Mapping Ambiguity:
Using stray finds data to determine archaeological potential in North Yorkshire

Dissertation for
MSc in Archaeological Information Systems
York University

2005

Nicholas Boldrini

Table of Contents

List of Illustrations

List of Abbreviations

ACAO	Association of County Archaeological Officers
ADI	Artefact Density Index
ALGAO	Association of Local Government Archaeological Officers
AOT	Archaeological Objects Thesaurus
CAA	Computer Applications in Archaeology Conferences
CBA	Council for British Archaeology
DC	Development Control
DCADI	Development Control Artefact Density Index
DCMS	Department of Culture, Media and Sport
EH	English Heritage
EMA	Events-Monuments-Archives Model
FISHEN	Forum on Information Standards in Heritage (England)
FISH	Forum on Information Standards in Heritage
GIS	Geographic Information System
HBSMR	Historic Buildings Sites and Monuments Record
HER	Historic Environment Record
MDA	Museums Documentation Association
MIDAS	Monuments Inventory Data Standard
NGR	National Grid Reference
NMR	National Monuments Record
NYCC	North Yorkshire County Council
OS	Ordnance Survey
PAS	Portable Antiquities Scheme
RCHME	Royal Commission on Historical Monuments for England
SMR	Sites and Monuments Record
TMT	Thesaurus of Monument Types

Acknowledgements

Neil Campling, Principal Archaeologist, North Yorkshire County Council - for his support throughout the course, and also for discussions about this issue which lead to me considering the topic for the dissertation

Julian Richards, York University – for help and advice throughout the MSc in Archaeological Information Systems, and in particular during this project

William Kilbride, Archaeology Data Service – for comments on parts of this dissertation

Steve Dobson, York University – for help and advice on the use of ArcView and other software issues

Ben Gourley, York University – for help with working out how to export raster to vector files

Barry Taylor SMR manager Greater London Archaeology Service; Louisa Matthews, Assistant Archaeologist (SMR), South Yorkshire Archaeology Service; Alice Cattermole, Assistant Records Officer Norfolk Historic Environment Record; Sarah Poppy Historic Environment Record Officer Cambridgeshire County Council;– for answering my queries about finds recording in their respective Sites and Monuments Records/ Historic Environment Records;

Victoria Bryant, Historic Environment Record Manager within the Worcestershire Historic Environment and Archaeology Service - for answering my queries about finds recording in the Worcestershire Historic Environment Record and also for background advice on Finds

Chris Cumberpatch – Freelance Pottery Specialist, for advice and information on finds

Daniel Pett – ICT Adviser at the Portable Antiquities Scheme - for providing me with the Portable Antiquities Scheme digital data and also answering my questions about the scheme and data

Simon Holmes – Finds Liaison Officer for Yorkshire - for answering my queries about the Portable Antiquities Scheme data from North Yorkshire more specifically

Chapter 1

Introduction.

The aim of this dissertation is to develop a methodology which allows stray finds data, which is often only imprecisely recorded within Historic Environment Records (HER), to be incorporated into Geographical Information System (GIS) mapping in such a way that it can usefully influence development control (DC) decisions.

In the following discussions it is important to understand the distinction between the recording of finds and the recording of find spots. A find could be recorded as part of an excavation record and linked to the generic site record for the excavation. However, findspots refers to finds which were found 'off site', either during structured field work or by chance. It is possible to record similar data about the artefacts recovered in both cases, but in the latter case, it is not possible to posit a 'site' from which the finds came – either because not enough data is known, or because, in fact, no such 'site' exists. In this project I am mainly concerned with these 'off site' or stray finds.

For various reasons, some finds are often only recorded to an area location – such as a Grid Square, Parish, or Quarter sheet location. This can be due to the location coming from antiquarian sources which it is impossible to re-interrogate to ascertain more detail, or to unwillingness to give more precise detail by more modern finders, such as some metal detectorists.

However, to fully understand the archaeology of an area, even 'fuzzy' data such as this need to be borne in mind, otherwise one may miss evidence of activity, which may not be visible from other complimentary sources, such as earthwork or aerial photographic evidence. An example of this is given by Phillips (1980:19 – 20) where he explains that the mapping of antiquities on Ordnance Survey (OS) maps (which usually relate to the site of discoveries of hoards or other finds) have often lead to the discovery of more extensive sites by later researchers.

The problem of mapping imprecisely located artefacts has been with archaeology for some time, though a satisfactory solution has never been found. In map and paper/card index based analogue systems the problem was not really solved. What actually transpired was that a number of pragmatic solutions were worked out.

These solutions can be classed as map based and non-map based. On maps, the pragmatic solution was Marginalia notes on Maps which informed an examiner that some Roman finds, or other finds of archaeological interest, were found somewhere on a particular map. The non-map based approach involved putting a record of the find in the relevant Parish of the Parish based filing system so that a similar note would be made by those researchers using parish based research methods. (OS 1978)

In the digital age, the constraints of data of having to be, to a certain extent, yes/no, off/on or 1/0 have re-kindled the problems, or, more precisely, re-cast them in a different medium. This problem has been hidden by the fact that often the computerised compilation of most Sites and Monuments record (SMR) was done with the focus on trying to achieve data entry as quickly as possible given the mass of data accumulated over the years of the SMR.

However, now the dust of the SMR digitisation stampede has begun to settle somewhat, more thought is being given to what data is actually in the SMR. This has also been prompted by the switch in emphasis from SMR to HER, as detailed in, for example, *Historic Environment Records: Benchmarks for Good Practice*. (EH ALGAO 2002).[1] Another influencing factor is the emergence within archaeology of the role of full time data managers (e.g. Baker 1999 records at least 21 officers in this role), as opposed to officers whose duties were split between DC work and data management, and who were quite often under too much pressure from their DC related activities to be able to worry about such issues. These various factors have helped engender a period of reflexivity regarding HER's generally, but also regarding finds recording within HER's (Fernie & Gilman 2000 Page C11).

Further prompting has come from the establishment and expansion of the Portable Antiquities Scheme (PAS). This has created a new system of recording finds, which often have imprecise location data, and has re-kindled the debate to some extent (Fernie & Gilman 2000 Page C11; HER Forum Email List Archives April 2001; Gilman 2004).

The final prompt has been the rapid increase in the use of GIS by HER's. In 1997 there were approximately 20 using some sort of GIS (Foard 1997) and this had risen to 22 (29% of HER's) by 1999 (Baker 1999: 18). The most recent survey showed 88% of HER's using GIS (Newman 2002: 16). The way GIS has mapped data has highlighted the problems mentioned above graphically. With data of, for example, four figure National Grid references (NGR), there have commonly emerged small clusters in the bottom left hand corner of a Grid Square, as a number of finds are lumped together at a generic point. The problem of quarter sheet and Parish level finds is even more complex. Often an arbitrary point within the Parish or Quarter sheet may have been chosen to represent these finds.

However, if an area just off screen from these points is interrogated, the data they represent are essentially ignored in decision making processes, giving distorted views of the archaeological potential of an area.

---

[1] In the following discussion the terms will be used as chronologically correct i.e. prior to 1998 the term SMR will be used, post 1998 the term HER will be used. Where the context crosses this date, the generic term SMR/HER will be used.

What I have experimented with in this project are alternative methods for mapping such finds so as to produce what I have termed an Artefact Density Index (ADI) for areas which will, hopefully, more usefully reflect the pattern of activities across the landscape. This Index has been divided into broad periods, and takes the form of a “Red flag” type model (Altschul 1990).

In attempting to look at possible solutions to the above issues, I have developed a methodology, and then tested this methodology on a subset of the North Yorkshire County Council (NYCC) HER also using additional data from the Portable Antiquities Scheme (PAS) for the same area.

However, the aim of this project is *not* to create a tool which can be used i*n isolation* to make DC decisions. The aim is specifically to deal with scattered and imprecisely located finds and to try and integrate these into the DC process. With this in mind, it is envisaged that the DCADI will be an additional layer of information to further inform DC decisions.

In Chapter 2, the context of the project is described, by looking at the background to the development of SMR/HER’s and their recording of Finds, looking at North Yorkshire in detail, and also looking at the PAS scheme.

Chapter 3 looks at the technical issues relating to developing the final product, such as how to map ambiguously located finds and how to categorise the finds, before setting out the chosen methods.

Chapter 4 details the actual process of data preparation and building the software model.

Chapter 5 compares the different Development Control solutions examined, comparing their strengths and weaknesses.

Chapter 6 is a round up of the overall project and details the main conclusions.

Chapter 2

Retrieving the past: Archaeology and Artefacts; Databases and GIS

In this chapter the aim is to give a brief history of the development of HER's, with particular emphasis on the recording and mapping of finds, to put the project into a wider context. These issues will then be examined with reference to the two datasets chosen for study – the NYCC HER and the PAS data.

Archaeological Database Development

A database can be defined as "A collection of data arranged for ease and speed of search and retrieval" (Dictionary.Com: 2004). Whilst in recent years, the term has increasingly been used in relation to particular types of software which perform the above task, it is worth remembering that manual systems have long been developed as forms of databases. These can range from files in a filing cabinet, to card indices for things from library catalogues to archaeological records. The fact that many paper based systems have been digitised over time has muddied the waters further, but the development of SMR type databases must be traced back to their paper roots to understand their present state. In this section I will focus only on a few key issues relevant to this project, for a fuller SMR development history, see e.g. Fernie and Gilman 2000; Robinson 2000; Gilman 2004.

The most important Monument Database for understanding the development of SMR's is that created by the OS Archaeology Division. The full history of the OS Archaeology Division is recounted elsewhere (e.g. Phillips 1980). However, in terms of the development of Monument databases the salient points are that in the years prior to World War II, the details of site antiquities were gathered by a network of local field workers, with the sources of the information being recorded in the appropriate OS Name Books (Phillips 1980). As much of this archaeological data was never published in any form, the loss of most of these books, from the bombing of the Southampton offices of the OS in the Second World War, meant much irreplaceable data was lost. Presumably it was as a result of these losses, that in 1947 CW Phillips instituted the OS Card Index system, whereby each site or find was given a unique identifier number which identified it both on the map on which it was marked, as well as in the Card Index system which recorded the sources of information for the site, as well as other pertinent information or comments (Murray 1999: 236; Phillips 1980: 54).

The OS system needs to be examined in some detail by any researcher trying to make sense of the current state of SMR's, as most SMR's used it as the basis for their development (RCHME 1993; Murray 1999: 236). In the OS system, the aim was always to try and record locational information about finds or sites as precisely as possible. However, it was also recognised, that in some cases it may not be possible to record data with sufficient precision. As a result the OS system of “marginalia” developed (ACAO 1978). In this system, if a find could not be located to within a 200 metres area, it was added to the marginalia. This meant that a card and index number were allocated as normal, but that the record was not mapped. Instead a note was made in the Margin of the appropriate map, or maps, to show that the find had been made in the vicinity of this village, or within the parish. On the Card the NGR was often given as a four figure reference. The system also involved, for some time, a colour scheme on the map sheets to help distinguish different types of record and their precision of mapping. For example, a yellow band was shown to indicate the area within which an item or site was found. However, this colouring system was abandoned circa 1978, and it is unclear how long it was in use for (OS 1978). On the cards themselves, as much detail about the find as was possible was recorded, including data about the type of finds (e.g. coin hoard, or pottery sherds etc); the material of the finds; the period of the finds, and other details where known. There is apparently no guidance as to what should have been recorded regarding finds, but the general aim of the OS to record as much data as available obviously influenced what fieldworkers noted.

Digital SMR Databases & GIS

Archaeological Computing is a huge subject in itself, best understood by reviewing the published Computer Applications in Archaeology conference proceedings (e.g. Laflin 1986; Lockyear & Rahtz 1991; Kamermans & Fennema 1996). However, a subset of this general development is the setting up of computerised SMR’s. The OS System was a key influence on this process, and formed the basis for the computerisation of the English National Monuments Record (Leech 1986). The design of the cards themselves often influenced the design of the SMR software databases which were used, and it has been argued that much computerisation in SMR’s has “centred on the computer replication of the original manual card-index systems” (Harris & Lock 1992a: 187; Gilman 2004). This is an important point with regard to the recording of finds within computerised SMR’s. The main guidance on establishing SMR’s suggested consideration of finds recording, but also suggested that this would be of less a priority than the recording of sites, and that finds from sites would be relegated to the archive sections. (ACAO 1978: 5). Whilst written at a time when most SMR systems were still manual, it can be seen that this approach will have influenced the software design, as noted above.

In the early 1990’s, as SMR’s became increasingly computerised, the advent of desktop GIS meant that there were increasing calls for the adoption of GIS by SMR’s in particular, and archaeologists generally. At this stage in SMR computerisation, it is worth remembering that often the computer system was usually simply a database, cross referenced to paper maps (Lock & Harris 1991; Lock & Harris 1992a; Lock & Harris 1992b), and that even now most HER’s are hybrid paper and digital systems (Robinson 2000: 90).

Even at this early stage in the use of GIS in archaeology the problems of the precision of the mapping of sites was recognised. However, there were suggestions that this may not necessarily be an insurmountable problem for SMR's (e.g. Harris, T M; Lock, G R 1992a). In terms of mapping in SMR's, the view of Wheatley and Gillings is likely to be the true situation: "In many records, the spatial referent in the conventional database does not mean 'this site is at this co-ordinate with a given precision' but instead means something like 'the site is in the grid square whose lower left hand corner is described by this co-ordinate'. " (2002: 222). In the context of this project, the term 'Site' could also be taken to apply to findspots.

Whilst the above issues apply to all sites and finds records, it is most pertinent with regard to the Marginalia sites mentioned before. The turning of these into GIS points has highlighted the problems for how these are mapped once again. In some cases, the creation of the record from the OS card means the generic OS NGR has been used to map the find, though sometimes the metadata about the precision of the find has been lost.

Finds

As should be beginning to be obvious, the recording of finds within SMR/HER's is very much the poor relation of other aspects of the Historic Environment. This stems back to the fairly early days of the development of SMR's (ACAO 1978: 5), as noted above, and this view can be seen to have been agreed and fossilised by later standards documents (see below) (RCHME 2003).

However, recent years have seen this issue re-kindled for a number of reasons. One is the maturing of the SMR/HER community to a point where there is now increasingly an officer whose task is to manage the SMR/HER as a primary task (Baker 1999:12), rather than an officer where this is just one of many tasks where those with strict time constraints (i.e. Development Control) inevitably took priority (Addison 1999). The separation of the roles of Development Control and Countryside Management from SMR/HER management and enhancement has given time, space and resources to consider aspects of recording which were not possible before. This reflexivity has also been given impetus by the campaign to give SMR/HER's statutory basis (e.g. Department of National Heritage 1996a: 46; Darvill & Fulton 1998: 235; EH 2000: 39) and the associated move to define acceptable standards (EH ALGAO 2002).

Portable Antiquities Scheme

Another development has been the setting up of the Portable Antiquities Scheme, which is purely focussed on the recording of finds and their locations (Pett et al 2003), and has involved liaison with SMR/HER's. This new project has caused SMR/HER's to re-consider how they record finds data (Addison 1999; Fernie & Gilman 2000: C 11).

The PAS was set up in October 1997 by the Department of Culture, Media and Sport (DCMS), in five pilot areas, with a further pilot area funded by the British Museum. In 2003, with Heritage Lottery Funding support, the scheme was rolled out nationwide (Chitty & Edwards 2004). The scheme was the outcome of a long process which started with growing concern over the loss of archaeological data due to the growth in metal detecting as a hobby (Dobinson & Dennison 1995; Bland 1999; Chitty & Edwards 2004).

The PAS records finds in a database, which is a relational database design. Originally developed as an Access database, the current version has been re-designed to be web enabled as part of the design specification (Pett et al 2003: 7). The PAS database uses some PAS defined Look Up Tables of approved terms for certain key fields (e.g. Pett et al 2003: 64), and also includes a number of INSCRIPTION approved Thesauri (Pett et al 2003: 44).

Whilst one of the aims of the PAS was simply to record Portable Antiquities, another was"To advance knowledge of the history and archaeology of England and Wales by systematically recording archaeological objects found by the public" (Chitty & Edwards 2004: 10). It was hoped that the information gathered would be fed into HER's to inform the development control process (Bland 1999). However, this process never seems to have developed into the routine system originally envisaged (Chitty & Edwards 2004). Attempts at data exchange have usually been done at an informal level at a local level (e.g. Linda Smith: Pers comm), with more structured attempts to come up with an import routine for the Exegesis software, for example, breaking down due to the process being more complex than envisaged. (Paul Cuming: Pers comm). Whilst moves have been underway to develop a national approach for exchanging the data formally between the PAS and HER's (Sargent 2002), this process has stalled for various reasons.

However, the PAS has made HER officers more aware of the problems of incorporating findspot data into HER's in a way that can help DC decisions. This has been highlighted by the fact that in many cases, data recorded by the PAS is only to a parish level, despite the fact that the importance of findspot locational data is recognised within the PAS (Pett et al 2003: 21). This problem holds particularly true for the Yorkshire area (Simon Holmes: Pers comm) and is generally due to the fact that many reporters to the scheme, usually metal detectorists, are often reluctant to give precise locational details about where the reported artefacts come from.

Standards

A related issue to the computerisation of SMR data is the one of standards for data entry. Standards for recording archaeology have long been a preoccupation of archaeologists, the first example of which dates back to the 1907 Congress of Archaeological Societies (Burrows 1985: 6). Since then the issue has bubbled under with periodic attempts to establish Standards.

As more systems began to be computerised during the 1980's, a data standard review was begun in 1987, with the establishment of a Working

Party on Information Standards (RCHME & ACAO 1993). Similarly, the development of common terminology was recognised as being important with the publication of various Thesauri (e.g. RCHME 1986; RCHME & EH 1995; MDA 1997). This work on standards was continued by FISHEN (Forum on Information Standards in Heritage (England)) from 1998, which became FISH (Forum on Information Standards in Heritage) in 2001 when its role expanded to include Scotland, Wales and Northern Ireland. The main achievement of FISHEN/FISH has been the setting up of INSCRIPTION in 1999 – an agreed list of thesauri and word lists covering many aspects of archaeological data.

The major outcome of this increased focus on standards was the first publishing in 1998 of the Monument Inventory Data Standard (MIDAS) (RCHME 2003: 3rd Reprint). This has become the model for the development of most HER's and was confirmed as such in various documents relating to HER development (Baker 1999; Fernie & Gilman 2000; EH ALGAO 2002).

The Events-Monuments-Archives Model

One key point about these standards drives was that the Events-Monuments-Archives (EMA) model was created, developed and began to be implemented within HER's. The following overview of the model is drawn from a number of sources, to which the reader is referred for fuller details (Bourn 1999; Catney 1999; Fernie & Gilman 2000; RCHME 2003).

The EMA model (which sometimes has the Archives section referred to as Sources) is important because it is based on the primacy of Archaeological Events. An Archaeological Event can be defined in slightly different ways, but is essentially a discrete activity in time and space which gathers primary data, or (depending on the definition) involves re-interpretation of primary data. An event will always create an archive or source of some sort (whether this is an unwritten-up site archive or a fully published Monograph).

It is only from sources that Monuments records can be created, and Monuments are not necessarily 'objective' descriptions of the archaeology discovered, but are interpretations of the sources to create a record. This distinction is subtle, but important, as it implies that monuments are not fixed entities, but can change as new data is added to the interpretation.

The key issue for this project is that the EMA model deals very superficially with finds. Finds are seen as being part of the Archive Section of the model (RCHME 2003: 49 - 50), and this has created problems. For example, it is not uncommon to have differential recording of finds. A chance find of a particular artefact may be recorded and indexed against the appropriate Museums Documentation Association (MDA) Thesaurus entry, but the same type of artefact found within an archaeological excavation may not be indexed at all, and only included under a generic assemblage heading (Fernie & Gilman 2000). However, this omission of guidance on how to index or record finds in standards documents means that, in practice, each SMR/HER has developed its own unique, pragmatic standard, which may differ significantly from one side of a SMR/HER boundary to the other.

It is noteworthy that the PAS database does not appear to have followed the MIDAS standard; there is no mention of this in the supporting manual (Pett et al 2003). This may be due to the poor coverage of finds recording within MIDAS. MIDAS states that its aim is to record monuments rather than individual artefacts (RCHME 2003: 49) though it also notes that Findspots of individual artefacts may sometimes be useful to record (RCHME 2003: 18; 49). Furthermore Object Type information is an optional field within MIDAS, though it does refer to other standards in terms of artefact recording, if it is decided to include them. One reason given for this policy is that the level of detail needed to record finds is much greater (RCHME 2003: 50). The extra detail needed is evident in the PAS database which includes fields to record weight, length, width etc (Pett et al 2003). However, this is because it is assumed that the PAS record may be the only record ever made of an object, hence as much data as possible need to be recorded. In SMR/HER's. the usual level of detail recorded is find type, material, data and location, the assumption apparently being that if more detail is needed the artefact itself or other sources will be interrogated to get the relevant information (RCHME 2003 48 – 50).

Exegesis Software

One result of the drive to digitise SMR's, given impetus by the push for standards, was the proposal by Royal Commission on the Historical Monuments of England (RCHME) in early 1996 to set up a "PC based software package for SMR's" (Foard 1996:1). The original design specification was written by the RCHME in consultation with ALGAO, and went to tender in 1997 based on "SMR'97" (Crispin Flower: Pers comm). SMR'97 (RCHME 1997) was a logical data model for SMR's, which was MIDAS compatible. In 1997, the first Exegesis software, named "SMR" was launched. First installations began in early 1998, and their was a fairly rapid take up of the software (e.g. 15 users were reported in 1998 (Exegesis 1998), with 24 users by January 1999 (Exegesis 1999)), suggesting the proposal was timely and popular.

Since then the software design has seen continued upgrades and enhancements, as well as a creditable level of support and use by SMR/HER's. In 2001, the software was renamed HBSMR (Historic Buildings Sites and Monuments Record) with the release of Version 2, which reflected a widening of emphasis. All design changes subsequent to the initial version were led by the deliberations and suggestions of the User Group, with input from English Heritage (EH) and Exegesis. One key area of improvement, most evident with version 3 of HBSMR, has been the increasing emphasis on enhancing the GIS capabilities of the software.

Within the Exegesis system, the various data types are recorded in separate modules. E.g. there is a data module for Monuments, one for events, one for Sources and Archives. Somewhat anomalously, given that MIDAS suggests Finds are part of the Archives section, there is also a separate module for finds. This probably reflects the fact that the types of data which could be recorded about finds are different from many other types of source, though the level of detail recordeable in pre-defined fields in the Exegesis software is more limited than that in the PAS database. In the Finds Module it is possible to record the Finds type (based on the MDA Thesaurus (MDA 1997)), Period of the find (based on the Archaeological Periods List (RCHME 1998)); material it is made from (using the English Heritage materials thesaurus) as well as text data. Records in the Finds module are not directly mappable, but can be given geographic expression by linking them to a record in the monument module. It is worth noting that linking the Find to a Monument record is not quite the same as defining a site from the Find data, as one Monument Record Type is Findspot, which is aimed to be used for stray findspots. Therefore a Monument record could be something along the lines of a point showing that some Roman pottery was found in a particular area, without necessarily having to interpret that data into the site of a Villa or other Roman site.

