Archaeological Predictive Models for the Jarrah Forests of South Western Australia

Robert Corner, Tom Lee, Curtin University
Fiona Hook, Archae-Aus
Alison White, ALCOA
Aims

• Build inductive and deductive predictive models of likely aboriginal archaeological site locations
• Used to select sample areas for heritage surveys prior to mining to reduce the risk of disturbing potential sites
• Used for Archaeological research to discover sites and preserve their heritage value.
Impact of sites on mining operations

- Discovery of sites during mine development causes problems (e.g., delays - need approval (3-6 months) to disturb sites, redesign of roads or mining boundary)
- If identified early sites can be avoided
- Sites commonly occur near water course so higher risk of site disturbance when roads cross water courses
- Ore bodies are located on the flanks of hills (mid slope) so risk of disturbance is lower
Impact of mining on Archaeological sites

- Disturbance of site which is an offence under the Western Australian *Aboriginal Heritage Act 1972*.
- Sites are evidence of habitation and are therefore used in research by archaeologists.
- Sites are important and significant to local Aboriginal people as they indicate past use of the Darling Range area by their ancestors.
Study area

- Alcoa’s Huntly Mine site 75km SE of Perth.
- Medium to dense Jarrah forest
- Mediterranean climate (cool, wet winter and hot, dry summer).
- Study area is 25km x 25km.
Evidence of Occupation

• Only evidence of sites is stone artefacts as organic material has degraded.

• Radiocarbon dating of selected sites indicates that artefacts were deposited at least 5,000 years ago on the Darling Range, however there are Pleistocene dates (30,000 years ago) for archaeological sites on the Swan Coastal Plain.

• Most sites on the Darling Range are artefacts scatters.
Site definition

- A site is the occurrence of artefacts or other evidence of human occupation of an area.
  - The type of activity at a site is interpreted from artefacts types, lithologies and number of artefacts present.
- Site data was obtained from multiple sources
  - Different site databases recorded different site attributes
- For modelling purposes each artefact scatter is defined as a site and recorded as a point.
Ground Visibility

- Affected by the amount of leaf litter on the ground.
- Poor visibility greatly reduces chance of site discovery.
- Forest that has been recently burnt has improved ground visibility due to minimal leaf litter.
Deductive Models

• Knowledge driven
• Use expert opinion of subjects known activity and habits to determine likely site location.
• Only data required are the environmental variables identified by the expert.
• Can be quickly and cheaply implemented in an unexplored area without extensive sampling.
Deductive models

- Activity of past Aboriginal groups based on theory
- Statistics compiled from archaeologists notes on DIA database
- Activity seasonally dependant
Inductive Models

• Data driven
• Analyse relationships between known existing sites and environmental variables to determine factors that affect site location.
• Requires extensive sampling
• More difficult to produce than deductive models but potentially more accurate.
Cold weather created bias towards areas of northern aspect.

High rainfall made water readily available in all streams.

Rainfall created bias towards well drained soil.

Edible vegetation grew near rock outcrops.
Winter Model Input

• Sites likely to be:
  – Within 500m of water,
  – Within 1000m of granite or dolerite outcrops,
  – Northern aspect,
  – Slope less than 10 degrees,
  – Within 500m of sandy or gravelly soil,
  – Within 500m of well draining soil.
• Low water supply, drinkable water in only fourth or above order streams.
• Warm weather creates a bias towards southern aspects for shade.
Summer model input

- Sites are likely to be:
  - Within 1000m of fourth or higher order streams,
  - Southern aspect,
  - Slope less than 10 degrees,
  - Within 500m of sandy or gravelly soils.
Available Data

- Primary data
  - Surface Hydrology
  - Roads
  - 5m contours
  - Rock outcrops
  - Vegetation
  - Soils (low resolution)
  - Sites on DIA database
  - Sites found by Alcoa

- Derived data:
  - DEM
  - Slope Aspect
  - Hydrology with dam removed and streams interpolated and Strahler stream ordering carried out.
Building deductive models

• Weighted overlay model.
• Binary (likely/unlikely) data layers produced for each attribute
• Arithmetic Overlay to produce final result.
Model schemas

Summer Model

- 4th+5th order streams
- Buffer to Within 1000m of water
- Aspect
- Buffer to Southern aspect
- Slope
- Reclass to Slope <10 degrees

Winter Model

- Streams
- Buffer to Within 500m of water
- Aspect
- Reclass to Northern aspect
- Slope
- Reclass to Slope <10 degrees
- Outcrop
- Buffer to Within 1000m of outcrops

Arithmetic overlay

Summer Model

Winter Model
Inductive model

- 67 known sites, 56 used in building model, 11 used in testing.
- Variables available for analysis:
  - Slope
  - Aspect
  - Distance to Water
  - Distance to outcrop
Slope

- Mean slope (study area): 4.31 degrees
- Mean slope (known sites): 3.61 degrees
- 95% of sites on slopes less than 7.5 degrees
- Less than 7.5 degrees adopted as threshold
- Aspect may be seasonally dependant.
- No clear bias towards any aspect.
- Aspect not used in the model.
• Mean distance to rock outcrop (study area): 5344 m.
• Mean distance to rock outcrop (known sites): 6827.37 m.
• 8 sites within 1900 m of outcrop,
• 2000 m adopted as threshold
• Mean distance to water of sites: 150m
• 82% of sites are within 210m of streams.
• 210m adopted as threshold
Distance to $4^{th}/5^{th}$ order streams.

