Appendix C-21

Classification of Vine Thicket on Dampier Peninsula using TM Imagery
Report to Woodside Energy Limited

Classification of Vine Thicket on Dampier Peninsula using TM Imagery

February 2010

Jeremy Wallace
CSIRO Mathematics Informatics and Statistics
Floreat WA 6014
Jeremy.Wallace@csiro.au

Note. This report summarises analysis and validation data of a study conducted for Woodside Energy Limited. It is intended as a summary of that work for limited circulation only. It is not a fully referenced or peer-reviewed scientific paper. Please contact the author before reproducing or citing material in this report in digital or print form.
Classification of Vine Thicket on Dampier Peninsula using TM Imagery

Executive Summary

The objective of this work was to provide

- a broad regional mapping of the extent of vine thicket vegetation on the Dampier Peninsula, and
- an estimate of the area of vine thicket vegetation beyond the limited areas which have been mapped previously in fine detail.

The James Price Point area had been mapped in detailed survey by Biota Environmental Sciences (2009). Packer Island and Perpendicular Head had been mapped by ENV Australia (2008).

Landsat imagery was analysed and processed for this purpose. Independent validation of the mapping was carried out by a vegetation expert using high resolution photography and targeted ground visits.

Results:

- Using the Landsat imagery, a total area of 830 hectares was classified as vine thicket on the northern region of the peninsula
- The validation process showed that this area estimate was very conservative. The validation produced spatially detailed maps for the areas which were checked
- Classification and validation data have been archived in GIS.

Findings and recommendations:

Effective application of remote sensing to provide vegetation information requires collaboration between remote sensing analysts and vegetation experts in the context of clearly articulated information questions – here the area and location of vine thicket.

Variation of spectral response within vegetation types is a physical reality. It resulted in inaccurate classification when only single-date Landsat imagery was used. Images also varied with seasonal condition. Classification results from higher resolution imagery are likely to be affected by variation in a similar way.

Combining results from three carefully selected Landsat dates produced a mapping of vine thicket which was conservative in area with few commission errors. This approach was restricted to the northern region on the basis of available ground information. It provided an area statement and a basis for validation and improvement by vegetation experts.

The mapping and results are accurate but conservative in area. To produce a comprehensive baseline map, the validation process could be progressively extended – this approach combined Landsat classification as a detector of possible vine thicket with expert interpretation of photography to define precise boundaries.

For future management and reporting of changes in VT, monitoring becomes possible once an agreed baseline map is established. To provide ongoing monitoring of changes in vine thicket over broad areas, consistent time series data is essential. Landsat imagery has been demonstrated to be an effective tool for monitoring and a long time series exists. Continuity of Landsat-scale imagery or cross-calibration of historic series with newer high-resolution imagery should be considered for ongoing monitoring.
Classification of Vine Thicket on Dampier Peninsula using TM Imagery

The aim of the study was to produce a spatial mapping and area estimate of vine thicket extent on the Dampier peninsula using Landsat TM imagery, and to provide assessment of results. In practice, the analysis was guided by existing knowledge and ground surveys of vegetation type and focussed on the northern area of the peninsula.

1. Study area and ground-survey training data.

The region of interest was the Dampier Peninsula (Figure 1). Monsoon vine thicket vegetation occurs as semi-deciduous and evergreen vine thicket communities on landward slopes of coastal sand dunes on the Dampier Peninsula.

Two main sources of located ground vegetation data were available to train and assess the classification:

1. Detailed spatial mapping in the immediate vicinity of James Price Point (JPP) produced by Biota [1] in the wet season survey of 2009 (Figure1, Figure2).
2. Point location data on threatened ecological communities (TEC) recorded by DEC from survey data from the peninsula. These data record occurrence of TEC’s, principally vine thicket, but do not provide indication of spatial extent. These data are concentrated in the coastal regions in the north of the Peninsula (Figure 1). In the northern region, vegetation at Packer Island and Perpendicular Head had also been mapped by ENV Australia in 2008 [2,3].