A brief history of the development of the North Yorkshire SMR/HER

The above discussion gives a broad brush overview of what I consider to be the main factors affecting the development of SMR– the influence of the OS system, the digitisation of SMR's and the development of standards. The next section will give a brief history of the NYCC HER in light of these main drivers.

In North Yorkshire detailing the history of the SMR/HER is hampered by a lack of documentation, so the presented history has been compiled by looking at the available documentation (e.g. Smith, 1997; Griffiths 1982) and discussion with former employees, whose memories sometimes seem to conflict.

The Early Days

A county archaeological officer, with a technician to assist him, first took up post on 7th October 1975 in the Planning Department of North Yorkshire County Council, with one aim being to compile the SMR.

From an early stage, it was realised that computerisation would help officers to cope with the quantity of data: Investigation of options started in 1980 and led to the development of a bespoke software system; thus the "North Yorkshire Archaeological Record System" was born. It was envisaged that it would take just over two years to transfer the basic record, to be followed by a second stage of detailed records in sample areas. The SMR was to include the Yorkshire Dales and North York Moors National Parks, although it was only compiled for the Moors, which had its own mainframe link.

The development of the North Yorkshire System took place on a mainframe system and was a complex piece of software designed to allow inputting of data in 3 main modules – the Basic SMR Record; an Aerial Photograph Record; and a Bibliographic record. The system was based on a hierarchical numbering system, where sub components of a site were given numbers in an ordered way which linked them together. The system used pro forma record cards which were filled in by the archaeology section members before being sent off for addition to the record by data input staff. The system allowed a high level of detail to be entered, down to individual finds. The system also allowed searches to be run by inputting grid reference co-ordinates to determine if an area contained any records.

The only difference between a site record and a finds record was that a different record type was entered in the appropriate field. Apart from that, all the data fields were used in the same way, allowing finds records to be geographically recorded individually, and recorded in some detail. This has lead to some quite complex records in some areas. For example, the hierarchal nature of the system meant that a find could be created as the top level record, and in this way was usually used to map stray findspots. However, at lower levels a find could be recorded as being linked to particular site, though the find would also be geographically recorded separately. In some cases, the level of detail was such that a find could be mapped to the particular trench that it came from within an excavation – with the trench being hierarchically linked to the event, and archaeological site at higher levels. As each of these records could also be geographically recorded, when GIS became available, the plots for such areas was extremely confused, or potentially had a large number of points all on top of each other for complex monuments such as excavated barrows.

In terms of recording find types, the same method was used as for monuments, in that according to various rules, a number of fields were filled in which collectively described the archaeological site or find. Again this allowed great detail to be added, but also was a source of later confusion, as the same type of site might have been input in slightly different ways by different people. Finally the system was made even more complex by the decision to record each find or archaeological element individually. In modern systems if there are three burials, the item is only entered once in the index (RCHME 2003: 96). However, the North Yorkshire System would have created three records.

The reason why data were recorded in different ways in different places represented a period of experimentation with the system as to how best to utilise it. However, the potential level of detail available was a mixed blessing as it made data entry an extremely complex and slow process, and this was the cause of further changes to the recording of data.

Over time the system further evolved into a number of discrete subsets, the relevant ones for this project being SMR Master and Site Index. The Main SMR (known as SMR Master) was the fully compiled version of the SMR. A second dataset was known as Site Index. This was started in 1984, and was based on a subset of the SMR Master table with only 14 of the database fields being filled in, and after 1988 not even all of these were completed. In later years, SMR enhancement projects often used the Site Index as their basis.

For both the SMR Master and the Site Index data set, the key main data source was the OS cards for the area, supplemented with data from the

Yorkshire Archaeological Society Survey Cards, and other sources where available. It should be obvious, therefore, that the OS system has influenced the data in the NYCC SMR.

The development of the original North Yorkshire system was ahead of its time, in many ways, by trying to be more than just a Development Control tool, and aiming to be used for research. However, the issues it had, particularly relating to the Hierarchal Numbering system, caused significant problems.

All Systems Change: Standards rear their head

The database was downloaded off the mainframe system in 1995 onto a Windows based PC, as part of a general restructuring of the County's IT systems.

Shortly after this, North Yorkshire was involved in the first phase of local government reorganisation which took place in April 1996. Following this review, and since April 1997, the SMR has been designated as covering the area outside the two National Parks and the City of York unitary council boundary.

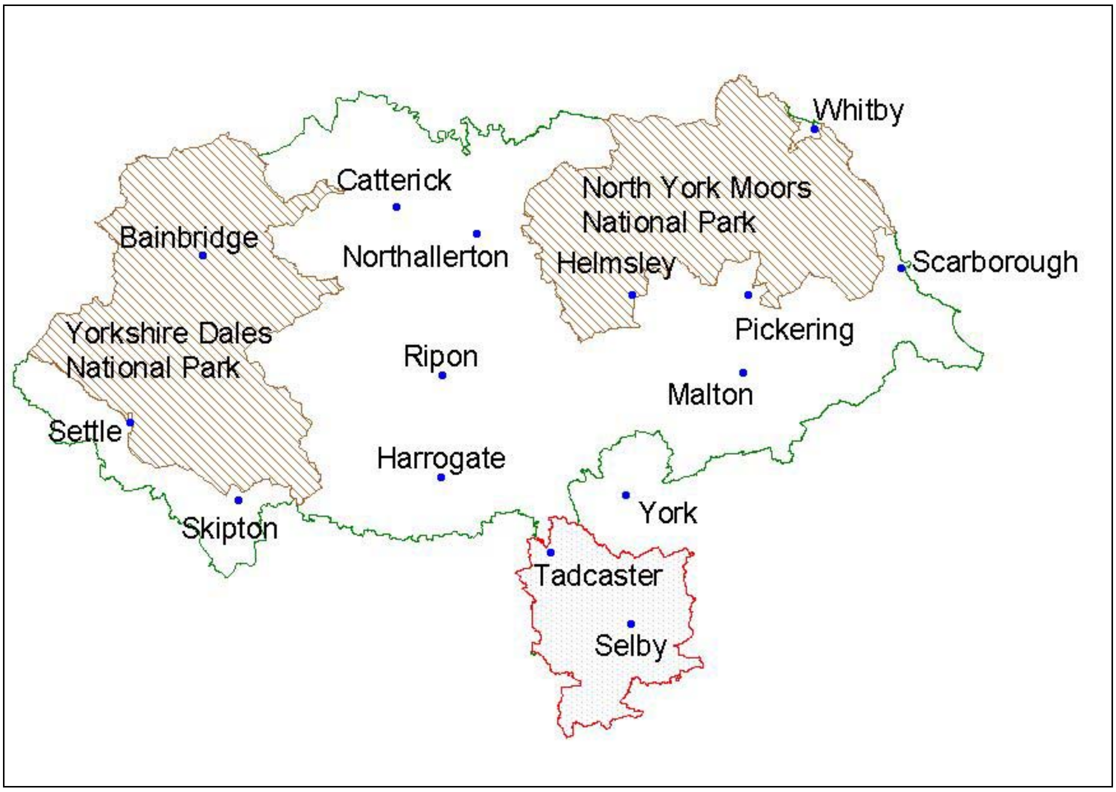


Figure 1. North Yorkshire County Council SMR/HER area from April 1997, with Selby District Highlighted.

For some time, the Archaeology section had been building up a significant backlog of data to input into the system due to pressures of work, and the complexity of the system. The move away from a mainframe system exacerbated this issue as data entry now had to be undertaken by the Archaeology Section staff, and the system was felt to be too unwieldy to allow this to happen easily.

Investigations began to take place within the service about how to take advantage of the possibilities new computerised systems offered, and also to try and deal with the known issues with the system. Between 1996 and 1998 an SMR Data Audit was carried out by NYCC staff. In 1999, Bullens Consultants were commissioned to carry out a review of the HER system and look at options for development, and the eventual outcome of these processes was the decision to adopt the Exegesis Software.

Exegesis saves the day

In April 2002 the NYCC System was migrated to the Exegesis HBSMR software version 2.10. The transfer involved the migration of a number of disparate Monument databases into two main modules of the Exegesis software; the Monuments Module and the Finds module. Similarly, other databases were migrated into the appropriate modules (Events and Consultations). In all cases, records were also created in other modules (e.g. the Source and Archives Module) as appropriate.

The Hierarchal structure of the data meant that the decision about which records to add solely to the Finds Module of the Exegesis system, and to link to Monuments record, and how this should be done was a complicated process. In the end, the Migration was done “as was” with no attempt to rationalise the data at the time, though it was recognised that this was a significant post – migration task.

In January 2004 it was agreed to re-designate the SMR as a Historic Environment Record in line with emerging thinking on the future roles and scope of SMR/HER’s.

THE NYCC experience

The recording of finds within the NYCC system can perhaps serve as a suitable microcosm example of the sorts of issues that SMR/HER’s have with finds generally. The system as originally designed allowed the recording of finds to a great level of detail, including to a particular excavation trench. There are also records showing locations of chance finds, as well as records showing records which are only known to a generic Parish level or rough geographic location. In reality, the recording of finds to a Trench level was rapidly abandoned, as it became obvious that the work of recording to this level of detail would mean compilation of the record would take too long. For example, the initial phase of digitisation begun in 1983 was expected to take just over two years, but had not in fact been completed by 1991 (Smith, 1997).

Furthermore, the fact that the SMR/HER was compiled to different levels of detail over different areas at different times means that there is little consistency in the recording of finds. For example, Site Index finds records are often very skeletal in detail, as opposed to those within the areas more fully compiled.

The backlog problem also means that there is also a significant element of finds data which is in the wider non-digital SMR/HER but has no index within the digital system.

Finally the mapping of finds represents the simple views put forward as being problematic earlier. Most findspots are recorded as point data within the GIS, even those which are only known to a level of precision which would suggest they would be better mapped against e.g. a parish polygon. OS based records which were from 'Marginalia' type records, cluster on the intersection of kilometre square grid lines.

Chapter 3

Developing the model

In this chapter I aim to examine the technical issues which are of specific relevance to the project, and detail possible options that were considered, before detailing the final option chosen. This will be looked at in four main sections: Reasons for the model; Area of Study; Finds Mapping; Finds Data.

Reasons for the model

The aim of the model is to create a measure of the archaeological potential of an area based solely on the relevant stray finds, with particular emphasis on incorporating imprecisely located finds data.

Personal experience of DC work, and discussion with colleagues, suggest that workloads are such that speedy decisions have to be taken about the archaeological potential of a development area, based on data in the GIS, and that if potential is thought to exist then more in-depth assessment of the area will be undertaken before a decision is reached (Gail Falkingham: Pers comm). However, as highlighted previously, it is likely that certain types of data are likely to be missed out from this assessment of potential due to the way they have been mapped.

Discussions with a select number of HER colleagues (a questionnaire was circulated – see Appendix 1) has suggested that they are aware that even imprecisely located finds can add to the archaeological potential of an area (Barry Taylor; Vic Bryant; Louisa Matthews: Pers comm).

As the initial assessment of a development's potential is based on visual inspection of the GIS data, it was envisaged that what would be needed to address the issue at hand is a simple, visual clue to the potential of an area as indicated by finds, one which would flag up an area as being worth some more in depth appraisal.

Again, liaison with colleagues has suggested that creation of such a system would be found useful (Barry Taylor; Sarah Poppy: Pers comm), though some colleagues were less definite, and noted that they would not expect such a system to be useable to determine DC decisions on its own (Louisa Matthews; Victoria Bryant: Pers comm). These latter answers reflect, in my view, the fact that the appropriate question could have been better phrased, as it was never the intention to create a system which would answer DC questions on its own, but rather to create a system to allow the types of data which are not routinely included in such decisions to be taken into account, in addition to other data in the GIS.

As such, I believe that there is some support for the creation of this model for use by DC archaeologists. Obviously, whether that support becomes stronger depends on the success of the model.

Area of Study

In determining an area to use as the trial area for the model, a number of criteria were considered. The area chosen needed ideally to have finds which were mapped to a variety of precisions, to OS quarter sheet only; to parish level only, to a kilometre square only and more precisely located finds.

An OS quarter sheet covers a 5 kilometre square, and the original thought was to test the model on a 20 kilometre square, which would enable a total of 20 quarter sheets to be looked at. However, this would also have involved data preparation for up to 60 parishes, and the possibility of preparing over 700 finds spots for the model, from the NYCC data alone. As the PAS data on top of this would have added a significant extra element of work, it was decided to scale down the area to a 10 kilometre area. This would still allow an area to be chosen which included 4 quarter sheets, but the number of parishes was cut down to 22, which was a more manageable number given the time available.

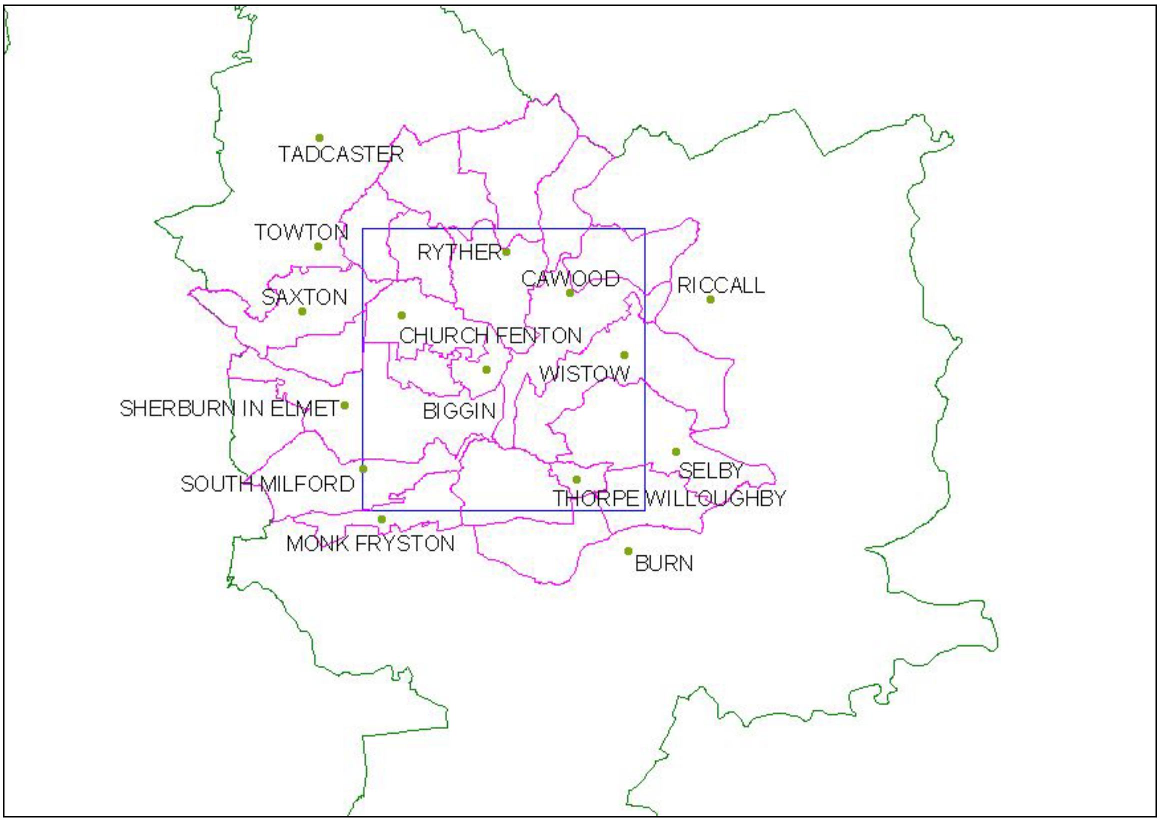


Figure 2. Showing Project Area, with Parish boundaries and local centres.

The decision on how to choose the area was based on some crude analysis of the data as existing i.e. prior to its fuller preparation. A mapping of finds only recorded in the PAS to a parish was made showing only the number of finds. This data was then mapped against point data showing the following: PAS data located more precisely than Parish level; NYCC HER point finds data actually referring to Parish level data and NYCC HER point finds data relating to chance findspots. These layers of data were then examined through a 20 kilometre square template, and this was moved around and visually checked to pick an area where all the varying types of data were present to allow the model to be used against varying precisions of data. The area chosen was in the Selby district and had the Minimum Bounding Rectangle National Grid Co-ordinates of Easting 450000 Northing 430000 to Easting 460000 Northing 440000.

Finds Mapping

The mapping of ambiguously located finds has found varied solutions within HER's. Personal experience of working in three HER's, as well as discussions with colleagues, suggests it is common for most findspots to be mapped as a point within HER GIS systems, (Louisa Matthews; Alice Cattermole: Pers comm), though some are also beginning to experiment with the mapping of finds to polygons (Vic Bryant: Pers comm). In some HER's, finds are only recorded in the GIS if they can be mapped to a kilometre square. Finds located less accurately than this are either recorded in the database and need to be retrieved by a separate query (Alice Cattermole; Sarah Poppy: Pers comm), or are simply put into a paper Parish information file (Louisa Matthews: Pers comm) which will need to be manually searched. In both the cases detailed, for finds mapped only to a kilometre square, a small cluster will appear at the intersection of the relevant grid lines and it is here that the finds mapped only to a kilometre square will be displayed.

In the NYCC system, all finds are mapped where possible. Similarly to the examples above, finds known only to a kilometre square will be mapped to the intersection of the relevant gridlines. However, for finds only known about to a parish or area, these will also be mapped to a kilometre square. As the NYCC System is based on the OS cards, the NGR usually used is that given on the cards, but it is unclear how this was arrived at. For example, for finds only mapped to a parish a NGR is given on the cards – but it is difficult to know how this was decided on as being the most appropriate point in the parish, especially when not all the finds for a particular parish are always mapped to the same kilometre square.

Similarly, for finds only known about to an area, the mapping usually displays the findspot to a kilometre square. In this case, it appears that the appropriate grid intersection is decided upon by choosing the kilometre square which encloses the area detailed. For example, a find recorded as being from near Cawood would be recorded against Grid Reference SE 5636 as this is the kilometre square which contains Cawood town.

There are other options for mapping which could be considered. For example the PastMAP website of the Royal Commission on Ancient and Historic Monuments of Scotland, has experimented with mapping which shows the accuracy of the location of the find (or other record) based on the colour it is displayed at. Finds only located to 10 kilometre precision are mapped at light blue, with darker shades used to represent records mapped to a kilometre, 100 metres and 10 metres, with records of a precision of 1 metre being shown in black (Royal Commission on Ancient and Historic Monuments of Scotland, 2005). Other ideas mooted have been to randomise the distribution of points over the area they could fall within (Julian Richards: Pers comm).

Another possibility is the creation of buffers around points, the size of the buffer representing the area that the find could fall within or even the recording of the finds location as a polygon which covers the area of its possible location (Fernie & Gilman 2000: B 30). This system could work, but might have the disadvantage of having many different overlapping polygons for finds, which would make it difficult to see other layers of data. Known as 'stacking' this is a common issue with GIS (Fernie & Gilman 2000: B 31). Even if these layers were mapped in such a way so as not to obscure other data in the GIS, personal experience suggest that interpreting data presented in this way can be confusing. At NYCC, experimentation of mapping events has revealed that if the various events which can overlap in one development area (e.g. a combination of any or all of the following techniques: desk based assessment, geophysical survey; other field survey; Evaluation trenching, excavation etc) are mapped as polygons then it becomes difficult, with sites with many activities, to disentangle the boundaries of any single event. Our solution has been to map as polygons the first event (usually a DBA) and then merely record the other events as points within this area. In light of this experience, it is felt that the use of numerous polygons has significant disadvantages and is felt not to be the best solution to the issue.

The issue of mapping finds in HER's is not solely connected to imprecisely located finds. As more modern finds recording techniques have been used - for example the use of modern technology such as total stations and Geographic Positioning Systems to 3 dimensionally record finds during landscape surveys, fieldwalking or excavation – it is possible for finds to be plotted to a high degree of accuracy. Incorporating finds recorded to this level of precision within an HER raises a variety of issues, not solely in terms of data management, but also in terms of whether this level of detail is needed, consistently achievable, or even desirable in an HER. A pragmatic solution is that in the case of excavation sites, the individual finds are unlikely to be plotted (though see previous discussion of NYCC system) but the finds information may be indexed and link to the generic site record to allow information on them to be retrieved. Fieldwalking finds are more problematic – it is possible to create a separate record for each artefact in the area surveyed. In the NYCC HER the current pragmatic solution is to link the finds data to one generic record for the area surveyed, and similar methods appear to have been used elsewhere (Alice Cattermole: Pers comm). However, in the context of this project dealing with precise data is less of an issue, as the policy and solutions developed to deal with such data are likely to be developed with the needs of DC HER users in mind.

So the problem of mapping finds in an HER can be seen to occur at both ends of the spectrum – for imprecisely recorded finds, and for very precisely located finds. The key point in regard to the current project is that linking precisely located finds to another record will allow them to be indirectly incorporated into DC decision, as the evidence they give should be reflected in the type of record created. However, for imprecisely recorded finds, it should be obvious that if a search occurs just out of view of the recorded find spot, then even though the search area could still be within the possible area of a finds location, it is unlikely to be taken into account.

In this model it was decided that the appropriate level of recording would be to link the finds data to the appropriate type polygon, depending on the precision of the locational data e.g. link the find to a parish, quarter sheet, kilometre, or lesser sized polygon as appropriate.

Finds Data

The other key aspect of the model to be considered is what information about the finds themselves to incorporate into the model and how. With finds, as all archaeological data, there is an almost infinite amount of data that could be recorded about the artefact. However, in practice there is a common subset of data which is usually recorded within the HER data format and the PAS data formats.

The types of data which are available regarding finds in the Exegesis system are listed in Appendix 2, as are the data available from the PAS system. Of these fields, a number were considered to be of particular relevance to this project: namely – location; period, type of artefact and material. Location has already been discussed (see above) and we will now look at the others in more detail.

Period

The settling of how to deal with Period was fairly easily resolved, as it was decided to use the standard Archaeological Periods List from INSRIPTION (RCHME 1998). Whilst it might be possible to split these terms into more precise categories, it was felt this level of detail was not needed to achieve the aims of the project. Identification of the broad period would be enough to indicate a potential area of interest, and a broad period. One decision made was to use the broad term Prehistoric rather than the separate period terms for Bronze Age, Neolithic etc. This decision was taken due to the small amount of data for the various sub-Periods in the project area, and also to overcome the issue of finds which were recorded as the broad term Prehistoric, as opposed to the more specific terms.

Finds Type

The main criteria for determining the Finds type data centred around the use of the Archaeological Objects Thesaurus (AOT) (MDA 1997). This is because this is the Thesaurus recommended to be used by INSCRIPTION, and also because it is what is actually used in the NYCC HER and PAS data.

