This led to a simplified classification being performed using the following three classes: cultivated, woody and unclassified. The classification was as conservative as possible, to ensure that false positives were kept to a minimum (< 5%), which resulted in large areas remaining unclassified.

Data preparation for the classification followed these steps:

- All field data was entered into a shapfile (cwl_field_data.shp). Separate shapefiles containing the field data present in each Landsat TM scene were then created.
- Each field data shapefile was simplified into 5 classes: pasture, modified pasture, bare ground, cropping and woody. A separate shapefile was then created for each class, to make it easier to import and view in Imagine.
- Each field data shapefile was then simplified further into 3 classes: cultivated (modified pasture, bare ground and cropping), woody and unclassified. These classes were recorded as a numeric code to facilitate the accuracy assessment.
- The Landsat TM scenes were then prepared for the classification. This included the following: improving the rectification; reprojecting to MGA Zone 55; and clipping to the Central West-Lachlan CMA boundary. Summer 2009 scenes were also resampled into 25 m pixels to be consistent with the other data. The following images were used:

2008 winter	2009 summer
9081_270908_mga55_cwl.img	9081_060309_mga55_cwl.img
9083_270908_mga55_cwl.img	9083_060309_mga55_cwl.img
9181_200108_mga55_cwl.img**	9181_250209_mga55_cwl.img
9183_180908_mga55_cwl.img	9183_250209_mga55_cwl.img
9281_250908_mga55_cwl.img	9281_150109_mga55_cwl.img
9283_250908_mga55_cwl.img	9283_310109_mga55_cwl.img
9383_160908_mga55_cwl.img	9383_070209_mga55_cwl.img*
9483_230908_mga55_cwl.img*	9483_290109_mga55_cwl.img*

- The 3 images marked with an asterisk above did not contain any field data points and were classified according to a slightly different method. This involved defining signature areas in regions of obvious cultivation (rectangular paddocks with a very high near infra-red response) or woody (dark response with mottled texture and higher relief). Accuracy assessment for these images was only performed through visual checking.
- The October image marked with a double asterisk above appeared to contain different response in certain areas that overlapped with other scenes. Some paddocks that were clearly being cropped in a September scene looked bare in the October scene, while the opposite was also observed. To check that these changes were real and not due to the image being the wrong date, MODIS NDVI and EVI images from 16-30 October were examined. While the coarse pixel size prevented a paddock scale assessment, it appeared to confirm that the October Scene was not wrong.

The classification procedure followed these steps:

- Each image was classified separately, using some field data points as seed pixels for spectral signatures. The accuracy of the classification was assessed against all field data points.
- In the Erdas Imagine Classification Tool Signature Editor, signature pixels were defined for classes, using seed pixels identified from field data. Extra signatures were also defined for other obvious classes (e.g. Urban, mining, clouds, water, woody forests and clear-fell).
- Signatures were then evaluated using the contingency matrix (maximum likelihood) and separability (Transformed Divergence) tools. Where problems were identified signatures were modified or deleted.

- A parametric classification was run using the Maximum Likelihood rule, and the result compared to the original false colour image. Problem areas were identified visually, and signatures were added or deleted as necessary. The main problem faced by the classification was a similarity between bare ground, cropping, pasture and modified pasture that had a strong visible red (band 3) response, likely due to bare soil showing.
- Once no further problems could be identified, the classified image was modified using the Threshold tool, where the maximum likelihood distance image was used to re-class any pixel too far from its signature as unclassified (the confidence level was varied between 0.05 and 0.001 in an effort to be conservative but also allow sufficient pixels to remain classified).
- The threshold image was then recoded into three classes: cultivated, woody, and unclassified. The recoding was performed such that the two main classes were defined with a minimum extent.
- An accuracy assessment was then carried out against the field data. If false positives were greater than 5% the thresholding tool was re-run to reduce the number of classified pixels. Statistical accuracy helped guide the classification process but was not a priority for two reasons:
    1. The classification procedure has been designed to be as conservative as possible, with cultivated and woody pixels being intentionally under-classified. Minimising false positives was deemed to be more important than classification accuracy.
    2. The classified pixels will be used to attribute polygons, and accuracy should really be assessed on the polygon attributes. This will allow thresholds to be set to the polygon attributes which will improve the classification.

The resulting raster images were then recoded into binary format, with a separate image being produced for the woody and cultivated classes.

2.3 Classification of NDVI images

Each of the winter Landsat TM scenes was classified into cropping and modified pasture through applying a NDVI threshold listed in the table below. The summer 2009 scenes were only classified into cropping as modified pasture was not distinguishable.

		NDVI threshold	
Season	**Image**	**Cropping**	**Modified pasture**
Summer 2009	9081_060309	0.48	
	9083_060309	0.41	
	9181_250209	0.40	
	9183_250209	0.38	
	9282_040309	0.37	
	9283_310109	0.34	
	9383_070209	0.40	
Winter 2008	9081_270908	0.53	0.63
	9083_270908	0.45	0.53
	9181_201008	0.48	0.54
	9183_180908	0.48	0.61
	9281_250908	0.42	0.51
	9283_250908	0.44	0.55
	9383_160908	0.35	0.44
	9483_230908	0.44	0.51

The resulting raster images were then recoded into binary format, with a separate image being produced for the cropping and modified pasture classes.

2.4 Digital elevation models

As slope is an important factor in determining land use, it may assist in discerning areas of cropping. Unfortunately, as the 25 m NSW DEM covers only the eastern section of the CMA

region, the 250 m NSW DEM had to be used in the western section. Slope grids in degrees were generated in Imagine (Image Interpreter > Topographic analysis > Slope). The 25 m slope grid had to then be edited to remove unwanted effects from the western edge of the grid. To ensure that every polygon was attributed with an elevation value, the 250 m elevation and slope grid was resampled to 25 m pixel size before being used in the zonal statistics commands that are outlined in the following section.