Figure 1. Landsat 2008 mosaic of the Dampier Peninsula. Area shown is 170 by 180km. White rectangles show the locations of the JPP area and the northern rectangle which were classified using TM imagery. Pink squares show point locations of Threatened Ecological Communities from the DEC database.
Reports from these surveys, and other documentation on vegetation mapping was provided by Woodside Energy Limited (WEL). These contain lists and descriptions of mapped vegetation types. Vine thicket vegetation is described in two categories in the Biota report; evergreen vine thicket (EVT) and deciduous vine thicket (DVT).

The Northern Peninsula area classified in the study (shown in Figure 1) is approx 54 by 54km in extent and includes a land area of approximately 120,000 hectares. As spatial mapping of the JPP area already exists, this was area was used to test and evaluate methods for classification of the Northern Peninsula region.

2. Landsat Imagery and considerations for classification

Landsat imagery available for the study were:

- 2008 Landsat 7 ETM ‘Composite Mosaics’ processed by Geoimage P/L for WEL.
- Historic Landsat 5 TM and Landsat 7 ETM mosaics processed and calibrated to consistent standards for 10 time periods 1989-2006 from the DCC NCAS archive.

The imagery is resampled to 25m pixels and corrected to standards specified in the NCAS system [4]. See Appendix for more details.

The analysis and classification concentrated on images from 1989, 1992 and 2008. Visual inspection of all images was first carried out along with the ground data. As noted, most extensive vegetation types (e.g. Pindan) were highly variable in space and time. EVT vegetation appeared quite consistent with a strong ‘green’ response in all images – however some pixels within the mapped EVT at JPP had a bright ‘sand-like’ response indicating very little vegetation cover (Figure 3). The DVT however was highly variable and at all dates contained areas of similar spectral response to some of the Pindan and other
vegetation types. It was concluded that accurate classification of DVT was unlikely to be achievable, but that EVT classification was possible and would be most successful in images where green responses were minimal in other vegetation types. The 1989 and 1992 mosaics were judged most suitable. These and the 2008 imagery (which was most recent and which corresponded most closely with ground data collection) were analysed and classified as described in Section 3.

This issue of spectral variation within and between vegetation types is a physical reality which affects the accuracy of digital classification and the selection of an appropriate methodology. The seasonal conditions at the time of overpass greatly influence the spectral signal of areas of vegetation – particularly through variations in greenness associated with rainfall or phenology, and through impact of fire. The small format images in Appendix 1 provide some indication of the temporal and spatial image variation.

These variations in spectral response significantly affect the ability to classify vegetation types – especially with a single-date image. Image classification relies on recognising relatively consistent numerical spectral signature(s) associated with a ground cover or vegetation type which are different ("spectrally separable") from those of other cover types. If two different vegetation types have similar spectral responses, then the classification will be unable to produce an accurate mapping of both. If mapped vegetation types are variable in spectral response, then ‘overlaps’ of this kind are more likely. It can be seen in Figure 2 and Figure 3 that vegetation polygons are in general not strongly associated with consistent colours and many are highly variable in this 2008 image.

Statistical analysis tools (discriminant analysis etc) exist to quantify the spectral separation of ground cover types in defined training areas. These methods were applied to selected images prior to classification, based on the ground information. Ground data is crucial to evaluating the spectral separability of classes (e.g. vine thicket vs other vegetation types). Spatially accurate & representative data are required for training and validation. Here spatial maps were only available for a very local area around JPP. Interpretation of the site TEC data and the imagery provided a basis for application of the approach in the north.

3. Classification and Validation methods

A summary of the classification methods is provided here. Please contact the author for further information, references relating to these methods. The basic steps were