Individual Finds Type

An initial idea, quickly abandoned, was the possibility of creating an ADI for each Find type in the AOT However, given that there are over 2000 terms presently within the AOT, this was though to be practically impossible, and it was quickly realised that some method of Grouping the finds data would be needed.

Finds Ranking

One possibility for Grouping was the ranking of finds, so that for example certain types of finds which might be more indicative of certain types of activity in an area would be weighted within the model to reflect this. For example, a poorly located find of a Roman altar might be seen as being more indicative of Roman activity in area, than sherds of Roman pottery, which might be dismissed as background noise. However, preliminary, informal email discussions with some finds specialists (Chris Cumberpatch and Victoria Bryant: Pers comm) suggested that the attempt to rank finds in this way would be at best problematic, but more likely to be somewhat controversial and frowned upon by finds specialists due to the variety of issues which can affect the survival and recovery of finds in the archaeological record. It was therefore decided at an early stage not to pursue this route.

Material Type

Another option considered for Grouping the finds was to try and map the objects based on material of construction. This was thought to have some merit as it is often how finds are dealt with in archaeological reports – with specialist reports on pottery, lithics, glass etc. This method was developed to such a degree that a trial version of the ADI process was developed and created for a Currency ADI (using Metal Coins and Jettons) and plans were underway to try and do similar mappings for lithics, pottery and ceramic building materials. Whilst this was a useful exercise in developing the final model, it did not succeed, ultimately, because major difficulties were encountered in trying to determine how to deal with metal objects.

It was felt that merely mapping all the Copper Alloy finds, or Lead finds would not yield useful information about archaeological potential, and the metals issue revealed the limitations of this Material based approach.

Whilst some materials often only have a limited number of uses and therefore there presence could give some clue about potential archaeological activity on a site, for metals this is not true, as they can perform a wide variety of functions.

For example, lithics usually take a limited range of forms as tools or weapons – it is rare (though not impossible) to find a lithic container. Similarly, pottery usually forms containers of various forms and rarely (if ever) makes weapons. So it is possible to infer a certain amount about possible activities by the material found. Metal has no such limitations, and its qualities mean it can be made into almost anything – tool, weapon, container, etc.

One approach to deal with Metal finds was considered in that it might be possible to map metal objects using the broad classes of the Archaeological Objects Thesaurus. However, this then raised the issue that different materials were being dealt with in different ways.

Monument Type

A second option considered for grouping the finds data was to try mapping the AOT Terms to the Thesaurus of Monument Types (TMT) Thesaurus (RCHME/EH 1995) broad terms e.g. Agriculture and Subsistence; Defence etc. So for example, it was considered it might be possible to say Finds Type X indicated Broad Monument type Y.

At first this method seems promising as it is arguably what archaeologists do all the time; define a monument by the artefacts and their associations with each other.

However, following discussions with Julian Richards, it was realized that this model was starting to become too complex and this process would be problematic. It was felt to be likely that some classes of artefact could be mapped to more than one TMT term. For example, a sword could be viewed as being related to military activity, ritual activity, or industrial activity, at least.

If other finds were present then it might be possible by combination to determine the most appropriate class, e.g. a sword found with slag could be more indicative of industrial activity. However, this would further complicate the model, and would leave questions over how to deal with e.g. a sword on its own. It was therefore decided to discontinue this line of developing the model.

Finds Class Type

A third option was eventually adopted. This evolved from the consideration of the previous options, as well as the realization that the system needed to be relatively simple. The idea of linking the artefact type to functional categories was originally prompted by reading an article by Haselgrove (1985), and had lead to consideration of the Monument Type option detailed above.

In this article, Haselgrove discusses the possibility of trying to determine the nature of any possible sites identified in fieldwalking by looking at the functional categories the types of artefact might represent. For example, it might be argued that remains of weaponry could suggest a military function in the area; or slag remains indicate industrial activity. Haselgrove suggested that some functional categories would need to be determined, and mentioned some which could be considered.

Since the time of his work, however, functional categories have been determined for most finds, by the creation of the AOT. This breaks down finds into classes which could be used to indicate function. Given the difficulties of using Materials and Monument types, as noted above, it was felt that an approach which used these functional classes would work reasonably well. It meant that the Finds were usefully Grouped, but there was still enough information contained in the name of the broad class type to help with interpretation of the data and to indicate potential.

Uncertainty

Once the categorizations of finds have been determined there still existed the problem of data which, for whatever reason, didn't fit the categories. As well as location, there may be other details of the stray finds which may be ambiguously recorded. For example, the number of finds may simply be recorded in words (e.g. some, a few etc). With records of this sort it might be possible to go back to the original sources and get more details, but it may not be possible to give an exact number. In that case either the data must be ignored, or some number must be assigned. This issue was only encountered with one record which was recorded only as being a "small quantity" of finds. This data was given the nominal value of 1 find. However, on a larger scale implementation this issue could become more significant, and more problematic.

Similarly, the object type may be ambiguous. It may not have been recorded in sufficient detail, or it may not have been possible to identify the item at the time the record was made. This is less problematic than the numbers issue, as there are some generic terms which can be used to cover some of these eventualities within the AOT – for example the terms Unidentified object; Lithic Implement or simply Sherd for pottery could be used in many of these circumstances.

However, a third type of uncertainty may exist. A find may have been given a generic date which spans a number of period classes, or it may have a date range which happens to cross the somewhat arbitrary division between e.g. the Medieval and Post Medieval periods. With regard to date, it was decided that it was likely to realistically to fall in one category or another, as generally the date given an object is a creation date. In this case, to resolve any uncertainty, the various possible categories would be given an equal share of the Find. So for example, if three coins of Medieval or post Medieval date were recorded, then each Period would be given 1.5 'points' value for those records.

Uncertainty can also exist due to the nature of the categories used. For example, a spatula comes under both the Class terms Tools and Equipment and Medicine and Pharmacy. In this case it would appear reasonable to argue that the spatula is both at the same time, with the apparent uncertainty reflecting the logic of Thesaurus Construction. In this case, where a find type was linked to two classes, it was decided that the full value should be given to each category.

In the end, issues of uncertainty may not be resolvable due to poor data. In this project any such data has been removed from the model, but it may be desirable to find alternative methods for dealing with this type of uncertainty.

Other issues

During the course of this project, thought has been given as to whether it is desirable to treat the two datasets (i.e. NYCC HER data and PAS data) separately before combining them. This idea was considered as it was hoped to be able to include some analysis of the influence of recovery factors on the data. By treating the data separately, it would have been possible to examine differently recovered data as, broadly speaking, most PAS data is from Metal detecting, and NYCC data incorporates much Antiquarian data. However, in reality there is overlap between the two datasets, and as presently recorded, it is not always obvious how NYCC HER data was recovered, and this would have involved a significant amount of work to gather this data. Furthermore, time would not have allowed anything but a crude analysis of such factors, which it was felt would have been counter-productive. However, the issue of recording recovery methods of finds to allow this sort of analysis is one which is worth pursuing as a project in itself.

ADI - Developing the model

Bearing in mind the various points mentioned above, it was determined to try and create a product that fulfilled the following criteria:

- Was a simple way of showing the presence of archaeological potential based on stray finds data
- Gives some basic information about the nature of the potential
- Consisted of one main GIS layer to make interrogation easy

The process whereby these aims were achieved, i.e. the actual model and processes used, are detailed in the following chapter.

Chapter 4

Building the Artefact Density Index Model

In this chapter, the aim is to present the technical detail behind the process of creating the ADI for each data set, and to discuss how these ADI's were used to create the Development Control ADI's.

Technical Information

The solution was to some extent dictated by the software available and its functionality. The project used ESRI's ArcView GIS version 3.3 with the extension Spatial Analyst version 2.0, which incorporates a module called Model Builder. A further Extension, Grid Machine 6.53 was also used. For the manipulation of data, Microsoft's Access 2003 was used, and some operations were carried out using Excel 2003.

Unit Size

Before developing the model some decisions about the output units had to be made. It was decided that the base unit of output would be an area 100 metres by 100 metres. This unit was chosen for a number of reasons.

An area 100 metres by 100 metres is known as 1 square hectometre (Wikipedia 2005), and has the same area as a hectare. Hectares are a common unit in archaeology, often being the basis for determining the area of survey of a geophysics project. Hectares are also used at NYCC to measure the areas of Parishes; and hectares can also be easily calculated for kilometre squares or 5 kilometre grids (representing OS quarter sheet maps).

This all meant that when calculating the weightings of the various different sized polygons to be used in the model, the unit chosen would help to make the calculations reasonably simple, as it all could be done in hectares from data already available (for e.g. parishes) or easily calculable (for e.g. kilometre squares).

Another reason for selecting a hectometre was that for stray finds located more precisely than this, there is no real difficulty in incorporating them into Development Control decisions, so a smaller grid size was not needed. Furthermore, whilst any unit would involve the possibility of development sites straddling the boundary between units, a hectometre would theoretically allow smaller developments, even up to an individual house plot, to fall within one unit area only, making the decision making process simpler. If a 10 metre grid had been used, most developments would have been bigger than this, meaning many ADI units would need to be assessed before a decision could be reached.

Talking to colleagues also revealed that is at approximately the scale of a few hundred metres being visible on screen that they tend to examine the GIS when making DC decisions (Gail Falkingham: Pers comm), so a larger grid size was felt not to be useful. A kilometre square base unit was considered briefly, but was felt to be too crude a unit to use.

Furthermore, with stray finds recorded to only 6 figure grid references, the area they can fall within is somewhere within a 100 metre square, showing the level of uncertainty of their location. So it was felt that a 100 metre square would usefully reflect the uncertainty of location of stray finds recorded only to 6 figures.

Overall, it was felt that a hectometre was fine grained enough to show changes in stray finds levels across an area, but also of a useful practical size for DC and data management purposes.

The ADI Process

To fulfil the aims of the project, the stray finds needed to go through a number of processes. These can be broken down into 4 main phases, which will be examined in turn (Note this section gives a detailed overview of the process. Additional information details are also in Appendix 5). The four phases were:

Phase 1 - Data preparation
Phase 2 - Vector Conversion
Phase 3 - Creating the Dataset Specific ADI
Phase 4 - Creating the Development Control ADI

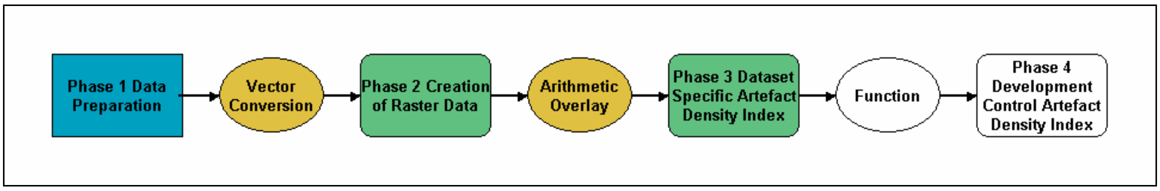


Figure 3. Overview of the ADI Process.

Phase 1 Data preparation

In preparing the data, the process was complicated by the fact that two distinct data sets were being used. The NYCC HER data has been edited to try and make it fit into the Events Monuments Archives model, whereas the PAS data model appears to have been created solely for that project. As such, whilst there are some common problems relating to the data preparation, there are also some issues which are to some extent the result of the way the data has been created in the different systems. NOTE: It should be borne in mind that in all cases the number of records is not the same as the number of finds, as in both data sets the ability to record quantities of artefacts either existed or was added. For example, in the HBSMR database, the ability to record quantities against finds has been added in the latest Version of the software (Version 3). In the PAS data it was assumed that each record represents one object, but closer examination showed this not always to be the case, so a quantity field was added to the cut down database used in this project to record this.

As much of the data preparation was done manually, and time was limited, some pragmatic decisions were made about what data to include and exclude. As the aim of the project was to show how poorly located data could be incorporated into development control decision making, and also to merge data recorded at different precision levels, datasets which would show this process best were prioritised. In effect, what this meant was that data which was only recorded within one dataset (i.e. only in the NYCC HER data OR the PAS data), to only one level of precision (i.e. only to ONE of 5 kilometre, Parish, kilometre or 100 metre precision levels) were not included. Obviously, in a full implementation of this model, all data would need to be included.

NYCC HER data

Apart from the finds data, all the GIS data which was not generated by the project came from NYCC, such as the parish and county areas shown in various figures, and the vector mapping. This data was exported from MapInfo Professional 7.5 in MIF format, and then converted into ArcView Shape files using the MIFSHAPE utility.

All the NYCC Finds data was exported from the Exegesis software HBSMR version 3.06 using the XML export function shipped with the program. This resulted in a number of tables being exported, not all of which were relevant to this project. For example the Export Routine automatically creates tables which can be used to link records to Sources, Designated sites (e.g. Listed Buildings, or Scheduled Monuments) and Archaeological Events, none of which were used in this project. Of the tables actually used, some also had fields which were not relevant deleted. For example, from the Monument table, the Summary field (which is rarely used at the moment in the NYCC HER) was deleted. Other fields, whilst not directly used in the GIS data, helped make sense of some of the ambiguous data: for example, the Description field often helped make sense of other fields, as the text sometimes contained additional information. The data actually used can be found at 00Data\Databases\NYCC1.mdb.

A total of 106 finds records were exported from the NYCC HER for the project area. However, due to the way the HER has evolved, these finds records are not all stray finds records, and so after cleaning, only 56 records remained. Finds were removed from the project data for the following reasons:
Finds records related to stratified monuments or sites (27 Records) – these records were not included in the model as they are already taken into account for DC decisions by the presence of the Monument Record
Finds related to assemblages treated as sites themselves (e.g. Coin Hoards) (15 records) – these records were not included for the same reason given above
Records which represented present object locations rather than finds spots (e.g. A Grave Slab kept at a church) (3 Records)

Poor data (i.e. not enough information to make sense of the record without checking sources which were not in the HER library) (4 Records)
Out of Project Area (1 Record) - originally this record was selected, but on cleaning it was realised it represented a find on a kilometre square grid intersection on the edge of the project area, though the actual square was outside of the project area.

PAS Data

The PAS Scheme very generously provided me with all their data for North Yorkshire. This data was used to help select the Project area, and following this, all the irrelevant finds records had to be removed from the data. This data can be found at 00Data\Databases\ PAS1.mdb – though to conform with a PAS request, the data has been degraded to only show precision to a 4 figure NGR. Some of the data used in the project had more precise locational information, which was what was actually used.

First all the records with Parish level data only were moved into a separate table (0ProjPasParish) in the PAS1 Database. All other records were then mapped using the given NGR, and only those mapped within the project area were selected to be used. These records were then further cleaned after being exported into a separate table within PAS1 (Table 0ProjPasGIS). Some records were deleted due to not enough detail being included with them to be able to determine certain key data (7 records); other records were split as they represented multiple find types. After cleaning, there were a total of 117 records to use in the Project.

There were found to be some issues with the PAS data over and above the general issues with both datasets, though some of these may have been at least partially due to being less familiar with the compilation of the dataset. There were some discrepancies within the data, for example district fields were left blank even though the parish was known – once the Parish is known the District should be able to be entered automatically. Also, some records had an NGR which placed it within the search area, but the parish recorded it as outside the area. These issues of slightly dirty data are fairly common, but did create some minor problems when cleaning the data. In the former case, the issue was ignored, as the district was not relevant to the project. In the second case, it was assumed that the NGR would be more accurate, and this was used in the implementation.

Also some assumptions had to be made about the data, for example the level of precision of recording within the data was not available. For many records it was possible to work out the precision from the NGR recorded. But for other records, the Easting and Northing in GIS format (i.e. twelve figures) only was given, with no NGR, so it was unclear what the precision of recording was. It was therefore assumed that the more zeroes at the end of the number the less precise the record. For example, a record with Easting and Northing of 432000 456000 was assumed to be recorded only to a kilometre square whilst one recorded as 432100 456700 was assumed to be recorded to within 100 metres. However, it is possible that in both cases the actual record was recorded to a precision of 1 metre, but it was impossible to tell.

These issues were not insurmountable but might mean that additional data about the precision of the data might be desirable to be recorded by the PAS.

Data into GIS for ADI Processing

Once the data was prepared, it needed to be imported into the GIS in the appropriate format for the ADI Process. In a larger project this process would need to be an automated process, however, to be able to achieve this, a number of issues (see below) would need to be able to be dealt with by an automated process. In this project the data was entered manually.

There were basically four precision levels that the data could be entered at:
The Quarter Sheet level – representing a possible 5 kilometre square area and equating to the table and theme called proj5k in the Project GIS (all this data is available on the CD in the folder 00Data\zProjectData).
The Parish level – representing a variably sized parish area and equating to the table and theme called projparish in the Project GIS.
The kilometre Square level – representing a one kilometre square area and equating to the table and theme called proj1k in the Project GIS
The 100metre level – representing those finds located to within 100 metres or more precisely located finds and equating to the table and theme called 100mgrd in the Project GIS.

Within each of these tables the appropriate fields need to be created to allow data to be entered. This was done manually. To help with the model creation process, a naming convention was created for the fields. The basic convention was that the name would have two elements with two letters representing the type of data, and two letters for the period. So, for example data from the Armour and Weapons Category from the Early Medieval Period would have the code AwEm; data from the Currency Category from the Medieval Period would be CuMe. The full lists of the codes chosen are given in Appendix 3 (though not all were eventually used – see above). Once the fields were created, the records were entered, with the totals being determined from the various tables using queries.

It may be questioned why the data needed to be entered into a vector layer in the first place – why was the raster data not created directly? The main reason for this is that the various vector mappings were already available from NYCC HER, and appending extra data to new fields to this would be relatively easy. Also, it would be easier to update this vector data and then re-run the ADI Models than re-create new Raster data as new data was entered into either the NYCC HER or PAS Datasets. Finally, the possibility of automatically importing the data into the vector layers has been briefly examined, though not implemented due to time constraints, and it is highly likely that this process could be automated, once the data was cleaned, speeding up the whole ADI Process.

Phase 2 Vector Conversion

Once the data had been entered into the appropriate tables, the actual processing of the data to form the ADI could begin. The first process was to create the raster data from the vector data.

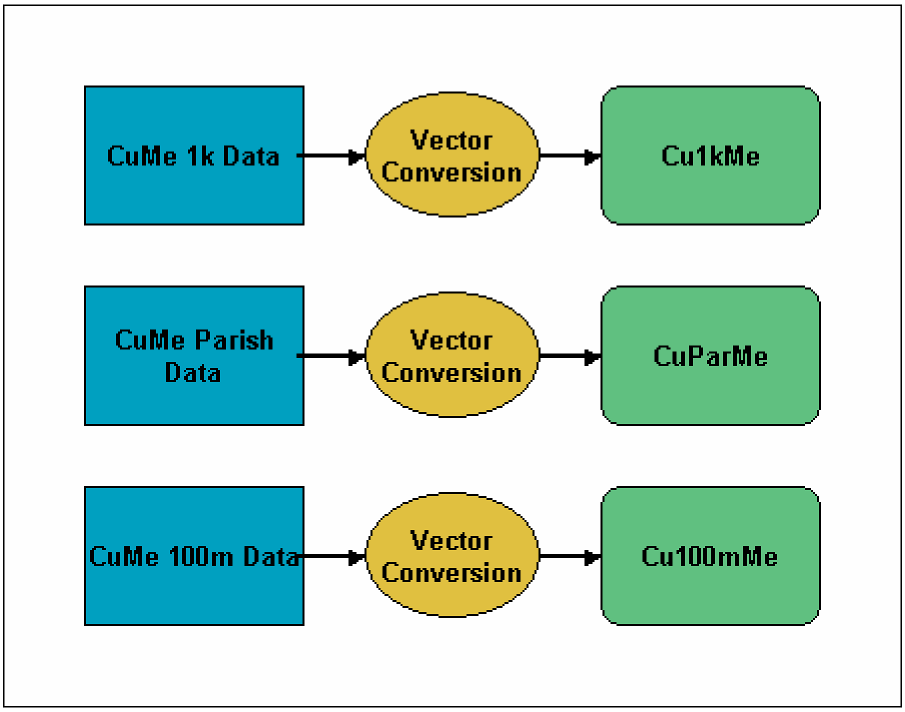


Figure 4. Example of the Vector conversion model for the Currency (Cu) Medieval (Me) ADI Process.

Essentially this involves converting data from the format in Figure 5, to the format shown in Figure 6.

This was done using the standard Vector Conversion functionality within ArcView. Again a naming convention was designed for the Themes created which contained 3 elements – the two already discussed above plus an additional one showing the precision of the dataset. So, for example, data for Currency of the Medieval period for 100 metre accuracy would be called Cu100Me (See appendix 3 for the list of codes used). In each case the vector data was converted into raster grids of 100 metre square size.

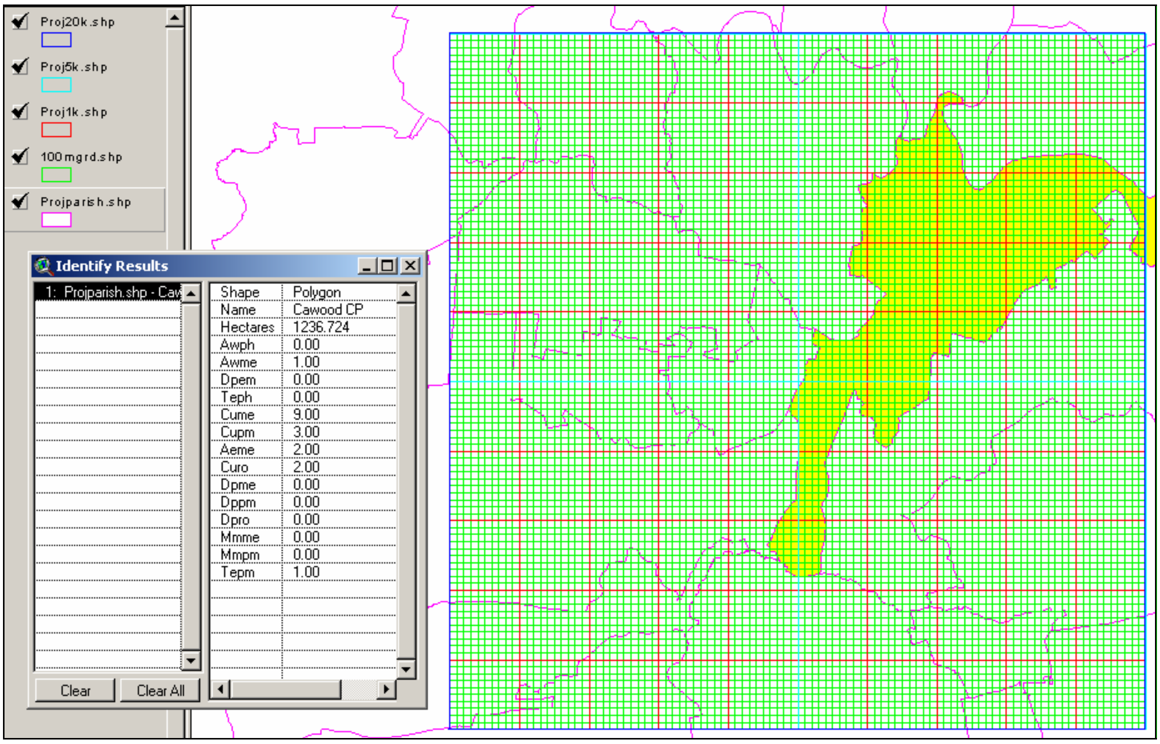


Figure 5. Screen Shot from ArcView 3.3 showing the vector data for the highlighted parish of Cawood.

