New developments in archaeological predictive modelling

ISA Summer School Tours

29 June 2007

Philip Verhagen
ACVU-HBS, Amsterdam

Martijn van Leusen
Groningen University

Benjamin Ducke
Christian Albrechts University, Kiel
Outline of lecture

- background: theory and history
- the basics of predictive modelling
- successes and failures
- new developments
What is archaeological predictive modelling?

a technique that, at a minimum, tries to predict:

“the location of archaeological sites or materials in a region, based either on a sample of that region or on fundamental notions concerning human behaviour”

(Kohler and Parker, 1986:400)
The fundamental problem of predictive modelling

known sites = roughly 1-10% of population

where are the other 90-99%?
What can predictive modelling do for us?

Archaeological Heritage Management (AHM)
- avoid destruction of archaeological remains
- help developers with planning
- improved resource allocation, risk reduction

research
- exploring settlement patterns and processes
- test hypotheses (models) against predictions
The beginnings (1975-1985)

- settlement studies
  - from site-based to regional studies
  - ecological approach

- Cultural Resource Management
  - National Historic Preservation Act (1966)

- New Archaeology
  - ‘scientific’ approach, applying quantitative analysis to archaeological data
  - the ‘inductive’ method
The golden years (1985-1995)

- GIS
  - quantitative spatial analysis made easy
  - pretty maps

- acceptance in Cultural Resources Management in the USA
  - selection tool

- export good
  - first Dutch predictive maps produced in 1990
Where is it done?

**Minnesota Predictive Modeling Project (Mn/Model)**
7,000 sites, 230,000 km² (France 338,000 km²)
financed by Department of Transportation

www.mnmodel.dot.state.mn.us/index.html

**North Carolina Predictive Model**
37,000 sites, 140,000 km²
derpartment of Transportation

www.informatics.org/ncdot
Where is it done?

Archaeological Predictive Modelling in Ontario’s Forests
Where is it done?

The Netherlands

Indicative Map of Archaeological Values – IKAW

66,000 sites, 41,000 km²
financed by Ministry of Culture

www.kich.nl
Where is it done?

Germany

Archäoprognose Brandenburg

8,000 sites, 30,000 km²
financed by Landesamt Brandenburg

www.uni-kiel.de/ufg/projekte_ug/Archaeoprognose/deutsch/sec_willkommen.html
Where is it done?

various other countries:

- Slovenia (Pomurje, highway project, Krištof Oštir et al.)
- Croatia (Island of Brač, academic study, Zoran Stančič et al.)
- Denmark (Eastern Jutland, academic study, Bo Ejstrup)
- Czech Republic (various regions)
- France (Argonne, Rhône Valley, Philip Verhagen et al.; Arroux Valley, Scott Madry; Roussillon, Jean-Michel Carrozza et al.)
- and probably many others …
‘inductive’ modelling

- 'inductive' modelling
- correlation with environment
- known sites
- available data
- model testing
- new data
- model development
- expert judgement
- withheld data
- model prediction
- prediction model
- model building
- correlation with environment
‘deductive’ modelling

known sites

MODEL TESTING

PREDICTIVE MODEL

MODEL BUILDING

hypotheses about site location
Modelling techniques (1)

- **expert judgement (‘intuitive’)**
  - ‘single variable’
  - classification into high/medium/low
  - no quantification

- **advantages:**
  - easy to produce and understand
  - ‘deductive’

- **drawbacks**
  - subjective
Example: Ede

soil and geomorphology

‘landscape units’
categorized into low/medium/high potential

source: Heunks, 2001
Modelling techniques (2)

- **density transfer**
  - ‘single variable’
  - classification into high/medium/low
  - based on relative site density (%sites / % area)

- **advantages:**
  - easy to produce and understand
  - simple classification rules

- **drawbacks**
  - no theoretical backup
  - sampling issues
Example: Castricum

soil map 1:20,000
relative site density calculated

source: Soonius et al., 2005
Modelling techniques (3)

- **weighted overlay**
  - ‘multi-variable’ (multi-criteria analysis)
  - based on expert opinion
  - individual factors are weighted
  - weighted factors are added to arrive at final classification

- **advantages:**
  - easy to produce and understand
  - ‘deductive’
  - simple classification rules

- **drawbacks:**
  - subjective weighting
  - danger of ‘overfitting’ (too many parameters)
### Example: Ontario

<table>
<thead>
<tr>
<th>CATEGORY (W)</th>
<th>SUBCATEGORY</th>
<th>VARIABLE</th>
<th>VALUE (V)</th>
<th>WEIGHTED VALUE (W x V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>proximity to water (W=3)</td>
<td>Order 4-5 Water</td>
<td>0-100m</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Order 3 Water</td>
<td>101-250m</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Order 1-2 Water</td>
<td>251m+</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>slope (W=2)</td>
<td>Slope</td>
<td>0-5°</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6+°</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>drainage (W=3)</td>
<td>Drainage</td>
<td>Dry</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mixed</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wet</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

source: Dalla Bona, 1994
Figure 1.2. Flowchart illustrating the development of a winter land use model

source: Dalla Bona, 1994
Modelling techniques (4)

- **logistic regression**
  - robust statistical technique, multi-variable
  - seeks the best model by step-wise regression
  - produces site and non-site model
  - final classification through intersection of site and non-site model
- **advantages:**
  - statistical, ‘objective’ method
  - weights of variables calculated instead of estimated
  - ‘overfitting’ can be analyzed and reduced
- **drawbacks:**
  - no theoretical backup
  - sampling issues
Example: Minnesota

source: Hudak et al., 2002

- post-processual archaeology
  - environmental determinism
  - the problem with induction