1. Inspect image and ground truth data, select best images for classification of EVT (images from 1989, 1992, 2008 were selected)
2. Select areas for classification on the basis of ground data for analysis and assessment (JPP area and the northern region were selected). See Figure 3 for an example.
3. Conduct discriminant analyses (canonical variate analysis CVA) of each selected image (3 dates) and area (JPP and North) to determine spectral classes relating to the ground training classes. These classes are based on the clustering and separation of the training site spectral responses and included ‘EVT’ classes and classes where EVT signatures overlapped with other vegetation labels (‘Mixed EVT’), and other non-EVT classes. See Figure 4.
4. Produce enhanced maximum likelihood classification of each image based on the classes defined in Step 3. The results retain measures of confidence of class membership of each pixel, and also identify areas for which are ‘atypical’ of all classes.
5. Display and assess the resulting classification map to determine which (if any) of these classes correspond to known EVT areas and associated errors of omission and commission. Rerun from step 3 as necessary.
6. Combine classifications over time using rules based on assessments of classes from each of the three years.
7. Produce maps of classified EVT pixels and area statements
8. Validation

**Figure 3.** Detail of JPP area, 2008 image, 8km by 8km. Lines are Biota survey vegetation boundaries, Red polygons are training sites selected for discriminant analysis. Yellow lines show mapped EVT extents (Biota), while DVT is bounded in orange. Note the variation within most polygons, and white ‘sandy’ areas which occur within the mapped EVT area.

Validation of the classification results in the northern area was carried out using interpretation of rectified digitised air photography from November 2007. As the accuracy of the area statement of EVT was of most interest, random point-based assessment was not undertaken. The validation was conducted in November 2009 by the vegetation expert, Russell Barrett, who also conducted subsequent field visits to confirm particular locations. A number of rectangular areas were selected and the spatial extent of all EVT within those frames was interpreted on hard copy prints of the air photographs. The interpretation was conducted without knowledge of the TM classification results and thus provided an independent spatial mapping for validation of EVT areas. Interpreted EVT polygons were classed as dense or sparse in this process. The polygons were introduced to GIS and the areas compared within each frame with the TM classification. In total, areas covering over 7000 ha were processed for validation; these included approximately 60% of the TM-classified EVT.

4. Results – Analysis and classification

A typical example of an ordination plot from the CV analysis for one year is shown in Figure 4 below. A group (but not all) of the EVT training sites are clustered and largely separate from other vegetation training sites indicating that classification might produce reliable results for this type of EVT (the dense green sites). Three DVT sites are spectrally separate from this cluster. Figure 5 (schematic only) shows how these classes were grouped
to produce a 5-class classification from the 1992 image. Two classes (green circles) contain EVT sites. The results for each class are then assessed (Figure 6).

Figure 4. Canonical Variate Plot. CV1 and CV2 means of analysis of northern area sites from the 1992 image. Sites which are distant from one another are spectrally separable. Colours indicate vegetation type. Green: EVT training sites; blue: mangrove; red: interpreted DVT sites; black: Other non-VT vegetation.

Figure 5. Schematic diagram showing grouping of sites for classification into 5 spectrally similar information classes on the basis of the CV ordination. EVT sites dominate two groups, while two ‘non-EVT’ groups are shown in red and orange.
Figure 6 illustrates a single year classification (here 1992) for a sample area. The two classes with EVT-like signatures are shown in green. The dark green class maps most of the dense green EVT, but some obvious errors of commission occur inland on seasonally green pixels. The light green class has mapped some sparser EVT vegetation (as in the training sites) but has large areas of commission. While such a map may be useful to indicate possible presence of EVT, these commission errors would result in overestimates of EVT area if this single-date classification was used.

Methods to reduce these single-date commission errors with regard to the area-estimation question were considered and examined. Errors from non-EVT ‘green’ vegetation varied in location over time and could be removed by using multiple years. Combination of class results over years using rules were tested and assessed and simple rule-based approach produced an EVT mapping which was visually assessed as reliable but conservative in area.

Using visual assessment (as described above) each spectral class in each of the three years was assessed and scored as either:
1. Likely – mapped EVT with small commission errors (score 2)
2. Possible – mapped some EVT, considerable commission error (score 1)
3. Non-EVT – other vegetation (score 0)

The total score across three years was then calculated (e.g. a score of 6 located pixels mapped as likely on all three years) and displayed (Figure 7). More sophisticated approaches to integration of multi-date classification are possible but were not applied because of the nature of the ground data.
Visual assessment indicated that the areas of score 5 or 6 (i.e., a spectral signature with EVT over all three dates) produced accurate but conservative maps of EVT extent (i.e., few and small commission errors with some omission errors). Within the northern region these pixels totalled 830 hectares (Figure 8). Pixels with a score of 4 (pink in Figures 7 and 8) totalled a further 1580 hectares but it is almost certain that this includes considerable errors of commission. It was suggested that quick assessment by ecologists of identified areas could remove the errors.