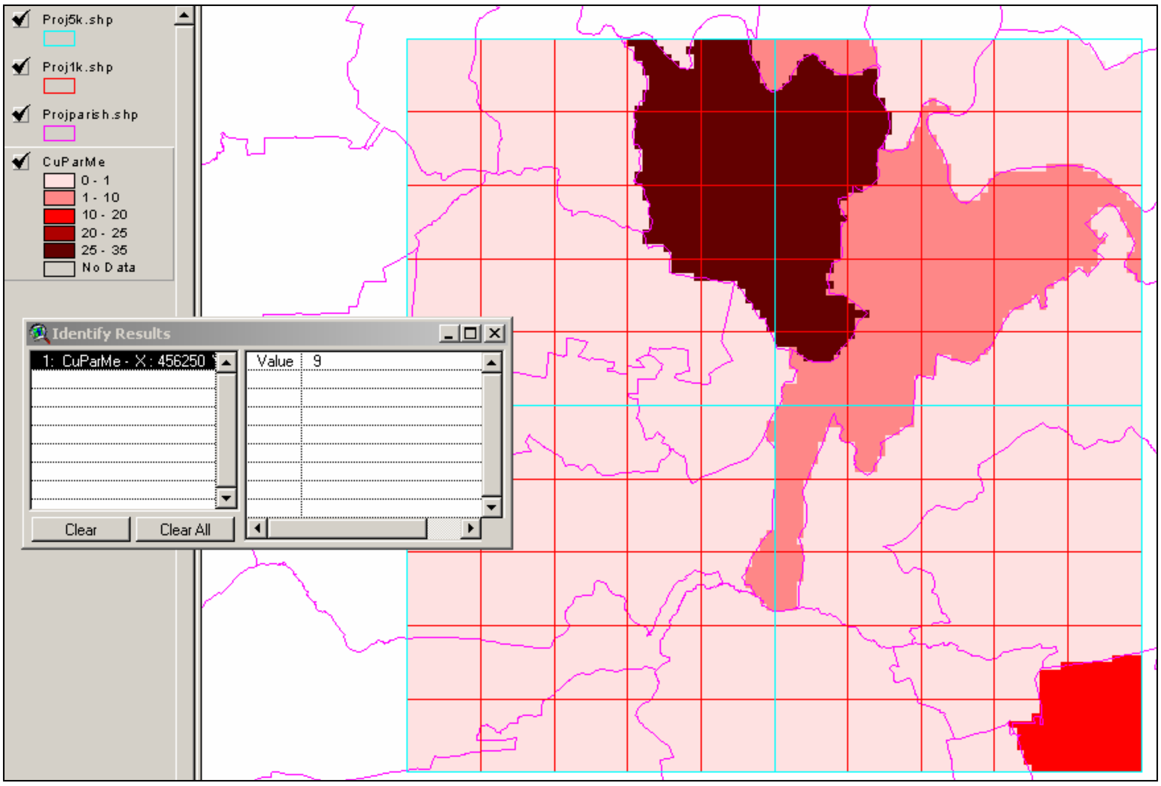


Figure 6. Screen Shot from ArcView 3.3 showing the Raster data for the CuMe ADI Process for the Project Area.

Phase 3 Creating the Dataset Specific ADI

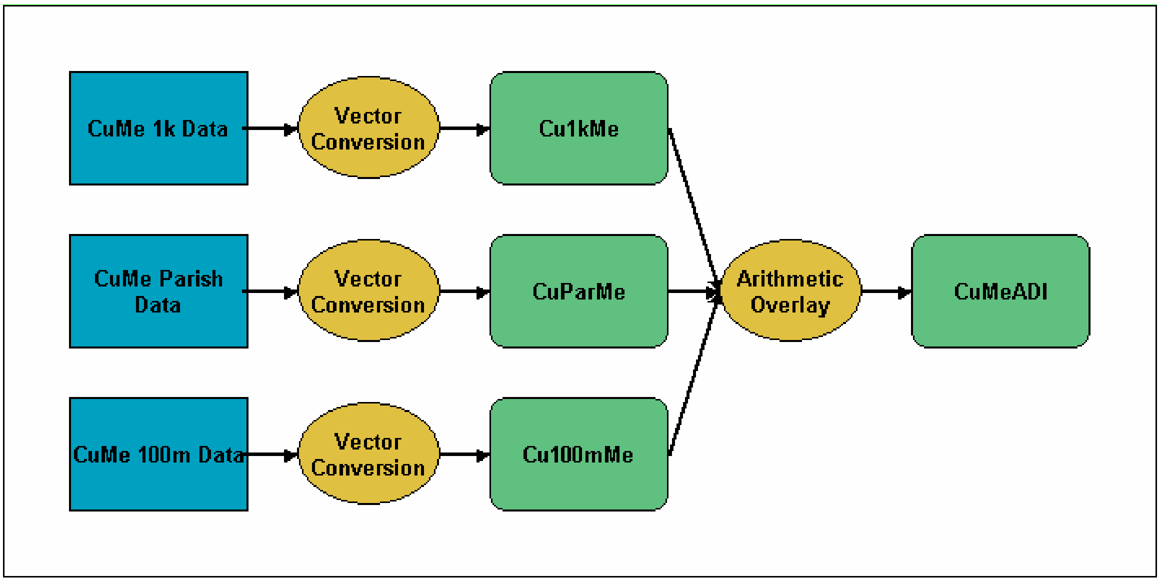


Figure 7. Example of the Dataset Specific ADI model for the Currency (Cu) Medieval (Me) ADI Process.

Once all the data had been converted into raster data, the dataset specific ADI could be created (i.e. separate ADI's for e.g. Medieval Armour and Weapons; Post Medieval Currency, Medieval Currency etc). The naming convention for this was the same as for the Vector Conversion, with the addition of a suffix of ADI. So, for example, the ADI for Currency of Post Medieval Date would be named CuPmADI.

The process of creating the overlay involved the adding together of the different raster layers, using the Arithmetic Overlay function in Model Builder. (This process is shown graphically in figures 8 to 11). This allows the different themes to be weighted as they were added together. The weighting was used to reflect the precision of the finds location based on the original size of the polygon. As the base unit was a hectometre (which has the same area as a hectare), hectares were used as the base unit for calculations.

The hierarchy which was used to determine the weighting was based on the relative sizes of the areas and was as follows: 100 metre square (base unit); kilometre square; Parish and then Quarter sheet. The logic of this was that the average size of parishes was calculated to be 894 hectares, which put them in between the 100 hectares of a kilometre square and the 2500 hectares of the quarter sheet. All parishes were weighted the same, regardless of size, for simplicity.

This gave the following multiplier values, which were used in the ADI:

100 metre Multiplier = 1
1 kilometre Multiplier = 0.01
Parish Multiplier = 0.0011
Quarter sheet Multiplier = 0.0004

Once the ADI had been calculated, the theme was saved as a grid Theme with a similar naming convention to that mentioned above, but with the addition of the term ADIGrd on the end. So for example, the ADI for the Currency Class of Medieval period was given the name CuMeAdiGrd.

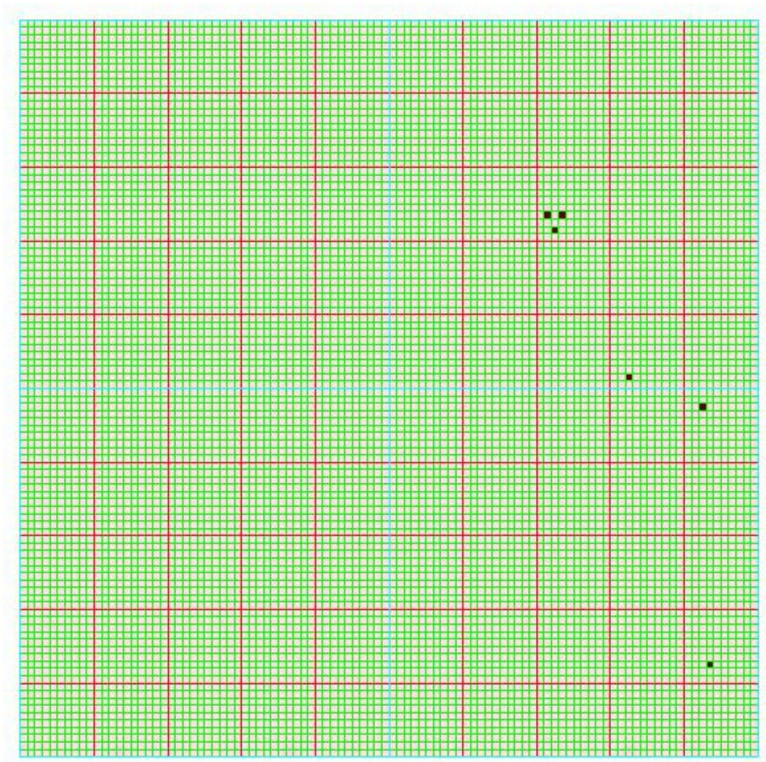


Figure 8. The Raster Grid created for the 100 metre Precision level for the Currency (Cu) Medieval (Me) ADI Process.

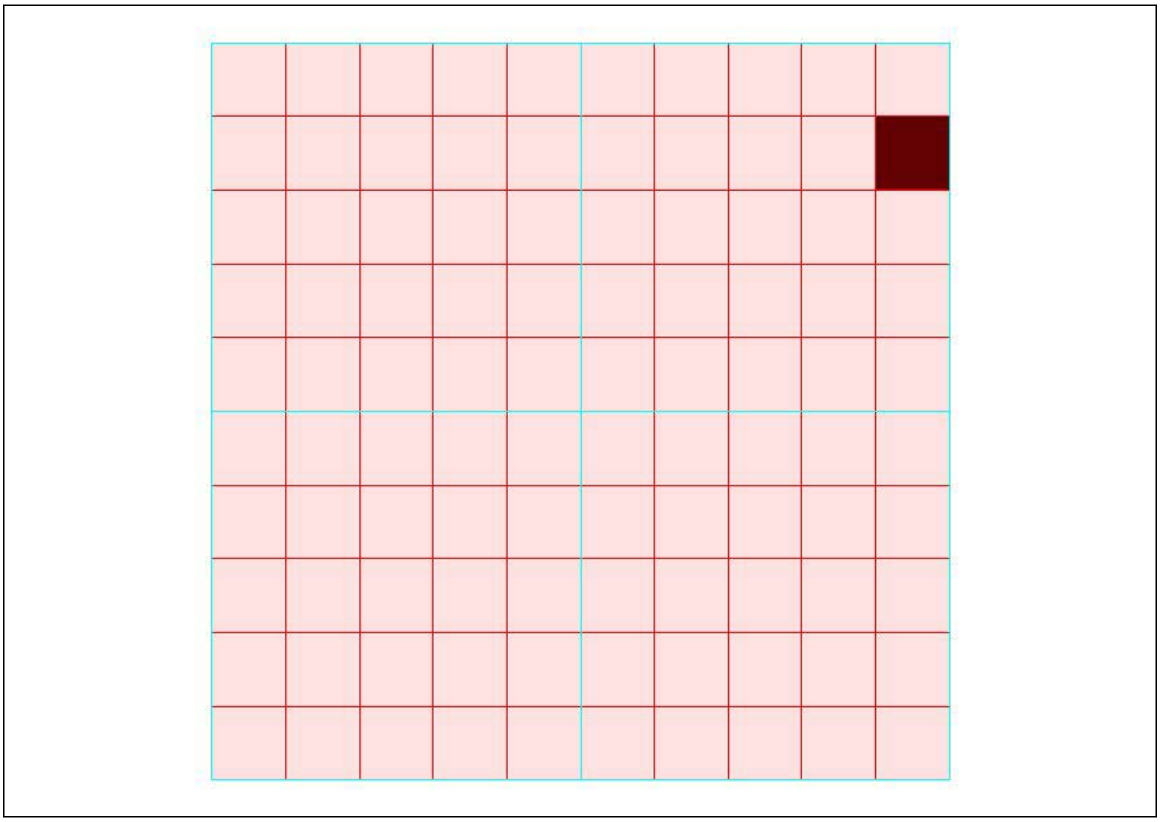


Figure 9. The Raster Grid created for the kilometre Precision level for the Currency (Cu) Medieval (Me) ADI Process.

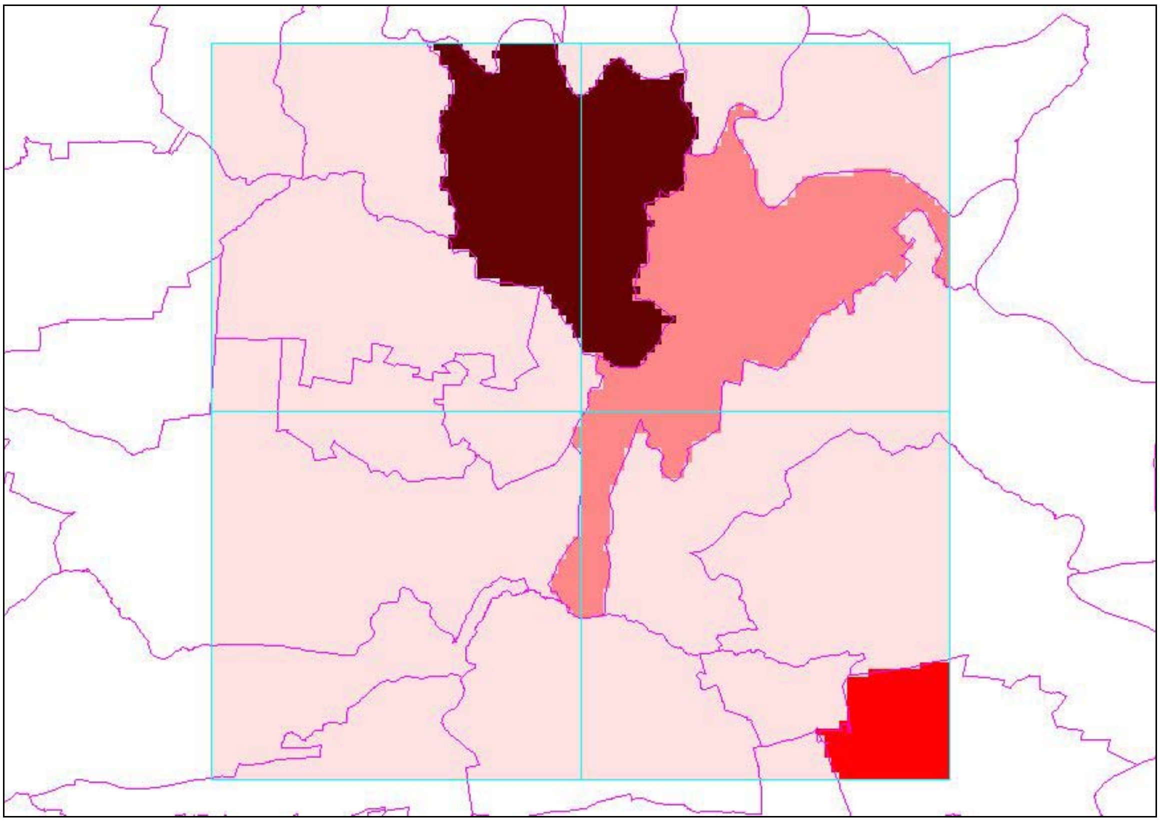


Figure 10. The Raster Grid created for the Parish Precision level for the Currency (Cu) Medieval (Me) ADI Process.

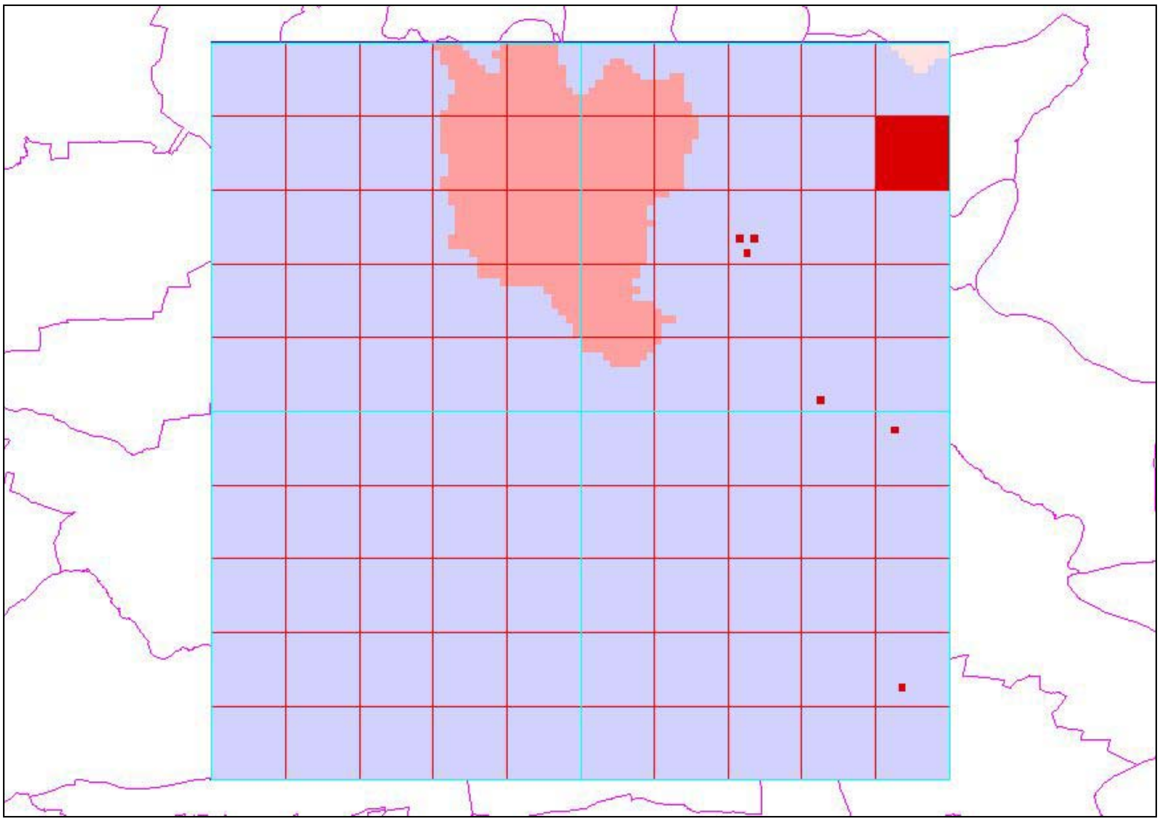


Figure 11. The Currency (Cu) Medieval (Me) ADI layer created from the data in Figures 8, 9 and 10.

Phase 4 Creating the Development Control ADI

Phase 3 results in a multitude of layers representing the various ADI's for different finds classes and periods. As the original aim was to create only one layer to be interrogated for DC purposes, a number of possible solutions were examined and are presented. In all cases the naming convention for the final Development Control ADI was simply DCADI with a Version number to differentiate the different examples.

This involved running another model named after the relevant ADI Model. In each case these new models were run based on the ADIGrd layers which were the outcomes of Phase 3 above.

Development Control ADI Version 1 (DCADIV1)

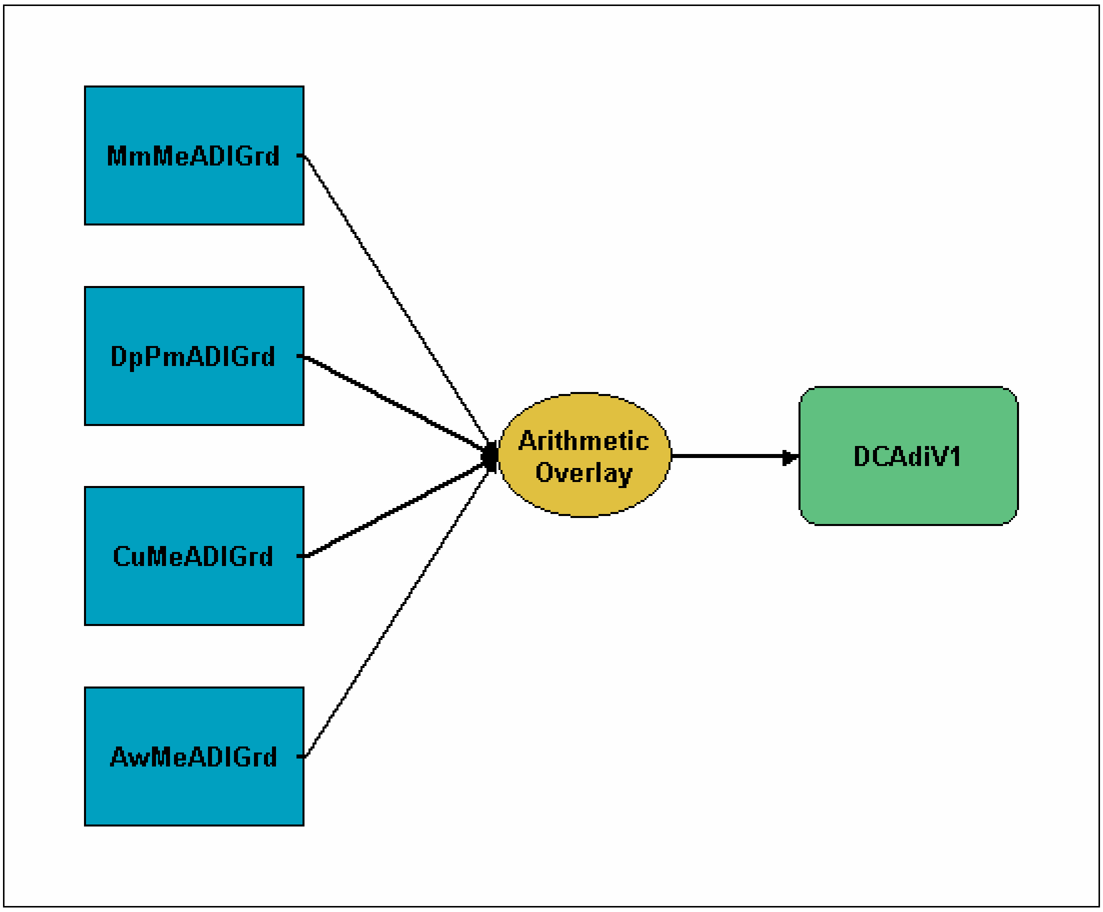


Figure 12. The Model used for creating the DCADIV1. Note that more than 4 ADI's were created in the project, only four are displayed for clarity.