- Summer water only available in $4^{th}$ of higher order streams.

- Sites where
  \[ D_{\text{allwater}} \equiv D_{\text{high order}} \]
  are within 370m of water.

- 500m adopted as threshold for fourth or higher order streams.
Inductive model schema

Digital Elevation Model

Derive Slope

Slope

Distance to water (1st – 5th order)

Reclass Slope <7.5 degrees

Streams

Distance Calculation

Distance to water (4th – 5th order)

Reclass Water <210m (1st – 5th order)

4th & 5th Order Streams

Distance Calculation

Distance to water (4th – 5th order)

Reclass Water <500m (4th – 5th order)

Rock Outcrops

Distance Calculation

Distance to Rock Outcrop

Reclass Within 2000m of Rock Outcrop

Overlay

Final Model
Field Verification

- Two days, four people
- Four transects 50m wide and 4680m long combined, covering 26.72ha.
- Sample area dependant on ground visibility.
- Medium to low ground visibility in area of burnoff in 2001.
Sampling and testing

- Transects covered variation of environmental variables and site probability.
- Three new sites were found.
- Deductive models were tested using all known sites (67 sites).
- Inductive model tested using sites not used in model building (11 sites).
• 7.5% of sites in high probability class which takes up only 3.2% of the area resulting in a high density of sites.
• Many sites in low probability classes
### Verification of Winter Model

<table>
<thead>
<tr>
<th>Value</th>
<th>Class</th>
<th>Total Sites</th>
<th>Total Sites (%)</th>
<th>Background</th>
<th>Sites per km²</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>V. Low</td>
<td>0</td>
<td>0.00%</td>
<td>0.21%</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>Low</td>
<td>3</td>
<td>4.50%</td>
<td>14.22%</td>
<td>0.032</td>
</tr>
<tr>
<td>2</td>
<td>Medium</td>
<td>40</td>
<td>59.70%</td>
<td>63.87%</td>
<td>0.096</td>
</tr>
<tr>
<td>3</td>
<td>High</td>
<td>23</td>
<td>34.30%</td>
<td>20.43%</td>
<td>0.173</td>
</tr>
<tr>
<td>4</td>
<td>V. High</td>
<td>1</td>
<td>1.50%</td>
<td>1.27%</td>
<td>0.121</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>67</td>
<td>100.00%</td>
<td>100.00%</td>
<td></td>
</tr>
</tbody>
</table>

- Fairly even distribution of densities across classes.
- High class 34% of sites in 20%
- Very high class does not cover much area.
Inductive model testing

<table>
<thead>
<tr>
<th>Value</th>
<th>Class</th>
<th>Sites</th>
<th>Total Sites (%)</th>
<th>Background</th>
<th>Sites per km(^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>V. Low</td>
<td>1</td>
<td>9.09%</td>
<td>0.39%</td>
<td>0.395</td>
</tr>
<tr>
<td>1</td>
<td>Low</td>
<td>2</td>
<td>18.18%</td>
<td>49.95%</td>
<td>0.006</td>
</tr>
<tr>
<td>2</td>
<td>Medium</td>
<td>4</td>
<td>36.36%</td>
<td>38.88%</td>
<td>0.016</td>
</tr>
<tr>
<td>3</td>
<td>High</td>
<td>2</td>
<td>18.18%</td>
<td>9.48%</td>
<td>0.032</td>
</tr>
<tr>
<td>4</td>
<td>V. High</td>
<td>2</td>
<td>18.18%</td>
<td>1.30%</td>
<td>0.236</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>11</td>
<td>100.00%</td>
<td>100.00%</td>
<td></td>
</tr>
</tbody>
</table>

- Very high class has high density of sites 18% in 1.3% of the area
- 36% of the sites are predicted in 10% of the model area (high and very high classes)
Comparison of models

- The summer and the inductive models have the greatest area covered by the low probability classes, which is representative of current understanding of the area.
- All models have high densities in the high probability classes.
- Inductive model has the best distribution of densities across class boundaries, and distribution of class boundaries across the study area.
Future work

• Use conditional probabilistic networks
  – Data driven, expert moderated approach

• Separate sites
  – Stone working sites
  – Others

• Investigate landscape evolution
Acknowledgements

• Tom and other co-authors for doing the work!
• Christine and Bruce for helping with field verification
• ALCOA for provision of data
• PNC 2003 for letting us present the work