It was concluded that 830 hectares represented a conservative estimate of EVT in the northern region.

Within the area, approximately 3200ha had been classified as mangrove in an earlier study (G Behn DEC). This mapping of mangrove was used in the final product.
5. Validation of EVT Area Classification – northern Peninsula

Validation of the classified areas was carried out to identify errors of commission and omission, and to provide estimates of confidence in the area estimate. The validation process was carried out as follows and as illustrated in Figure 9, results in terms of area comparisons are summarised in Table 1.

- Rectangular areas containing classified EVT were defined for validation (Figure 8)
- Air photographs were printed at large scale with a 100m (1ha) grid overlaid
- Independent interpretation of Vine Thicket areas within all of these frames was made by a vegetation expert familiar with the region (Russell Barrett). VT polygons were identified as dense or sparse in this process
- Both maps (TM classification and Interpretation) were imported into ARCinfo GIS by WEL, overlaid and the areas calculated and compared (Table 1). This process allows a comparison of total areas or areas by Frame, which can then be examined spatially on a pixel basis to identify discrepancies and reasons for these (see e.g. Figure 9).
Subsequently, targeted field visits were conducted by Russell Barrett to confirm interpretation boundaries.

In total, the validation areas represented approximately 7000 hectares of land area and included 522 ha of TM-classified EVT; i.e. over 60% of the TM-classified EVT. Within this total area, the interpretation mapped 738 ha of Vine thicket. In comparison with the validation data, the classification area was conservative (i.e. underestimated the area) in all validation areas, as summarised in Table 1.

Table 1: Comparison of the areas of TM-classified Vine Thicket and areas mapped by expert photo-interpretation for validation areas

<table>
<thead>
<tr>
<th>Validation Area</th>
<th>Area in hectares: TM classification</th>
<th>Area in hectares: Photo interpretation</th>
<th>Difference (ha) Photo interp – TM classification</th>
<th>Difference As % of Interpreted VT Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area 1</td>
<td>194.3</td>
<td>221.2</td>
<td>26.9</td>
<td>12 %</td>
</tr>
<tr>
<td>Area 2</td>
<td>129.8</td>
<td>141.2</td>
<td>11.4</td>
<td>8 %</td>
</tr>
<tr>
<td>Area 3</td>
<td>106.8</td>
<td>153.2</td>
<td>46.4</td>
<td>30 %</td>
</tr>
<tr>
<td>Area 4</td>
<td>52.4</td>
<td>161.2</td>
<td>108.8</td>
<td>68 %</td>
</tr>
<tr>
<td>Area 5</td>
<td>39.5</td>
<td>61.5</td>
<td>22.0</td>
<td>36 %</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>522.8</strong></td>
<td><strong>738.3</strong></td>
<td><strong>215.5</strong></td>
<td><strong>29 %</strong></td>
</tr>
</tbody>
</table>
In general, the TM classification omitted or underestimated sparse and small patches of VT, while classification of dense EVT agreed quite closely with the validation data. On the north-east of the peninsula (validation Area 4 in Table 1) less than half of the true VT area was classified. In this area there are relatively large areas of ‘sparse’ vine thicket. A large portion of pixels from this vegetation scored 4 on the combined classification (pink in figure 6) and so were not included in the area tally. Errors of commission did occur; see e.g. top of Figure 9. Commission errors were associated with dense green vegetation which was spectrally similar to Vine Thicket on all three dates but which was a different vegetation type. Greenness of pindan vegetation varies with season. On any single image, significant areas of pindan showed green signatures (see Figure 3). The temporal combination rules resulted in these commission errors being generally small and less than the omissions in all the validation areas. From validation, it can be concluded that the figure of 830 ha is a conservative lower bound for the area of VT on the northern Peninsula.