This method used a simple extension of the logic used to create the base ADI's – that the merging of raster data would give an indication of the potential for the area. To that end, an Arithmetic overlay was used which simply added together the values for all the ADI's to produce an overall score for each hectometre.

This data was then displayed using the Standard deviation Classification in the Legend Editor Dialog box, as the default display made it appear as if there areas with no data.

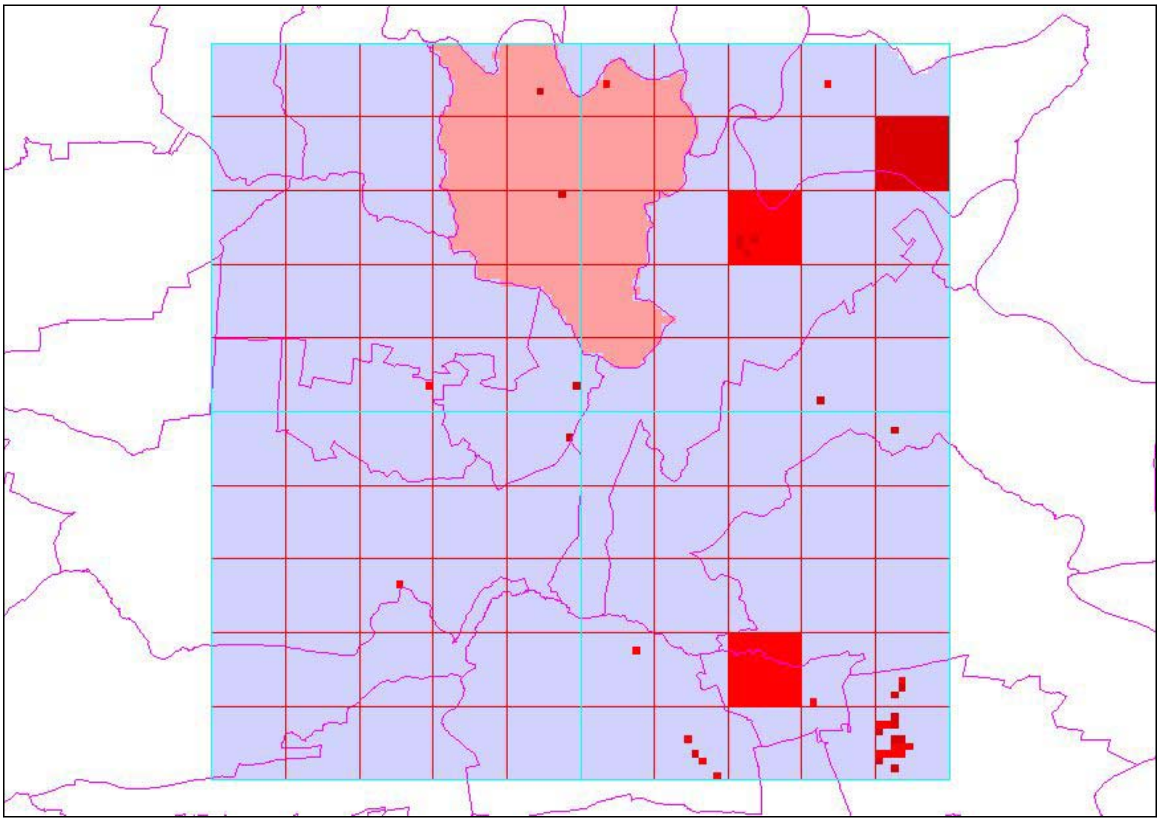


Figure 13. The full extents of DCADIV1.

In a real life situation, the ADI would need to be saved as a Grid theme and incorporated into the normal HER GIS, as shown in Figure 14.

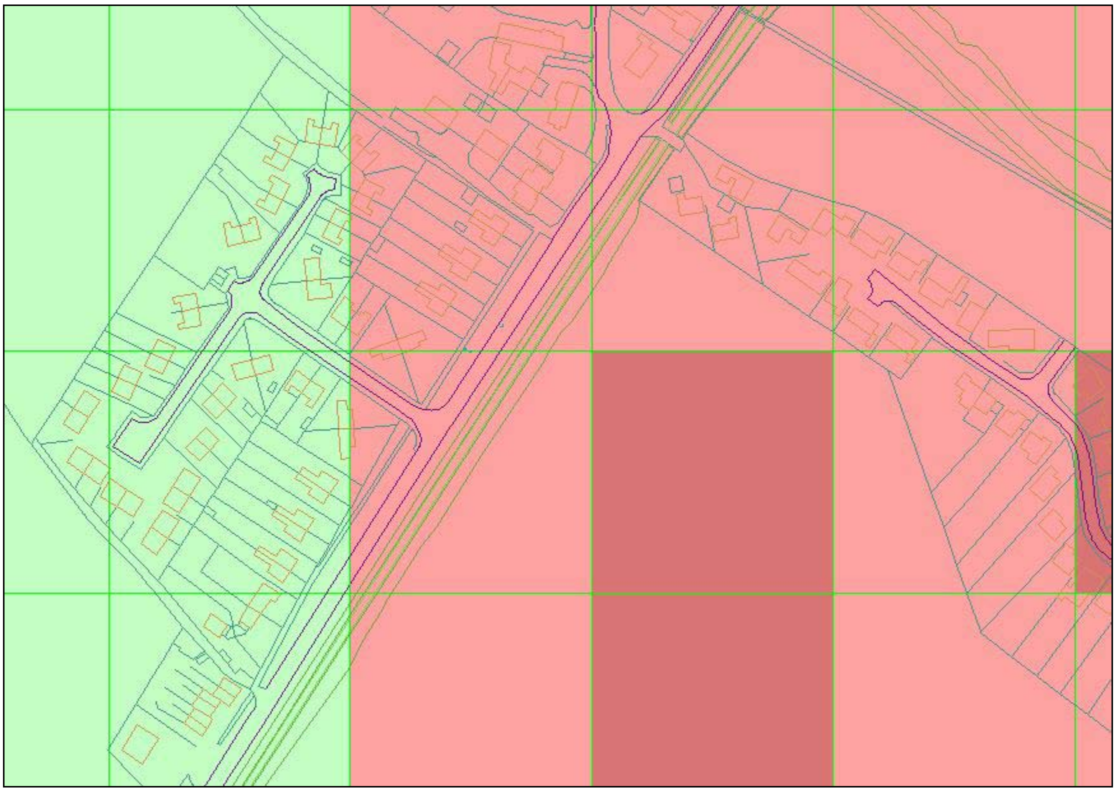


Figure 14. The DCADIV1 displayed against vector mapping. Crown Copyright North Yorkshire County Council Licence No. 100017946 (2005).

Development Control ADI Version 2 (DCADIV2)

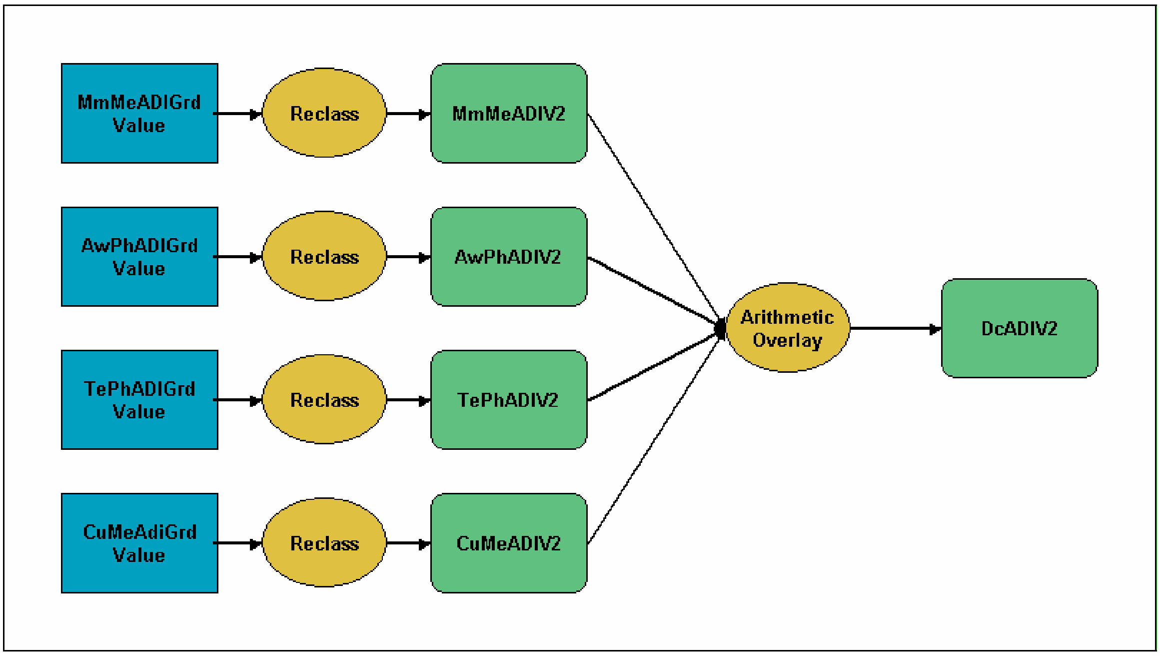


Figure 15. The Model used for creating the DCADIV2. Note that more than 4 ADI's were created in the project, only four are displayed for clarity.

In this DCADI Version, the idea of a 'Red flag' layer is used, with hotspots being shown in one colour, and other areas as another colour. This is a two part process.

In the first part, each ADI is reclassified in to a 'Red Flag' model, with all scores over a set threshold value being reclassified with a value of one, and all scores below this threshold being reclassified as a value of 0. Each reclassified dataset specific ADI is then saved as a grid theme, using the established naming convention with the suffixes Rf to denote Red Flag and Grd for Grid. So established the Red Flag theme for Medieval Currency is called CuMeRf and saved as the Grid Theme CuMeRfGrd, However, determining the Threshold level for the Red Flag is not a simple process.

Thresholds

Determining threshold levels was difficult. In the trial run ADI using Materials for Grouping already mentioned, a simple method to determine the background level of finds for any particular relevant class and period e.g. Roman pottery; Roman Coins, which was called the Flag level, was tried. This involved simply dividing the number of finds by the area of the project.

The idea was that the model would then compare the recorded level of finds in each area against this Flag Level. If the measurement exceeded this value then this would be enough to 'raise a flag' for that class of data. This proved problematic in that the method chosen produced very small thresholds. This meant that for some ADI's the entire dataset was above this low threshold level.

Eventually, another method was determined. This used the idea that anything within one Standard deviation from the mean was less likely to be significant. The value used to determine the Threshold layer was therefore the maximum value of one standard deviation from the mean. This can be calculated in ArcView by using the Standard deviation Classification in the Legend Editor Dialog box for each layer. This shows the values for the mean, one standard deviation, 2 standard deviations etc, within the Dialog Box (see Figure 16). The value for the interface between one and two Standard Deviations was used as the critical value, with anything over this maximum value being deemed significant and triggering the Red Flag. Any value over the Red flag trigger would be reclassified to have the value one, and anything below the value zero. This process is shown graphically in Figures 17 to 18.

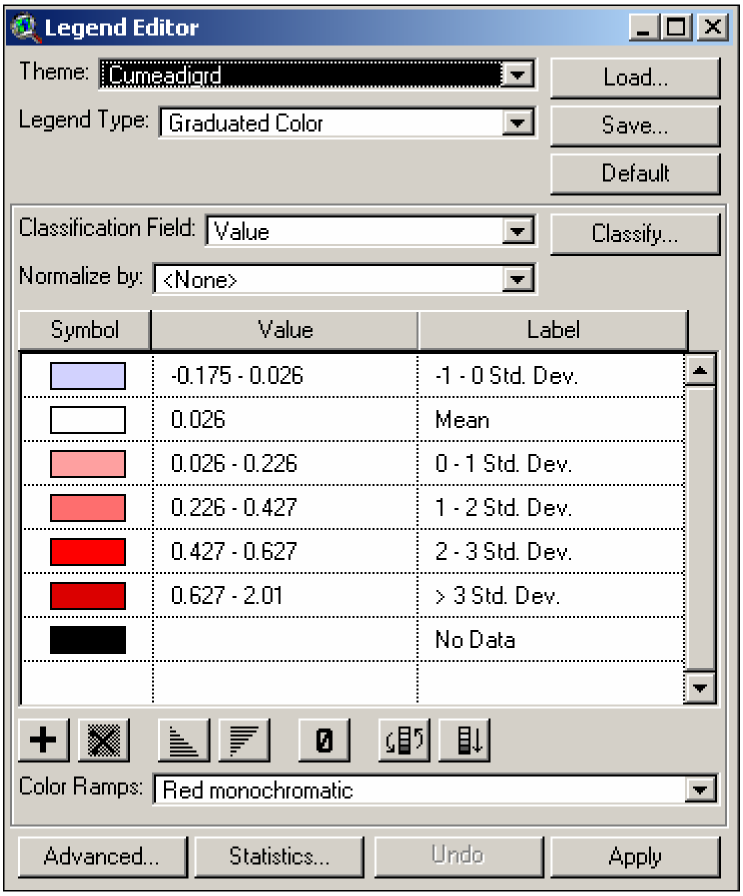


Figure 16. The ArcView Legend Editor Dialog Box with the Classification set to Standard Deviation. In this example, the critical value used would be 0.226.

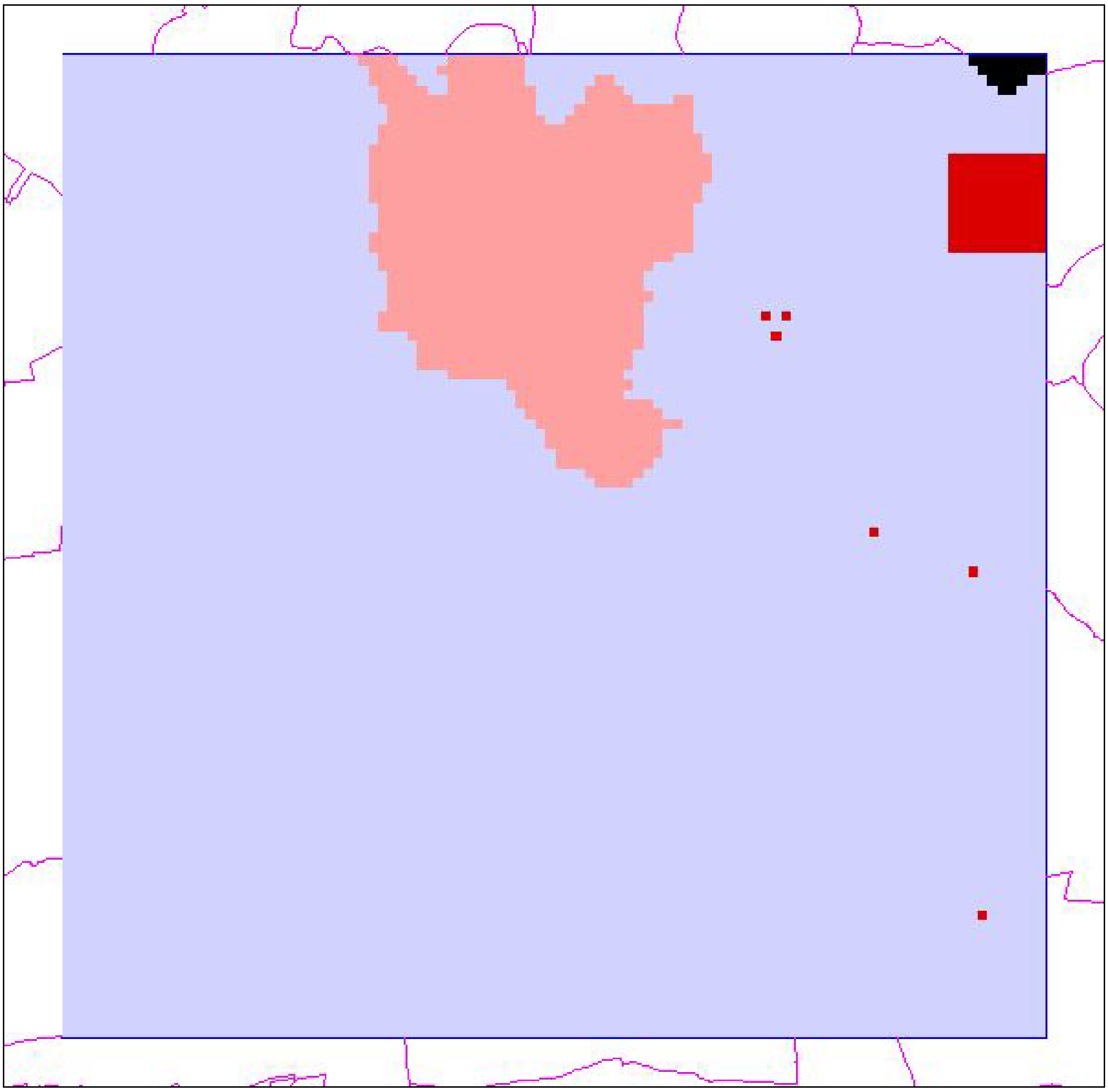


Figure 17.The CuMe ADI prior to reclassification.

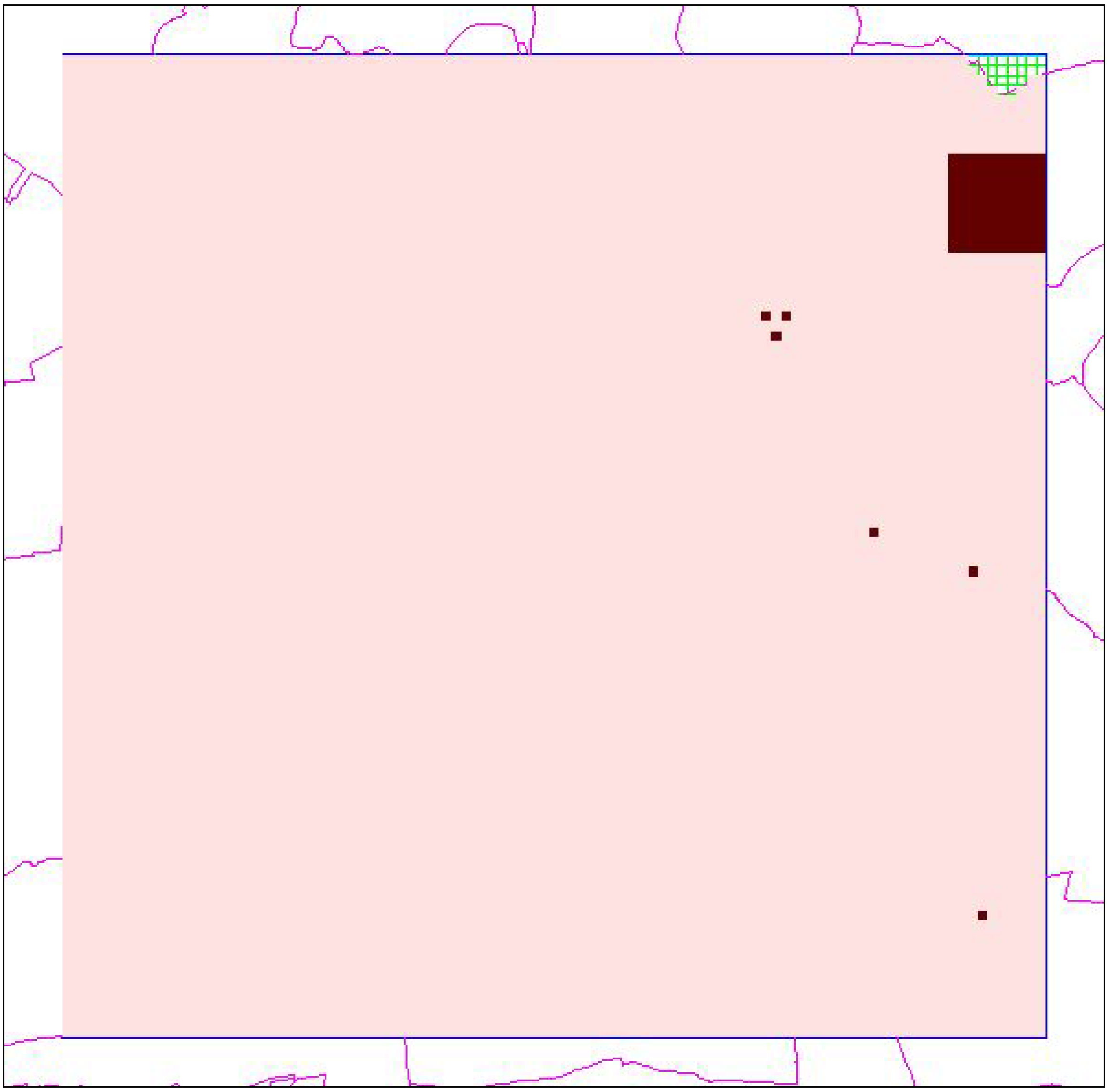


Figure 18. The CuMe ADI after reclassification.

Once each ADI has been reclassified in this way, a simple arithmetic overlay is used to merge all these values by simple addition. Mapping these values gives a simple way of showing Red Flag areas within the project area. This is shown in Figure 19.