- data problems
  - David Wheatley (2003): ‘archaeological reality is too complex to be modelled’

- quality control
  - how certain are we?
  - how do we deal with new data?
Reassessment (2000-2005)

- debate

- flaws
  - bad data produce bad models
  - limited theoretical perspective
  - lack of field testing
  - no quality norms

- opportunities
  - uncertainty mapping
  - loads of new survey data
Archaeological reality in the vicinity of my office?
New developments (1): Bayesian inference

\[ P(A|B) = \frac{P(B|A)P(A)}{P(B)} \]

- integrates expert judgement and empirical data in a quantitative framework
- model-based statistics, multi-variate
- uncertainty measures (confidence limits)
- ‘inductive learning’
- proved successful in radio-carbon dating, but not (yet) in many other archaeological fields
- problem: how do you quantify expert judgement?
Example: Rijssen-Wierden (1)

- experts asked for quantification (‘prior’)
- archaeological data added (‘conditional’)
- prediction (‘posterior’) + uncertainty mapping
New developments (2): Dempster-Shafer models

\[ \text{bel}(A) \leq P(A) \leq \text{pl}(A) \]

- needs two, mutually exclusive, hypotheses (site/non-site)
- belief = evidence in favour of hypothesis
- plausibility = maximum possible belief
- the rest is indeterminate (uncertainty hypothesis, ‘ignorance’)
- evidence from multiple sources combined through *Dempster’s rule of combination*
- only works if evidence from multiple sources is not in conflict
Example: Rijssen-Wierden (2)

- 3 maps:
  - site prediction
  - non-site prediction
  - uncertainty
Dempster-Shafer models: problems

- how do you decide whether the evidence ‘supports’ a hypothesis?
- sampling issues
- role of expert judgement
Looking at model quality

- how do we decide what modelling procedure produces the best results?

- we need criteria to judge the model’s performance
  - explanatory framework
  - transparency
  - best possible prediction with given dataset
  - good prediction in future
  - assess uncertainty in prediction
Model performance issues

- **accuracy**: how many sites in the model?
- **precision**: how small is the zone of high probability?
model performance measures

- popular model performance measures:
  - Kvamme’s gain \( 1 - \frac{p_a}{p_s} \)
  - relative gain \( p_s - p_a \)
  - indicative value \( \frac{p_s}{p_a} \)

- a model that captures 60% of the sites in 30% of the area has
  - Kvamme’s gain \( 1 - \frac{0.3}{0.6} = 0.5 \)
  - relative gain \( 0.6 - 0.3 = 0.3 \)
  - indicative value \( \frac{0.6}{0.3} = 2.0 \)
which model performs best?

<table>
<thead>
<tr>
<th>Model</th>
<th>i.v.</th>
<th>i.v. ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model A</td>
<td>0.6</td>
<td>0.3</td>
</tr>
<tr>
<td>Model B</td>
<td>0.8</td>
<td>0.4</td>
</tr>
<tr>
<td>Model C</td>
<td>0.9</td>
<td>0.6</td>
</tr>
</tbody>
</table>
How do Dutch predictive models perform?

Performance statistics for three different models:

- IKAW, Eastern Sandy Area
- Limes Gelderland
- Municipality of Ede
comparison of modelling procedures

- weighted overlay (0.41)
- logistic regression (0.29)
- Dempster-Shafer (0.47)

source: Ejstrud, 2003
model optimisation
intersection method

The diagram illustrates the relationship between the percentage of sites (% sites) and the percentage of area (% area). It shows two curves:

- The red curve represents the site model, which indicates an increasing accuracy as the % area increases.
- The green curve represents the 'non site' model, which shows an increasing precision as the % area increases.

The intersection point of the two curves suggests an optimal balance between accuracy and precision, depending on the specific needs of the analysis.
resampling

- re-uses sample data
- criticized in the past, but in fact good practice
- useful for error estimation and statistical inference
sampling, sampling, sampling

- unbiased samples of sufficient size needed

- potential sources of bias
  - surface visibility
  - artifact density
  - site size
  - preferential sampling

- can only be analyzed and corrected when we have sufficient information about data collection
## Data Sources in the Netherlands

<table>
<thead>
<tr>
<th>Method</th>
<th>Sampling Depth</th>
<th>Coverage</th>
<th>Preference for High Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field Walking</td>
<td>Ploughzone</td>
<td>Vegetation Dependent</td>
<td>Moderate</td>
</tr>
<tr>
<td>Core Sampling</td>
<td>&gt; 7 m</td>
<td>Small</td>
<td>Moderate</td>
</tr>
<tr>
<td>Trial Trenching</td>
<td>&lt; 2 m</td>
<td>Partial</td>
<td>Strong</td>
</tr>
<tr>
<td>Excavation</td>
<td>&lt; 2 m</td>
<td>Full</td>
<td>Very Strong</td>
</tr>
<tr>
<td>Watching Brief</td>
<td>&lt; 2 m</td>
<td>Full</td>
<td>Weak</td>
</tr>
</tbody>
</table>
erosion/accumulation modelling
discovery probability model

- built-up area
- heather
- forest
- surveyed zone
historical land use maps
conclusions

- predictive modelling is there to stay

- but model quality is insufficiently addressed

- we need
  - methods to incorporate uncertainty
  - source criticism
  - field testing
Rijssen-Wierden: area introduction
sandy plains
morainic hills
river valley
Rijssen
Enter
peat marsh
Wierden
sandy plains
morainic hills
reclaimed peat
plaggen soil

Regge

tumulus

Regge valley

De Borkeld