2.5 Attributing polygons

The following method using ESRI ArcGIS was developed to extract statistics from the raster files and populate attributes of the polygon shapefiles.

- Each polygon shapefile was converted to a 25 m pixel raster (.img format) based on the unique polygon id field using the PolygonToRaster command at the ArcCatalog command line. This raster was then checked to ensure all polygons had survived the conversion process.
- The "ZonalStatisticsAsTable" command was then run at the ArcCatalog command line for each rasterised polygon image with each input raster, using the NODATA option. This produced a series of tables that were named according to the input raster.
- After adding the polygon shapefile and all associated tables to map document, the following commands were processed at the ArcMap command line. A new field called PERCENT was added to each table that had been produced from a classified image. This attribute was then calcaulated as (([SUM]/[COUNT])/0.01) to give the percent of pixels within each polygon that were classified (which is the same as percent area as all pixels are 25 m). New fields were then added to the polygon shapefile to contain the relevant data from each table. Each table was joined to the shapefile, the field calculated, then the join removed.
- Below is a sample of the commands that were run, with path names removed to make the commands easier to read. The full list of commands is listed in the Excel spreadsheet commands.xls.

In ArcCatalog:

```
PolygonToRaster Forbes.shp ID Forbes.img # # 25
ZonalStatisticsAsTable Forbes.img Value 9283_250908_cultivated.img 9283_250908_cultivated.dbf NODATA
```

In ArcMap:

```
AddField 9283_250908_cultivated.dbf PERCENT DOUBLE 10 4
SelectLayerbyAttribute 9283_250908_cultivated NEW_SELECTION ("SUM">0)
CalculateField 9283_250908_cultivated PERCENT "(([SUM]/[COUNT])/0.01)"
AddJoin Forbes id 9283_250908_cultivated Value
SelectLayerByAttribute Forbes NEW_SELECTION "[9283_250908_cultivated.PERCENT] IS NOT NULL"
CalculateField Forbes cultiv_1 "[9283_250908_cultivated.PERCENT]"
RemoveJoin Forbes 9283_250908_cultivated
```

- This resulted in each polygon files having the following attributes:

Attribute	Description
area	Area of the polygon in metres squared
perimeter	Perimeter of the polygon in metres
id	Unique identification number for each polygon
dem_25	Mean elevation for each polygon in metres
slope_25	Mean slope for each polygon in degrees
vegtype	Majority class within each polygon from the interim vegetation map of NSW
ndvi_dif_1	Mean difference in NDVI between summer 2009 and winter 2008 images
cultiv_1*	Percent of each polygon classified as cultivated in one TM scene (supervised classification)
woody_1*	Percent of each polygon classified as woody in one scene (supervised classification)
modified_1*	Percent of each polygon classified as modified pasture in one scene (NDVI threshold)
cropping_1*	Percent of each polygon classified as cropping in one scene (NDVI threshold)
ndvi_w08_1*	Mean NDVI value from one winter 2008 Landsat TM scene
ndvi_s09_1*	Mean NDVI value from one summer 2009 Landsat TM scene
SLATS_2008	Percent of each polygon classified as woody in the SLATS 2008 results

* Note: The number suffix ranges from 1 to 8 corresponding to the number of Landsat TM scenes that are located with each polygon shapefile.

Four new summary attributes were added and calculated as the maximum of each numbered field.

Attribute	Description	Source
CULTIV	Maximum value of cultiv_1, 2, 3 and 4	supervised classification
WOODY	Maximum value of woody_1, 2, 3 and 4	supervised classification
MODIFIED	Maximum value of modified_1 and 2	NDVI threshold
CROPPING	Maximum value of cropping_1, 2, 3 and 4	NDVI threshold

Three new fields were also added to estimate the percentage of cropping and modified pasture in each paddock for each season. Means were calculated only from those polygons that had two non-zero values.

Attribute	Description
CROP_2008	If [cropping_1] EQ 0 then [cropping_2]; If [cropping_2] EQ 0 then [cropping_1]; If [cropping_1] > 0 AND [cropping_2] > 0 then ([cropping_1]+[cropping_2])/2
MOD_2008	If [modified_1] EQ 0 then [modified_2]; If [modified_2] EQ 0 then [modified_1]; If [modified_1] > 0 AND [modified_2] > 0 then ([modified_1+[modified_2])/2
CROP_2009	If [cropping_3] EQ 0 then [cropping_4]; If [cropping_4] EQ 0 then [cropping_3]; If [cropping_3] > 0 AND [cropping_4] > 0 then ([cropping_3]+[cropping_4])/2

A 2.5 m resolution SPOT5 image (ecw) from 2008 was subjected to an unsupervised classification, where it was classified into 50 clusters. Visual inspection of the clusters allowed the 50 clusters to be reclassified into 4 classes: crop, pasture, perennials and woody. Binary images of these 4 classes were produced, and using the "zonal statistics as table" command in ArcGIS the percentage of each class within each polygon was calculated.

Attribute	Description
s5_crop	Percent of each polygon classed as cropping in the SPOT5 image
s5_pasture	Percent of each polygon classed as pasture in the SPOT5 image
s5_perenni	Percent of each polygon classed as perennials in the SPOT5 image
s5_woody	Percent of each polygon classed as woody in the SPOT5 image

2.5 Preliminary classification of polygons in the Forbes map sheet

First classification

The field CLASS_1 was created, which labelled each polygon with the following simple classes:

Woody
Cropping
Modified