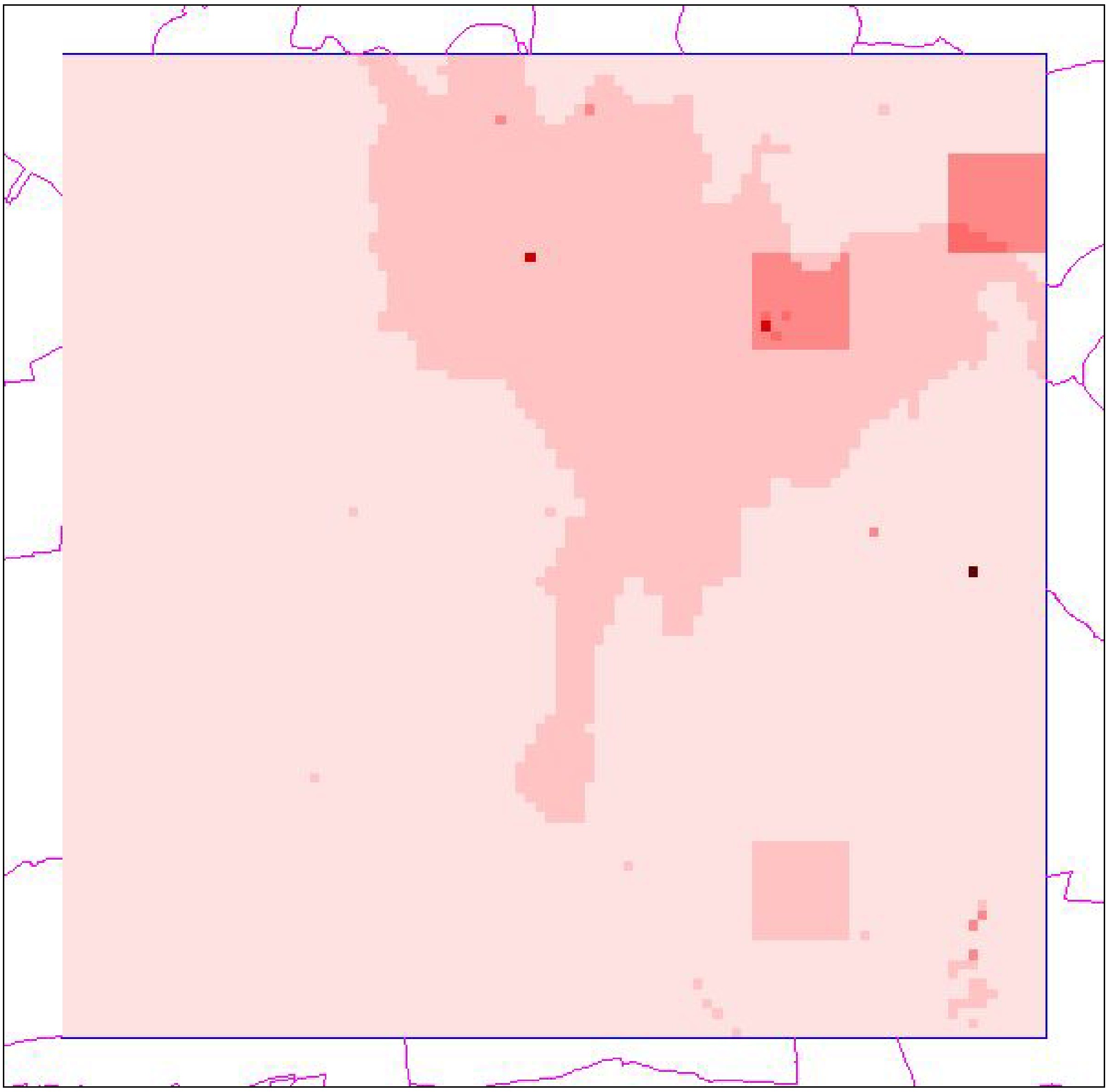


Figure 19. The DCADIV2, with all the reclassified ADI's having been merged into one layer. Experimentation has shown that the optimal way to display this layer appears to be using the Equal Interval classification, with the number of classes set to one more than the maximum value. E.g. in this example the maximum value is 7, so the Number of classes has been set to 8.

In a real life situation, the ADI would need to be saved as a Grid theme and incorporated into the normal HER GIS.

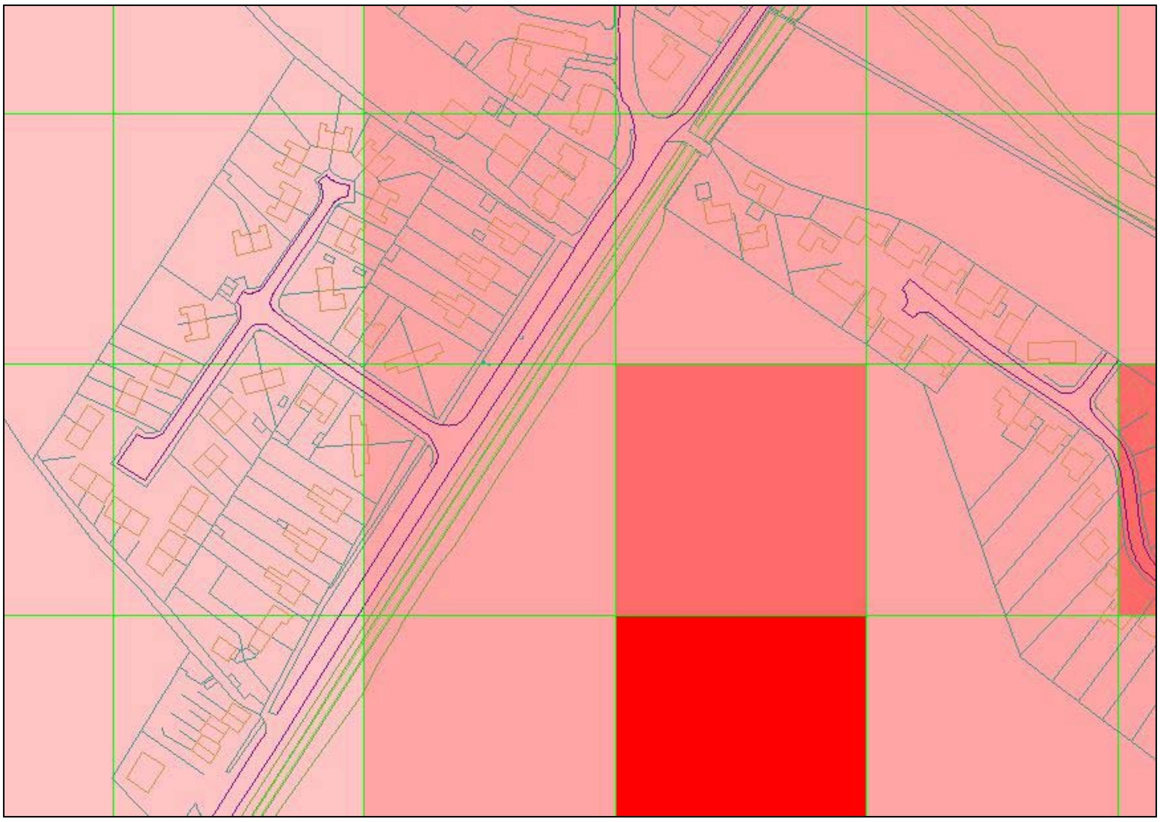


Figure 20. The DCADIV2 displayed against vector mapping. Crown Copyright North Yorkshire County Council Licence No. 100017946 (2005).

Development Control ADI Version 3 (DCADIV3)

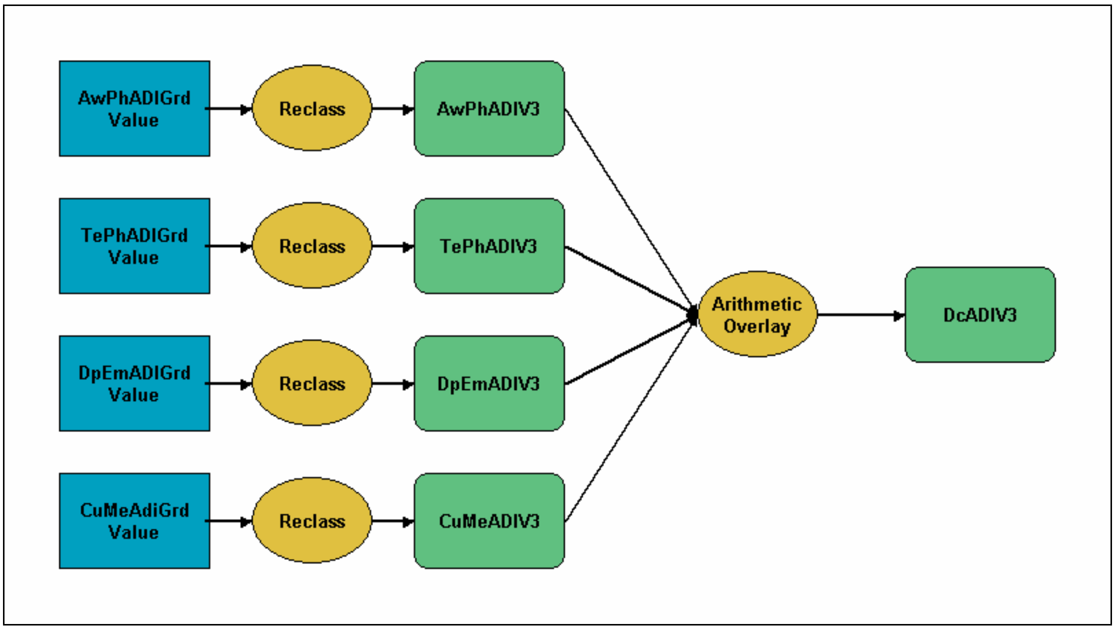


Figure 21. The Model used for creating the DCADIV3. Note that more than 4 ADI's were created in the project, only four are displayed for clarity.

DCADIV3 is a development of Version 2, which was suggested to me by Larry Theller of the Center for Advanced Applications in GIS at Purdue University, Indiana who I encountered via an email list for ArcView users called ARCVIEW-L list.

Originally, Mr Theller had used this method for classifying soil types, but its use could be applied to the DCADI.

In a similar way to Version 2, first each layer must be reclassified, and the reclassification methodology is the same as in Version, except for one important difference. Instead of the reclassification values always being one and zero, the values are zero and one for the first ADI, Zero and 10 for the second ADI, zero and 100 for a third ADI, and so on, with each additional layer being multiplied by a factor of 10 for the Yes value (1000, 10000 etc).

What this means is that when the layers are merged with each other, the resulting ADI has a value for each square which looks something like a string of binary e.g. 1001; 110 etc.

What this in fact means is that the Red Flag value is shown by each 1 and the no value by 0. This allows you to work out which ADI produced Yes values for the DCADI in each square and therefore go back to look at this original ADI for further information if need be. For Example 1001 means that the Red Flag was triggered for ADI's 1 (starting from right to left) and Four. If it is known which ADI's this represents, then they can be viewed for additional data.

The process for this is effectively the same the DCADIV2, except that in the Arithmetic Overlay which joins all the reclassified ADI's together, the multipliers for the different layers need to be altered as noted.

There were some problems with this DCADI, which will be fully discussed in the next chapter. However, a working version might appear something like that shown in Figure 22.

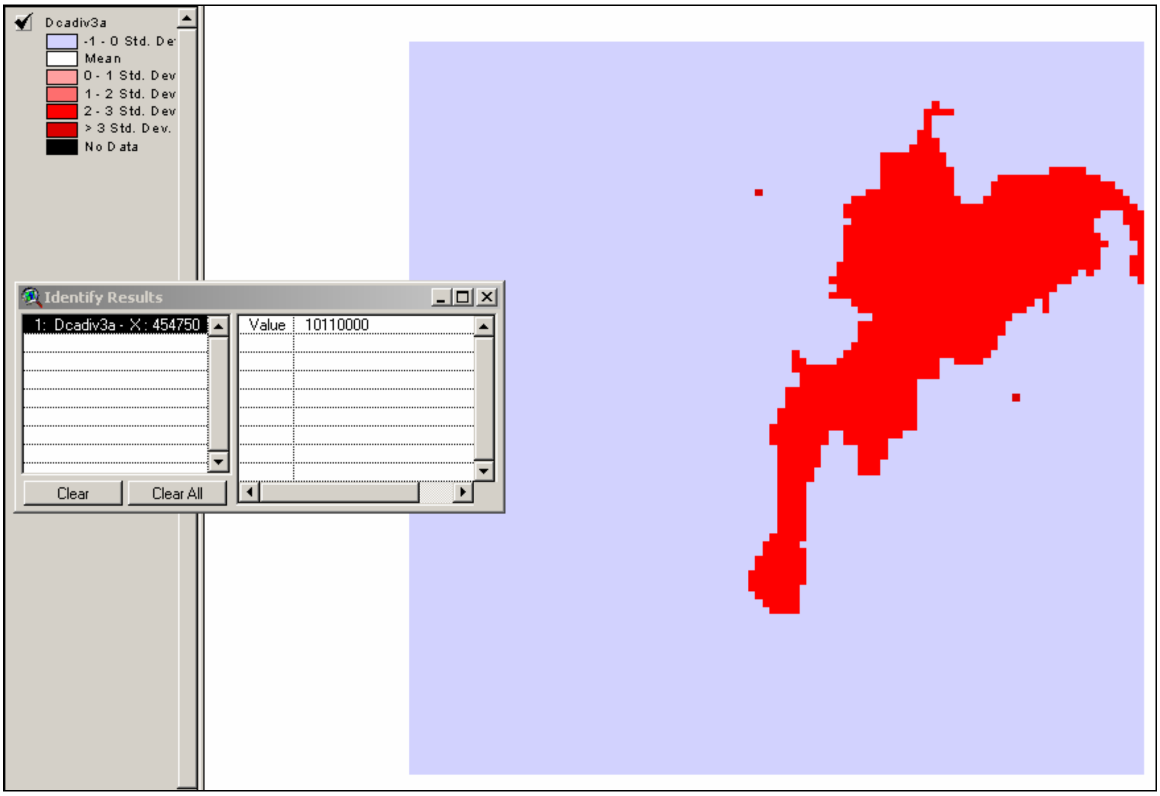


Figure 22. The full extents of a partial implementation of DCADIV3. (see next chapter for discussion of issues).

Development Control ADI Version 4 (DCADIV4)

The preceding ADI's were unsatisfactory for a variety of reasons, and through experimenting with them, it began to emerge that the best solution would be to try and get the data back into some sort of vector format. A number of possible methods were considered for doing this, but in the end the following method was used.

The Extension Grid Machine Version 6.53 was used to convert each of the ADI's into a points theme.

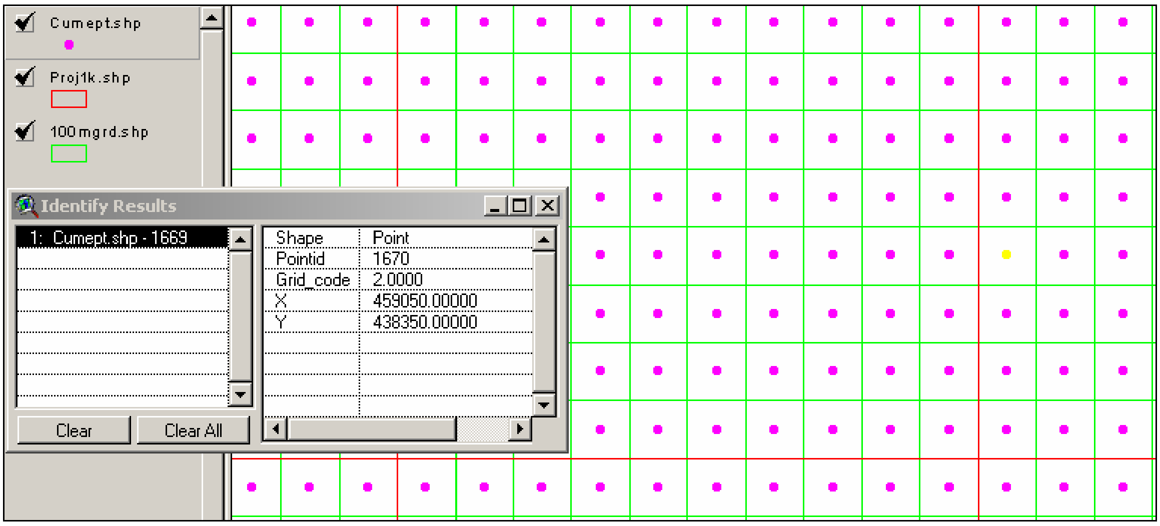


Figure 23. ArcView screen grab showing the point data as created by Grid Machine. The Grid_Code field shows the ADI value.

Each of these point themes was then exported into an Access database, where they were merged into one table, based on a common ID field, which was created automatically during the conversion.

This merging caused some minor problems, as not all the layers had the same number of points, due to areas of missing data. Some careful editing of the data was needed to make sure that the links were made between the correct points. This was fairly easily done, if a little tricky.

The result was an Access table which contained all the appropriate information – an ID field and the value for each ADI in a separate field.

0DcADIV4 : Table

ID	X	Y	AeMe	AwMe	AwPh	CuMe	CuPm	CuRo	DpEm	DpMe	DpPm	DpRo	MmMe	MmPm	TePh	TePm	UaMe
846	454550	439150	0	0	0.0004	0.035	0.0242	0.01	0	0.0011	0	0	0.028	0.0044	0	0	0
847	454650	439150	0	0	0.0004	0.035	0.0242	0.01	0	0.0011	0	0	0.028	0.0044	0	0	0
848	454750	439150	0	0	0.0004	0.035	0.0242	0.01	0	0.0011	0	0	0.028	0.0044	0	0	0
849	454850	439150	0	0	0.0004	0.035	0.0242	0.01	0	0.0011	0	0	0.028	0.0044	0	0	0
850	454950	439150	0	0	0.0004	0.035	0.0242	0.01	0	0.0011	0	0	0.028	0.0044	0	0	0
851	455050	439150	0	0	0	0.035	0.0242	0.01	0	0.0011	0	0	0.028	0.0044	0	0	0
852	455150	439150	0	0	0	0.035	0.0242	0.01	0	0.0011	0	0	0.028	0.0044	0	0	0
853	455250	439150	0	0	0	0.035	0.0242	0.01	0	0.0011	0	0	0.028	0.0044	0	0	0
854	455350	439150	0	0	0	0.035	0.0242	0.01	0	0.0011	0	0	0.028	0.0044	0	0	0
855	455450	439150	0	0	0	0.035	0.0242	0.01	0	0.0011	0	0	0.028	0.0044	0	0	0
856	455550	439150	0	0	0	0.035	0.0242	0.01	0	0.0011	0	0	0.028	0.0044	0	0	0
857	455650	439150	0	0	0	0.035	0.0242	0.01	0	0.0011	0	0	0.028	0.0044	0	0	0
858	455750	439150	0	0	0	0.035	0.0242	0.01	0	0.0011	0	0	0.028	0.0044	0	0	0
859	455850	439150	0	0	0	0.035	0.0242	0.01	0	0.0011	0	0	0.028	0.0044	0	0	0
860	455950	439150	0	0	0	0.035	0.0242	0.01	0	0.0011	0	0	0.028	0.0044	0	0	0
861	456050	439150	0	0	0	0.035	0.0242	0.01	0	0.0011	0	0	0.028	0.0044	0	0	0
862	456150	439150	0	0	0	0.035	0.0242	0.01	0	0.0011	0	0	0.028	0.0044	0	0	0
863	456250	439150	0	0	0	0.035	0.0242	0.01	0	0.0011	0	0	0.028	0.0044	0	0	0
864	456350	439150	0	0	0	0.035	0.0242	0.01	0	0.0011	0	0	0.028	0.0044	0	0	0
865	456450	439150	0	0	0	0.035	0.0242	0.01	0	0.0011	0	0	0.028	0.0044	0	0	0
866	456550	439150	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
867	456650	439150	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
868	456750	439150	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
869	456850	439150	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
870	456950	439150	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
871	457050	439150	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 24. Example data from the Access Table showing the various ADI fields with their associated data.

This table was exported from Access as Dbase V file called 0DCADIV4.dbf.

At this point, a copy of the Vector 100mGrd Shape file was made, and renamed DCADIV4. This file was edited in ArcView to remove all the fields except the ID field.

Finally, it was possible to bring the 0DCADIV4.dbf table into ArcView and link it to the Shape File DCADIV4, using the Join Command, as shown in Figure 25.

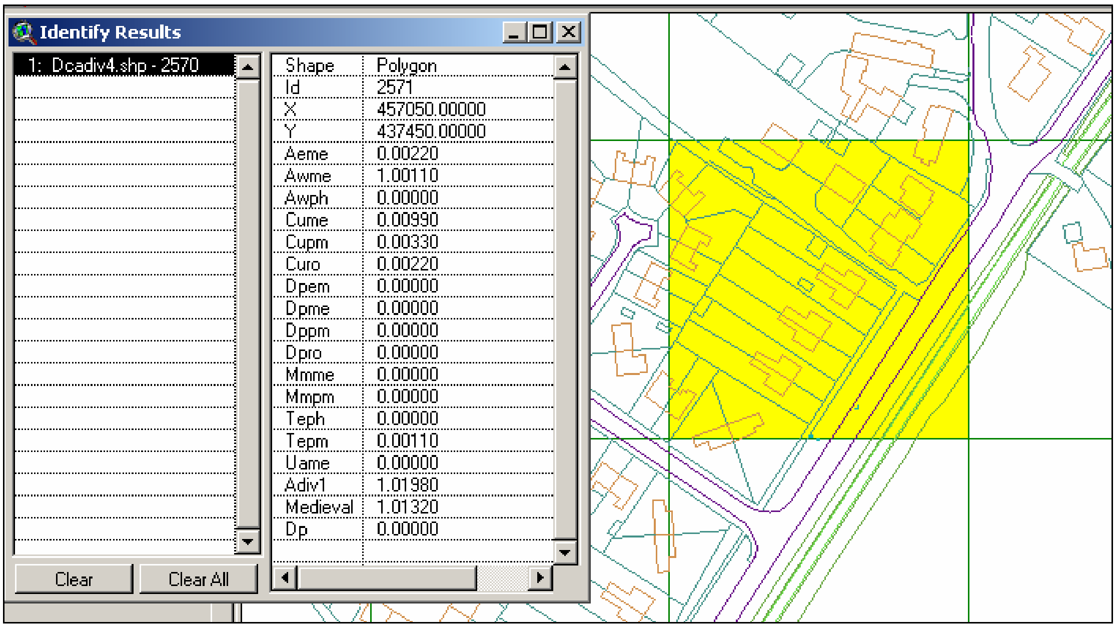


Figure 25. ArcView Screen grab Showing the DCADIV4 data for the highlighted 100 metre square. Crown Copyright North Yorkshire County Council Licence No. 100017946 (2005).

Chapter 5

The ADI Solution in Action.

In this chapter the aim is to compare the different DCADI's developed, and to compare their strength and weaknesses.

Results

As a result of the carrying out of the various processes detailed in chapter 4, 15 ADIs were produced (detailed below), which can be viewed as images in Appendix 4.

- AeMe (Animal Equipment Medieval)
- AwMe (Armour and Weapons Medieval)
- AwPh (Armour and Weapons Prehistoric)
- CuMe (Currency Medieval)
- CuPm (Currency Post Medieval)
- CuRo (Currency Roman)
- DpEm (Dress and Personal accessories Early Medieval)
- DpMe (Dress and Personal accessories Medieval)
- DpPm (Dress and Personal accessories Post Medieval)
- DpRo (Dress and Personal accessories Roman)
- MmMe (Measurement Medieval)
- MmPm (Measurement Post Medieval)
- TePh (Tools and Equipment Prehistoric)
- TePm (Tools and Equipment Post Medieval)
- UaMe (Unassigned Medieval)

These derived data have value themselves for determining the distribution of Finds types over an area, and are key to the DCADI process. However, the real test has been to produce the DCADI's. Four of these have been described, all of which have advantages and disadvantages, which will be examined below.

Advantages and disadvantages of the Different DCADI's

Development Control ADI Version 1 (DCADIV1)

The key advantage of this version of the DCADI is it simplicity. It follows the logic of creating the ADI's by merging data through to a logical conclusion, which give the final DCADI as being the sum of all the sub ADI's.

The other advantage this has is that, compared to the Flag models, it is the raw data which is being seen, so it is possible for the archaeologists to compare the data in the area being interrogated with the surrounding areas and interpret it in their own way, rather than having to accept a set interpretation (see below).

However, there are some issues with this DCADI. The first, is that the merging of all the data means some individual ADI's which may have high values in certain areas, could have these values masked by the merging with all the other ADI's which might have the overall effect of smoothing the data. A suggestion for dealing with this was unable to be examined but is mentioned here. It was suggested by Julian Richards that it might be possible to use an exponential scale when merging data, so that high values were exaggerated to make them more obvious in the final DCADI.