Conclusion and Recommendations

The application of the classification approach to the three selected image dates in the northern area produced a classification of EVT totalling 830 hectares. Validation confirmed that this area estimate was conservative, underestimating the ‘true’ VT area by an average of 29% over the total of all validation areas, and by at least 8% in all validated areas.

Thus omission errors were considerable in the classification. Commission errors were small and related to pixels with consistent green vegetation spectral signals in the images which were processed. Areas classified as ‘score 4’ by the approach were not included in the area total as they included errors of commission. This class did however indicate a large number of areas of sparse VT and thus could be use to guide detailed mapping as conducted in the validation process.

The validation process targeting areas classified produced detailed mapping quite rapidly around the TM-classified EVT areas. The results of both approaches are stored in GIS and can be overlaid and compared.

The combined classification-validation approach described here has been effective in producing a conservative mapping and area estimate of EVT. The findings highlight the limitations of digital classification alone, but suggest a process to produce a comprehensive and accurate baseline by combining image processing and vegetation expertise. Landsat TM classification might be improved using other dates, improved ground data and stratification. More sophisticated temporal approaches are also possible but would require stratification and ground truth. However, if more comprehensive and accurate mapping is required, it is suggested that the most appropriate practical approach would be to build on the validation approach here. That is, to use multi-date Landsat classification to identify likely or possible EVT areas, and then to apply expert photo interpretation to define boundaries, with targeted ground visits as required. High resolution digital imagery (spatial or spectral) was not available for this study. Such imagery, used by vegetation experts, is likely to assist in the validation process. Digital classification of high resolution imagery is likely to be affected by the problems of variation described here. This author considers that the new digital airborne imagery, acquired under suitable specifications and carefully processed, would be most suitable for evaluation.

The study has provided a digital classification of EVT on the northern Dampier peninsula, and demonstrated a rapid validation approach. The classified ‘EVT’ area of 830ha in the northern region is conservative.
References


APPENDIX – TM imagery, scene boundaries and image dates: Landsat 5 from NCAS and 2008 LS7 (supplied by GeoImage)

<table>
<thead>
<tr>
<th>Nominal NCAS Epoch</th>
<th>Path/Row 111/72</th>
<th>Path/Row 110/71</th>
<th>Path/Row 110/72</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989</td>
<td>27 Sep 89</td>
<td>20 Sep 89</td>
<td>07 Nov 89</td>
</tr>
<tr>
<td>1991</td>
<td>12 May 91</td>
<td>21 May 91</td>
<td>21 May 91</td>
</tr>
<tr>
<td>1992</td>
<td>17 Jul 92</td>
<td>24 Jun 92</td>
<td>24 Jun 92</td>
</tr>
<tr>
<td>1995</td>
<td>21 Jun 94</td>
<td>14 Jun 94</td>
<td>14 Jun 94</td>
</tr>
<tr>
<td>2000</td>
<td>15 Sep 99</td>
<td>07 Aug 99</td>
<td>24 Sep 99</td>
</tr>
<tr>
<td>2002</td>
<td>05 Jul 02</td>
<td>14 Jul 02</td>
<td>14 Jul 02</td>
</tr>
<tr>
<td>2004</td>
<td>16 Jun 04</td>
<td>25 Jun 04</td>
<td>25 Jun 04</td>
</tr>
<tr>
<td>2005</td>
<td>05 Jul 05</td>
<td>30 Jul 05</td>
<td>30 Jul 05</td>
</tr>
<tr>
<td>2006</td>
<td>24 Jul 06</td>
<td>17 Jul 06</td>
<td>17 Jul 06</td>
</tr>
</tbody>
</table>

2008**

Table: Image dates for Landsat scenes from the DCC NCAS archive covering the Dampier Peninsula; scene boundaries shown below.

** Note 2008 is a composite mosaic of Landsat 7 scenes purchased by WEL & processed by Geoimage P/L; metadata describing dates and mosaicing was supplied by Geoimage.
2006

LS7 2008 Visual mosaic supplied by GeoImage
LS7 2008 Mosaic_full extents