Another issue with this model is that as the data is merged, it is not possible to see which ADI's have contributed to a particularly high value – it could be a few key ADI's with extremely high values suggesting certain types of potential in key periods, or a generally high layer for many ADI's suggesting a more sustained, long term activity. The difference between these types of potential may well suggest different mitigation, but it would be hard to determine which mitigation is appropriate. This issue is compounded by the fact that it is not possible to go from the DCADI back to the appropriate ADI's to see what the key values are. In other words, there is no link between the DCADI and the ADI's which produced it, the only way to determine which ADI's produce a high DCADI value is to look through all of them – which defeats the purpose of the DCADI to some extent.

Development Control ADI Version 2 (DCADIV2)

This second ADI suffers from some of the same disadvantages of DCADIV1 – the inability to link back directly to the original ADI's for comparison.

Even the advantage of having a flag to highlight key areas, where it has been determined that the Value in at least one ADI is sufficiently high to be of note, is mitigated against by the whole issue of Thresholds. As well as the issue of how to calculate the Threshold level (see Chapter 4) there is also the issue that using Thresholds mean that values at or near the boundaries can be ignored or counted purely by the arbitrary drawing of a division, when in reality there may not be that much difference between them.

On the plus side, however, is the fact that the DCADI score at least give an indication of how many ADI's need to be checked to see what the score is. If the DCADIV3 value is 5, the Red Flag must have been raised in 5 ADI's., if 2, only in two, which is useful to a limited extent.

Development Control ADI Version 3 (DCADIV3)

As a development of DCADIV2, DCADIV3 shares some of the disadvantages of its predecessor – namely the issue relating to how threshold layers are determined and implemented.

However, its major advantage is that it does provide a direct link back to the relevant ADI which triggered flags, albeit it a slightly unwieldy fashion. The ‘Binary’ string generated needs manual decoding to arrive at the right layer, and this is likely to become more confusing as more ADI layers are generated, but it is still an important improvement on the other DCADI’s.

However, as noted in the previous chapter the implementation did not work fully. As more layers were added, the multipliers grew to such an extent that errors in the numbers generated began to emerge, with for example the predicted ‘binary’ string being interrupted with other figures. It is not known if this is a limitation due to the software or the process, but it does mean that this version can only be used for a limited number of ADI’s. This could solved by having a two or more layers of this DCADI with a limited number of ADI’s represented in each, but this could easily become confusing, and defeats the aim of a one layer solution.

Another issue related to this layer is one of display. As the numbers became increasingly large, it became more difficult to map them in a way that made hot spots obvious. This was not an issue with the cut down version shown in the previous chapter, so it may be related to the multiplier issue.

Development Control ADI Version 4 (DCADIV4)

This version is the one that fulfils all the main criteria hoped to achieve in this project, though it is not without some disadvantages. It uses only one layer, it is possible to map this DCADI in a variety of ways, dependent on the choice of which ADI field to colour map on. Also, at any point it is also possible to see all the ADI values for any given 100 metre square, allowing the user to make their own interpretations of the data.

It is possible to recreate the other previous DCADI’s in this implementation, also, if needed. Queries could be run in the database to produce the DCADIV1, or V2 (it is unlikely to be required to create DCADIV3 if using this version, but is possible), with the data being stored in additional data fields, and these could also be used to colour map with. In the final Table, this has been done for ADIV1, though caution must be exercise as areas with no data have had the value of zero apparently assigned – those this may be due to faults in the calculations.

It is also possible to create other types of DCADI, two examples of which have been included which involve summing all the Medieval Period ADI’s (Figure 26), and another which sums all the Dress and Personal accessories ADI’s together (Figure 27). These are examples showing options which could be explored as possibly having more value in different DC contexts. For example, it has been shown that that certain types and periods of archaeological remains are not always found by some types of assessment or evaluation (Hey, G; Lacey, M 2001). By using combinations of ADI values to develop the DCADI mapping, it might be possible to try and counteract this problem to some extent. To be fair, it would be possible to create these other types of DCADI as raster layers, but these would still have the main problem associated with the raster solutions – the difficulty of linking back to the original ADI data and the inability to see all the different ADI data regardless of what combination of values were used to generate the colour mapping. The vector solution allows more flexibility in dealing with mapping as well as allowing those issues to be overcome.

However, the preparation of this type of ADI was more involved than the others, and involved various stages of exporting and importing between programs to achieve the final version (see appendix 5 for full details), and whilst it might be possible to simplify this to some extent, it is still more complex.

Another minor issue at the time of writing, is that the Join between the DCADIV4 Shape file and the 0DCADIV4.dbf has only been implemented as a temporary link within whichever ArcView Project the link is made. In other words, the data has not been permanently joined to the DCADIV4 shape file. However, this is possibly an advantage. The fact that the data is not actually within the Shape file means that when the 0DCADIV4.dbf is replaced by a newer version, the data is automatically updated – even to the extent of having additional fields added in. However, it does mean that if the DCADIV4 implementation needs to be used in other projects, the Join must be re-created. This is a relatively trivial process to carry out, though.

Overall, given the advantages and disadvantages of all the DCADI's demonstrated, the DCADIV4 is my preferred version.

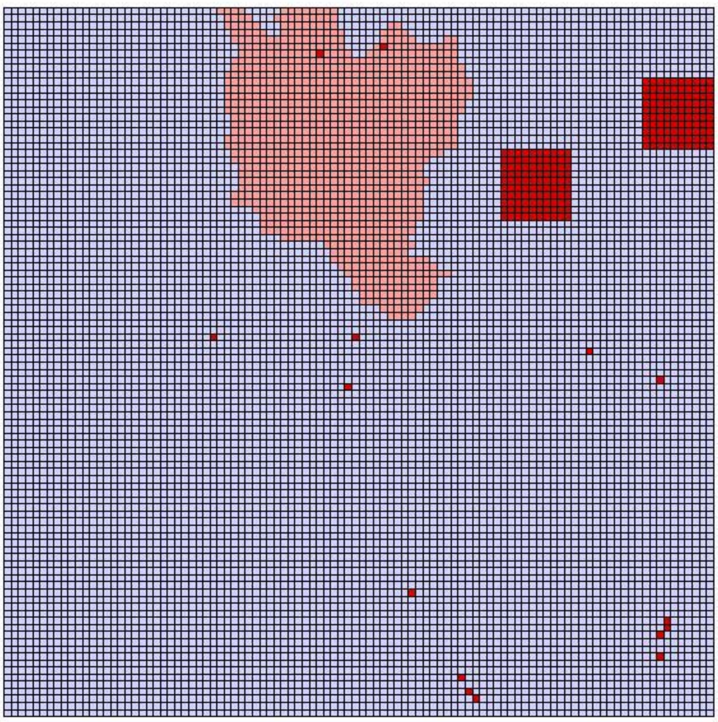


Figure 26. An example of the DCADIV4 mapping on a field which totals all the medieval ADI's.

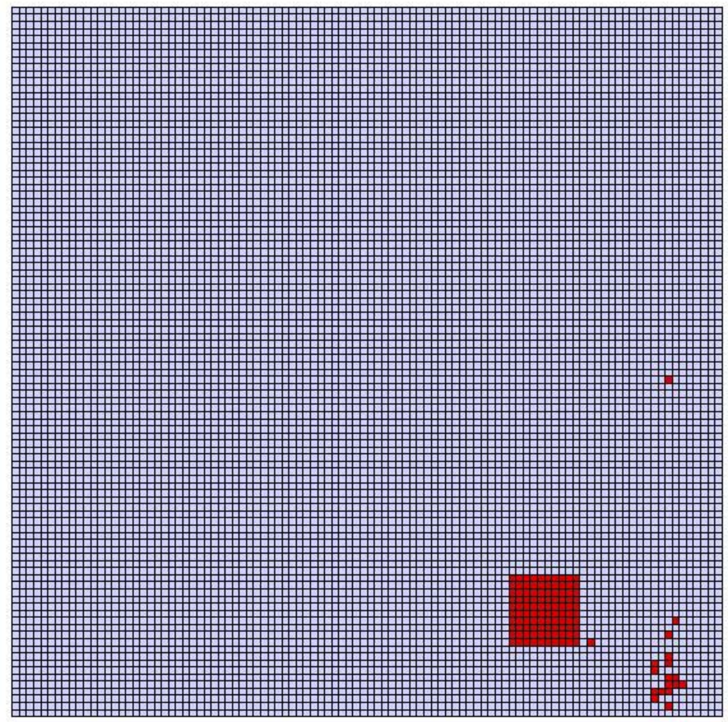


Figure 27. An example of the DCADIV4 mapping on a field which totals all the Dress and Personal accessories ADI's.

Chapter 6

Conclusions and The future of the ADI….?

The aim of this chapter is to look at the overall project and assess its success and possible improvements.

Overview

At the start of this project, the aim was to develop a system to allow imprecisely located stray finds data to be taken into account during Archaeological Development Control decision making. As far as I am aware, no work similar to this in finds mapping has been attempted before. This project has been successfully completed, as this aim has been achieved.

The key data used in the system - namely the finds type and period data - use nationally agreed Standards and terminologies for this data, which have been peer reviewed and thus show a certain level of robustness as categorising frameworks.

The actual mapping of the Finds data has also been fairly robustly done, and the finds have been shown to be able to be mapped usefully, and this Mapping merged to give an overall score for a particular area (the Artefact Density Index). It is hoped that the logic behind this process is seen as being sensible also, though it has had to be invented from scratch.

Finally, the data has then been transformed into a format which is aimed specifically to trying to make the data useful for development control purposes score (the Development Control Artefact Density Index) and a number of options regarding this have been explored and demonstrated.

What the model doesn’t address

Whilst the model has been successful, it doesn’t deal with certain issues which could benefit from attention. For example, the representativity of the data is not addressed – is the data skewed by fieldwork patterns? Is the data skewed by land use patterns? Is the data skewed by recovery method patterns? No attempt has been made to address these issues, partially because recovery method has not always been recorded for HER data. However, curatorial archaeologist often have to make decisions with other datasets solely on the data available without knowing the impact of these factors, so in that sense the ADI is no different. However, it is still an issue that would benefit from further research – both to determine any possible skewing of the data and also to develop methods to balance the data by further archaeological research.

The model also doesn’t really address the issue of how to deal with data which is ambiguous in other respects other than just location – such as quantity of finds. Other areas of ambiguity can be more readily dealt with, but this issue is likely to remain problematic particularly for antiquarian data. The only real remedy for this is more accurate recording of modern acquired data (e.g. PAS data) and an attempt to check sources of older Antiquarian data in HER's to fill in this information. Coupled with this is the notion of trying to persuade people to record new data's location more precisely, which would render the ADI process, for certain datasets, redundant. The better recording of recovery method would also allow better analysis of how this affects patterning of the data.

Is the ADI Useful?

I believe the DCADI can be useful, as it is based on the reasonably well accepted idea that stray finds can give an indication of the archaeological potential of an area. However, the usefulness of the ADI can only really be tested by use in a DC environment. To that end it is hoped to persuade colleagues at NYCC HER to use the data in the project area to see how it fairs as a tool. Even if this is not possible, it should be possible in a few years to test the ADI predictions against actual archaeological information recorded from fieldwork to see how well the predictions hold up. However, if the ADI does not appear to predict accurately archaeological potential we are left with the dilemma - is this due to limitations of the model, or the underlying assumption that stray finds can indicate potential? On this only further research will tell.

Can the ADI process be improved?

The actual development of the ADI involved various trial and errors, and it is likely that some of the solutions could be more elegantly produced, now that the solution is known. For example, not much time was spent in automating any aspect of the process of data entry into the vector tables, but if the ADI process were to be developed and promoted, user friendliness issues would obviously come to the fore, and the automation of parts of the model such as this would be desirable. Some more minor points of automation could also be introduced – for example rather than having to manually alter some parts of the data e.g. the Standard Deviation Threshold in the DCADI V2 and DCADI V3, this could be automated, the automation of the creation of the Grid versions of the various ADI's etc. Also, a user interface to control all the functions of the model would obviously be something to consider. For the DCADIV4, as noted, the issue of missing data means any automation would need prompts to ensure that care was taken by any user to ensure the correct data was being linked to the correct points.

One key aspect of the process that could be improved is the weighting of the data for parishes. In this model a generic weighting for all Parishes was used, which implies that all parishes are the same size. This is not the case, and the ability to weight the data based on the actual hectarage of each parish would be more desirable, and would result in, for example, smaller parishes having a higher weighting than larger ones.

Conclusion

The development of the Artefact Density Index and the Development Control ADI has been a long process, with some false starts and wrong turns along the way. However, the successful completion of the project shows that it is possible to develop a system to incorporate imprecisely located stray finds data into the Development Control process.

Perhaps this is one of the perils of trail blazing - the key thing is to try to find a route which gets you where you want to go, which may not necessarily get you there by the most efficient route, nor allow you to explore interesting things you uncover on the way. If I can stretch the analogy, then it might be that later travellers along an ADI type route might find easier routes to the same destination, or even a better destination. If so, I wish them luck!

Bibliography

Addison, C. (1999). Finds Recording and the changing role of SMRs. *SMR News* 8. NMR/RCHME. pp 7.

ALGAO/English Heritage SMR Working Party. (2002). *Historic Environment Records: Benchmarks for Good Practice.* London.

Altschul, J. H. (1990). Red flag models: The use of modelling in management contexts. In Allen, K. M. S., Green, S. W., and Zubrow, E. B. W. (eds.), (1990). *Interpreting Space: GIS and Archaeology.* Taylor & Francis: London and New York. pp 226 – 238.

Baker, D. (1999). *An Assessment of English Sites and Monuments Records.* Historic Environment Conservation: Bedford Report 97/20.

Bland, R. (1999). Portable Antiquities Initiative. *SMR News* 8. NMR/RCHME. pp 5 – 6.

Bourn, R. (1999). Events and Monuments: a discussion paper. *SMR News* 8. NMR/RCHME. pp 3 – 7.

Burrow I. (1985). The history of the SMR System. In Burrow I (Ed). (1985). *County Archaeological Records Progress and Potential.* ACAO. pp 6 – 15.

Catney, S. (1999). So what is an event? *SMR News* 7. NMR/RCHME. pp 3.

Chitty, G & Edwards, R. (2004). *Review of Portable Antiquities Scheme* Hawkshead Archaeology and Conservation Report.

Darvill, T; Fulton, A K. (1998). *MARS: The Monuments at Risk Survey of England 1995 Main Report.* School of Conservation Sciences: Bournemouth University.

Department of National Heritage. (1996a). *Protecting our Heritage A consultation document on the built heritage of England and Wales.* London.

Department of National Heritage. (1996b). *Portable antiquities: a discussion document*. http://www.britarch.ac.uk/cba/portant1.html Accessed 29/7/05.

dictionary.com http://dictionary.reference.com/ Accessed 01/11/04.

Dobinson, C. & Dennison, S. (1995). *Metal Detecting and Archaeology in England.* Council for British Archaeology and English Heritage. http://www.britarch.ac.uk/detecting/cont.html Accessed 29/7/05.

English Heritage (EH). (2000) *Power of place The future of the historic environment.* English Heritage: London.

EH & ALGAO. (2002). *Historic Environment Records: Benchmarks for Good Practice v 1.1*.

Exegesis Spatial Data Management. (June 1998). Sites and Monuments Database – Progress Report. *SMR News* 6. NMR/RCHME. pp 4.

Exegesis Spatial Data Management. (January 1999). Sites and Monuments Record database – recent developments. *SMR News* 7. NMR/RCHME. pp 10.

Fernie, K; Gilman, P (Eds). (2000). *Informing the future of the past: Guidelines for SMR's.* English Heritage: Swindon.

Foard, G. (1996). View from the Chair: The future of SMR Software. *SMR News*. RCHME. pp 2.

Foard, G. (1997). View from the Chair. *SMR News* 4. RCHME. pp 1.

Gilman, P. (2004). Sites and Monuments Records and Historic Environment Records in England: is Cinderella finally going to the Ball? *Internet Archaeology* Issue 15: Archaeological Informatics http://intarch.ac.uk/journal/issue15/gilman_toc.html Accessed 01/06/2005.

Griffiths, M. (1982). *Archaeological record System*. North Yorkshire County Council Computer Services.

Haselgrove, C. (1985). Inference form ploughsoil artefact samples. In Haselgrove, C; Millett, M; Smith, I (EDS). (1985). *Archaeology from the Ploughsoil: Studies in the collection and interpretation of field survey data* University of Sheffield. pp 7 – 29.

Hey, G; Lacey, M. (2001). *Evaluation Of Archaeological Decision-Making Processes And Sampling Strategies.* Planarch Project. Alden Group Limited: Oxford.

Kamermans, H; Fennema, K (Eds). 1996 *Interfacing the past Computer Applications and Quantitative methods in Archaeology CAA 95.* Analecta Praehistorica Leidensia 28. University of Leiden.

Laflin, S (Ed). (1986) *Computer applications in archaeology, 1986.* Centre for Computing and Computer Science: Birmingham.

Leech, R. (1986). Computerisation of the National Archaeological Record. In Laflin, S (Ed). (1986) *Computer applications in archaeology, 1986.* Centre for Computing and Computer Science: Birmingham. pp 29 - 37.

Lock, G R; Harris, T M. (1991). Integrating spatial information in computerised Sites and Monuments Records: meeting archaeological requirements in the 1990s. In Lockyear, K Rahtz, S (Eds). (1991). *Computer applications and quantitative methods in archaeology, 1990.* BAR International Series 565. Tempus Reparatum: Oxford. pp 165 – 173.

Lock, G; Harris, T M. (1992a). Towards a Regional GIS Site Information Retrieval System: The Oxfordshire Sites and Monuments Record (SMR) Prototype. In Larsen, C U (Ed). (1992) *Sites and Monuments. National Archaeological Records.* Copenhagen National Museum of Denmark. pp 185 – 199.

Lock, G; Harris, T. (1992b). Visualizing spatial data: the importance of Geographic Information Systems. In Reilly, D; Rahtz, S (Eds). (1992) *Archaeology and the information age – A global perspective*. One World Archaeology 21. Routledge: London. pp 81 – 96.

Lock, G; Bell, T; Lloyd, J. (1999). Towards a methodology for modelling surface survey data. The Sangro Valley Project. In Gillings, M; Mattingley, D; van Dalen, J (Eds). (1999). *Geographical Information Systems and landscape Archaeology*. The archaeology of Mediterranean landscapes 3. pp 55 – 63. Oxbow: Oxford.

Lockyear, K Rahtz, S (Ed). (1991). *CAA 90*. BAR International Series 565. Tempus Reparatum.

Miller, A P. (1996). Digging deep: GIS in the city. In Kamermans, H; Fennema, K (Eds). 1996 *Interfacing the past Computer Applications and Quantitative methods in Archaeology CAA 95.* Analecta Praehistorica Leidensia 28. University of Leiden. pp 369 – 376.

Murray, D. (1999). Ordnance Survey and the depiction of Antiquities on Maps: Past, present and future. The current and future role of the Royal Commissions as suppliers of Heritage Data to the Ordnance Survey. In Dingwall, L; Exon, S; Gaffney, V; Laflin, S; Van Leusen, M (Eds). (1999). *Archaeology in the age of the internet : CAA 97; computer applications and quantitative methods in archaeology*. BAR international series 750. Oxford: Archaeopress. Pp 235 – 239.

Museums Documentation Association. (1997). *Archaeological Objects Thesaurus.* http://www.mda.org.uk/archobj/archint.htm#Class . Accessed 01/11/04.

Newman, M. (2002). *SMR Content and Computing Survey 2002 v2.0*. EH NMR Data Services Unit.

Ordnance Survey. (1978) Archaeology Division Map Recording of Antiquities. Appendix 1. In Baker, D; Jeffrey, T; Hedges, J (Eds). (1978). *A guide to the establishment of sites and monuments records*. ACAO. pp 29 – 31.

Pett, D & Oxford Archdigital Team. (2003). *Portable Antiquities Scheme User manual for entering data onto database.* November 2003. Version 2.

Phillips, C W. (1980). *Archaeology in the Ordnance Survey 1791 – 1965*. The Council for British Archaeology. London.

Robinson, B. (2000). English Sites and Monuments Records – Information, Communication, and Technology. In Lock, G & Brown, K (Eds). (2000). *On the theory and practice of archaeological Computing*. Oxford University Committee for Archaeology Monograph Number 51. pp 89 – 106.

Royal Commission on Ancient and Historic Monuments of Scotland. (2005). *PastMAP.* http://jura.rcahms.gov.uk/PASTMAP/start.jsp Accessed 02/02/2005.

Royal Commission on the Historic Monuments of England (RCHME). (c 1986). *Thesaurus of archaeological terms.* Royal Commission on the Historical Monuments of England: Swindon.

RCHME. (1993). *Recording England's Past A Review of National and Local Sites and Monuments records in England*. Royal Commission on the Historical Monuments of England: London.

RCHME DSU. (1997) *SMR '97 A data Standard for England's Sites and Monuments Records.* Consultation Draft for Peer Review: Royal Commission on the Historical Monuments of England.

RCHME. (1998). *Archaeological Periods List* Version: 1.0 http://www.fish-forum.info/i_apl_e.htm Accessed 03/02/2005.

RCHME. (2003). *MIDAS: A Manual & Data Standard for Monument Inventories.* Royal Commission on the Historic Monuments of England: Swindon. 3rd Reprint.

RCHME & ACAO. (1993). *Recording England's Past A data standard for the Extended National Archaeological Record.* Royal Commission on the Historical Monuments of England: London.

RCHME & EH. (1995). *Thesaurus of Monument Types: A Standard for Use in Archaeological and Architectural Records.* Royal Commission on the Historical Monuments of England: Swindon.

Sargent, A. (2002). *Portable Antiquities Scheme report regarding the transfer of PAS data to SMR's*. English Heritage.

Smith, L. (1997). *NYCC SMR Data Audit.*

Wheatley, D; Gillings, M. (2002). *Spatial Technology and Archaeology The archaeological Applications of GIS*. Taylor & Francis: London.

Wikipedia 2005 http://en.wikipedia.org/wiki/Hectare Accessed 07/07/2005.

Appendix 1

Questionnaire

To determine the usefulness of the proposed project, a small questionnaire was distributed to a number of colleagues in HER's across the country.

The questionnaire was sent to the following people in August 2004:

Victoria Bryant Historic Environment Record Manager within the Worcestershire Historic Environment and Archaeology Service;

Louisa Matthews, Assistant Archaeologist (SMR), South Yorkshire Archaeology Service

Alice Cattermole Assistant Records Officer Norfolk Historic Environment Record

Barry Taylor SMR manager Greater London Archaeology Service

Sarah Poppy Historic Environment Record Officer Cambridgeshire County Council

Deborah Anderson, Assistant County Archaeologist, Durham County Council

Replies were received from:

Worcestershire Historic Environment and Archaeology Service;

South Yorkshire Archaeology Service;

Greater London Archaeological Advisory Service;

Norfolk Historic Environment Record (via phone conversation)

Cambridgeshire County Council

Questionnaire As sent August 2004

Are you a SMR or HER (self defined)?

Does your SMR/HER have a policy on how to record finds in the database and GIS, and what is it? (This could be an informal policy, but is basically what you are trying to get all your records like in the long term).

How do you map finds with poor locational information in your GIS, specifically:
If only the Parish area is known?
If only a generic Grid reference is known e.g. to a Quarter sheet or KM square?

Was your SMR based on the original OS Cards, or have these been subsequently added to the system?

How do poorly mapped finds get taken into account during Development Control decisions?

If a method for generating a map of archaeological potential for an area based on solely finds records, but which took account of poorly located finds, was available, do you think it would be useful to you for DC purposes?

Can you also let me have your official job title and place within the CC (e.g. we are Archaeology Team, Heritage Section, Planning & Countryside Unit, Environmental Services (!!)) so I can put it in the Acknowledgements

Reply From: Louisa Matthews, Assistant Archaeologist (SMR), South Yorkshire Archaeology Service November 2004

Are you a SMR or HER (self defined)?
SMR

Does your SMR/HER have a policy on how to record finds in the database and GIS, and what is it? (This could be an informal policy, but is basically what you are trying to get all your records like in the long term).
One was produced some years ago but does not appear to have been adhered to.

The general jist appears to have been to plot anything with a minimum of a 6 figure NGR and ignore the rest. I think it was inherited from the old paper-map system. An updated policy is expected to be finished in the next few months.

How do you map finds with poor locational information in your GIS, specifically:
If only the Parish area is known?
If only a generic Grid reference is known e.g. to a Quarter sheet or KM square?
If a km grid square is known the findspot is put on the SW corner where the grid lines meet. Anything less accurate than that is not given a point on the GIS and stored in a 'Parish Information' file until more details can be acquired.

Was your SMR based on the original OS Cards, or have these been subsequently added to the system?
Yes, but not exclusively. Other resources were used in setting up the SMR (the majority of which I have yet to identify...).

How do poorly mapped finds get taken into account during Development Control decisions?
Poorly mapped finds are used to indicate levels of archaeological potential. Type and date of finds are taken into consideration to help establish whether there are patterns of material showing up in a particular locale. This helps indicate the likelihood of a site in the vicinity.

If a method for generating a map of archaeological potential for an area based on solely finds records, but which took account of poorly located finds, was available, do you think it would be useful to you for DC purposes?
Potentially yes, but on it's own would not be enough to base decisions on.

Dinah had a few questions/comments of her own, I've copied what she sent me below:

Louisa - if by 'poorly mapped', Nick means finds without an accurate findspot, I would say that these help add to the picture of an area's archaeological potential, but no more than that.

Reply From: Victoria Bryant, Historic Environment Record Manager within the Worcestershire Historic Environment and Archaeology Service

Are you a SMR or HER (self defined)?
We are an HER (just because everyone else in this region is one)

Does your SMR/HER have a policy on how to record finds in the database and GIS, and what is it? (This could be an informal policy, but is basically what you are trying to get all your records like in the long term).
Our new system allows for the creation of finds and environmental indices (see documents attached to e-mail)

How do you map finds with poor locational information in your GIS, specifically:
If only the Parish area is known? By a parish polygon attached to an activity record
If only a generic Grid reference is known e.g. to a Quarter sheet or KM square?
By the grid reference and accuracy assessment within the record itself

Was your SMR based on the original OS Cards, or have these been subsequently added to the system?
It was based on County museum cards which were based on OS cards plus additional info

How do poorly mapped finds get taken into account during Development Control decisions?
They can be mentioned as indicating activity of a particular period in the general area.

If a method for generating a map of archaeological potential for an area based on solely finds records, but which took account of poorly located finds, was available, do you think it would be useful to you for DC purposes?
No - it might be useful if used with other evidence but would need a very great deal of synthesis with other records

Barry Taylor SMR manager Greater London Archaeology Service October 2004

Does your SMR/HER have a policy on how to record finds in the database and GIS, and what is it? (This could be an informal policy, but is basically what you are trying to get all your records like in the long term).
Yes. We only record finds where they are the only evidence for a particular period or activity type. So for example if you have a Roman ditch with residual iron age pottery in it and there is no other evidence for iron age activity then we would create two records:
1: Ditch Roman
2: Findspot Iron Age Associated Finds Pottery Iron Age

We do this for two reasons. First, because we break sites down into separate monuments relating to distinct periods of activity so the Iron age activity on the site would be considered a different phase. And second because if we attach iron age finds top a roman monument they wont come up in a search of iron age sites unless you search specifically for finds.

Finally if the excavator thought that the iron age pot wasn't residual but that the ditch was in use from the iron age to the roman (i.e. the finds were in situ) then we wouldn't record the pottery as a find but would just create one monument
1: Ditch Iron age to Roman

Interesting finds are recorded in the description.

The older SMR records don't always follow this rule (or in fact any recognisably consistent rule) such as the burial where the grave and the grave goods get separate records!

Our policy sort of breaks down when it comes to groovy finds like coin hoards which we would always record but then it could be argued that the coin hoard itself is a separate and distinct phase of activity and therefore deserves a record of its own.

How do you map finds with poor locational information in your GIS, specifically:
If only the Parish area is known?
If only a generic Grid reference is known e.g. to a Quarter sheet or KM square?
We note the level of accuracy of the grid coordinate

Was your SMR based on the original OS Cards, or have these been subsequently added to the system?
Yes I think so

How do poorly mapped finds get taken into account during Development Control decisions?

I'll ask my DC colleagues but they are probably taken into account as adding to the potential of the site. It probably depends on the finds as well and how many there are. One or two Roman coins from a general area might be background noise but a load more would possibly mean a site.
Similarly one axon brooch is nothing, 5 Saxon objects from an area probably means a cemetery. Its more difficult when you get Palaeolithic finds.

If a method for generating a map of archaeological potential for an area based on solely finds records, but which took account of poorly located finds, was available, do you think it would be useful to you for DC purposes?

Yes definitely.

Reply From: Sarah Poppy Historic Environment Record Officer
Cambridgeshire County Council August 2005

Are you a SMR or HER (self defined)?
HER, self defined

Does your SMR/HER have a policy on how to record finds in the database and GIS, and what is it? (This could be an informal policy, but is basically what you are trying to get all your records like in the long term)
Yes, an informal one. I will record and index everything. I only index unique find type and a quantity (where known), not individual occurrences of the same kind of find. I record only minimal information about finds, e.g. type and material and date.

How do you map finds with poor locational information in your GIS, specifically If only the Parish area is known If only a generic Grid reference is known e.g. to a Quarter sheet or KM square?
Many finds were originally mapped according to a 4 figure grid reference e.g. TL3557 (i.e. at the origin of the 1 km square). Some were even mapped at the origin of a 10km square, for which I have removed the mapped object. I am now only indexing to parish level where this is all we have, with no mapped objects. Those finds that are mapped to 1 km square I am leaving if the description justifies this level of precision. Often the description only refers to parish level, so the mapped object is removed, and the find is only indexed at parish level.

Was your SMR based on the original OS Cards, or have these been subsequently added to the system?
Based on the original cards

How do poorly mapped finds get taken into account during Development Control decisions?
Just spoken to our DC bod. Poorly located finds are used sparingly and cautiously in DC advice, but is sometimes used when they support more substantive and accurately located information.

If a method for generating a map of archaeological potential for an area based on solely finds records, but which took account of poorly located finds, was available, do you think it would be useful to you for DC purposes?
Yes, we would certainly be interested to hear about this, and could certainly be of use.

Reply From: Alice Cattermole August 2004
Notes of Phone conversation

Note this conversation took place before the questionnaire above had been fully worked out.

System Used = HBSMR with approximately 40000 records

Mapping
Event of Metal Detecting find spot of field – as a polygon

Will record individual findspots if got NGR’s

Antiquarian – record finds but not mapped, do GIS search then additional parish search

Not map to Quarter sheets, 4 figure grid ref minimum

Appendix 2

Relevant Fields within the PAS and HBSMR databases

Definitions given here are as I understand them from correspondence with the PAS and Exegesis, and may not be completely accurate.

List of data recorded in HBSMR for finds

Finds Module within HBSMR
Find Type – uses the Archaeological Objects Thesaurus
Material – uses the EH materials thesaurus
Period – using terms from the RCHME Archaeological Periods List
Quantity – each record can record more than one item e.g. a number of sherds.
Sci Dating – records if scientific dating of the artefact has taken place.
Summary
Detailed Description
Links to Monuments and Events modules
Museum/store information - where located (organisation and within organisation) and ref number.

Monuments Module
Record Type as Finds spot
Monument Type as Findspot
Through the link to this module it is possible to ascertain locational data such as the National Grid Reference; parish, District etc of the find(s).

List of data recorded in PAS Database for finds

This is only the data received from the PAS – it is understood more fields are recorded in the actual database (Pett, D & Oxford Archdigital Team 2003).

old_findID
culture – to record cultural type of find, e.g. Anglo Saxon, Anglo-Scandinavian.
discmethod – cross reference to a Discovery Method Look Up Table.
description
objecttype – uses EH/RCHME Thesaurus of Monument Types
class – gives extra detail to object type.
objdate1subperiod – qualifier to objdateperiod 1 e.g. late, early etc.
objdate1period – cross reference to Period Terms Look Up Table, From Date.
objdate2subperiod – qualifier to objdateperiod2 e.g. late, early etc.
objdate2period – cross reference to Period Terms Look Up Table, To Date.
broadperiod
numdate1 – Calendar date if known (From Date).
numdate2 – Calendar date if known (To Date).

material1 – cross references to a Material Type Look Up Table.
length
width
diameter
thickness
quantity
gridref
easting
northing
knownas
parish
district
county

Appendix 3

Codes used in naming aspects of the ADI's

MDA Class	Code for ADI
Agriculture and Subsistence	AS
Animal Equipment	AE
Architecture	AR
Armour and Weapons	AW
Container	CO
Currency	CU
Dress and Personal accessories	DP
Ecofacts	EF
Furnishings and Furniture	FF
Manufacture and Processing	MA
Medicine and Pharmacy	MP
Measurement	MM
Signs or Symbols	SS
Sports and Games	SG
Tools and Equipment	TE
Unassigned	UA
Written Communication	WC

Period	Code for ADI
Prehistoric	PH
Roman	RO
Early Medieval	EM
Medieval	ME
Post Medieval	PM

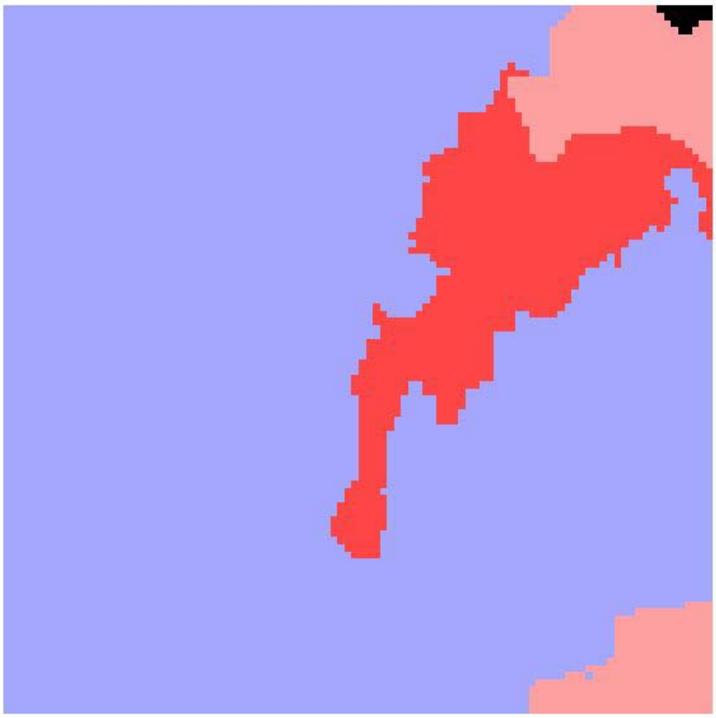
Precision Level	Code for ADI
Quarter Sheet	5k
Parish	Par
Kilometre Square	1k
100 metre or more precise location	100

Appendix 4

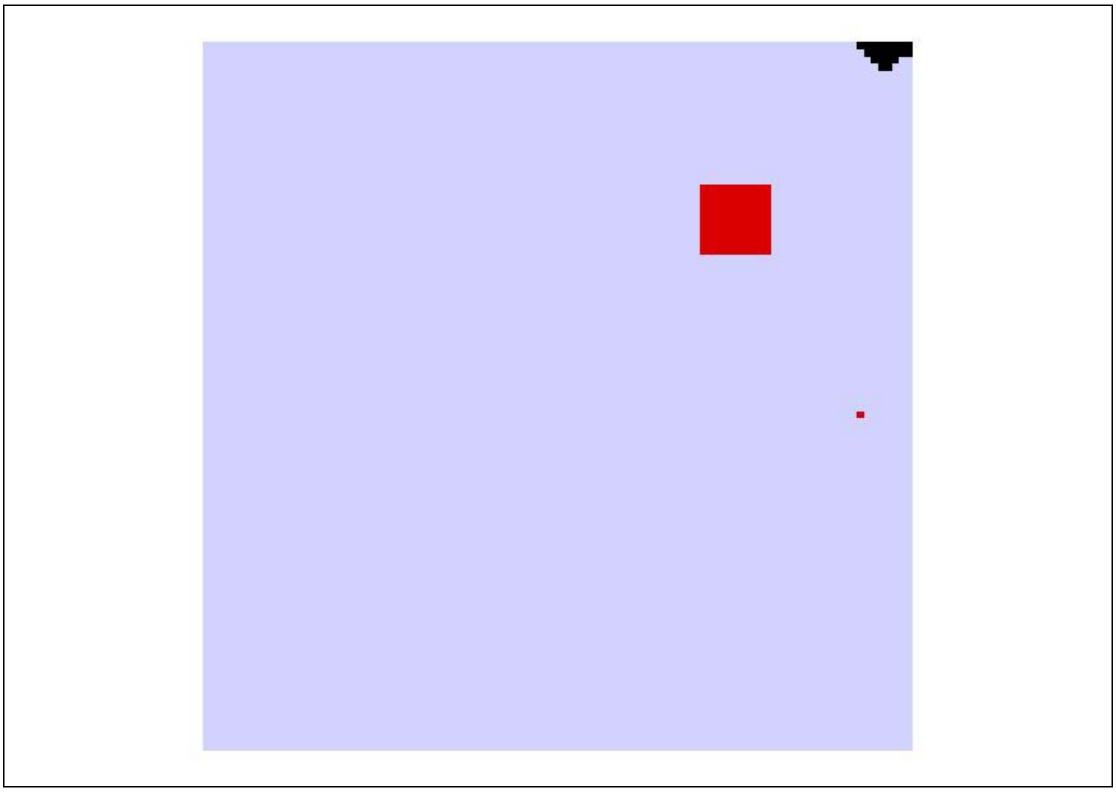
Gallery of ADI's created

This appendix provides images of all the ADI's created for this project.

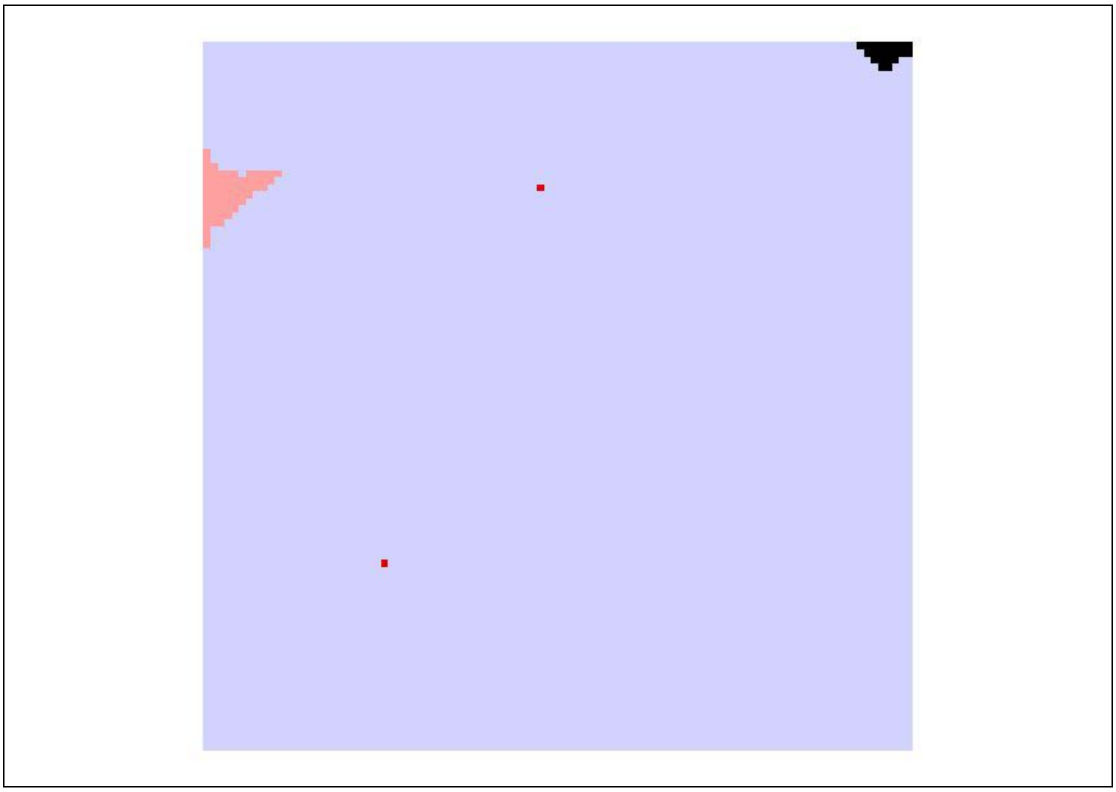
AeMe (Animal Equipment Medieval).



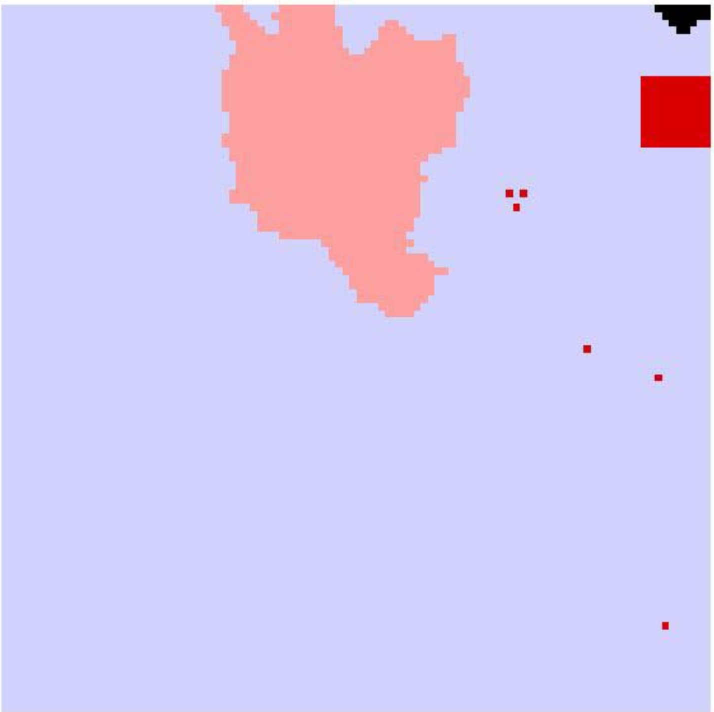
AwMe (Armour and Weapons Medieval).



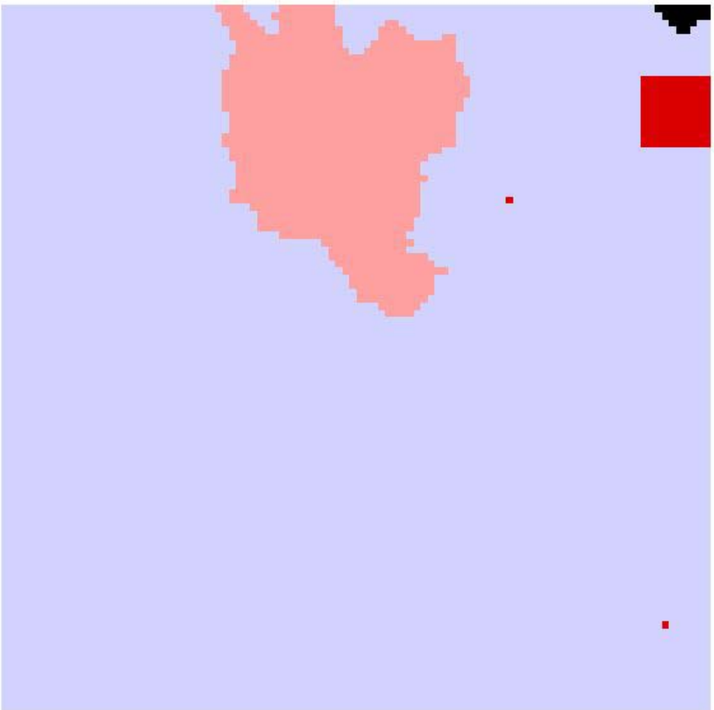
AwPh (Armour and Weapons Prehistoric.



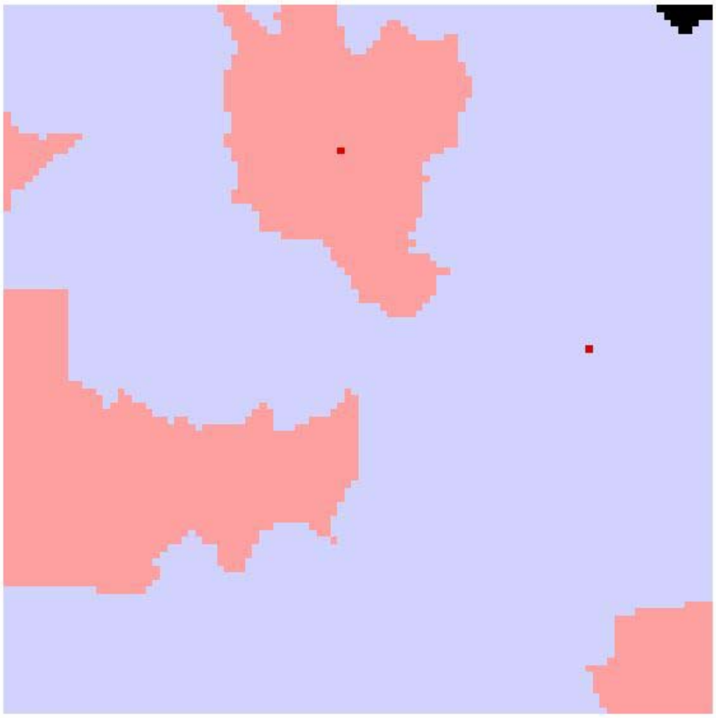
CuMe (Currency Medieval).



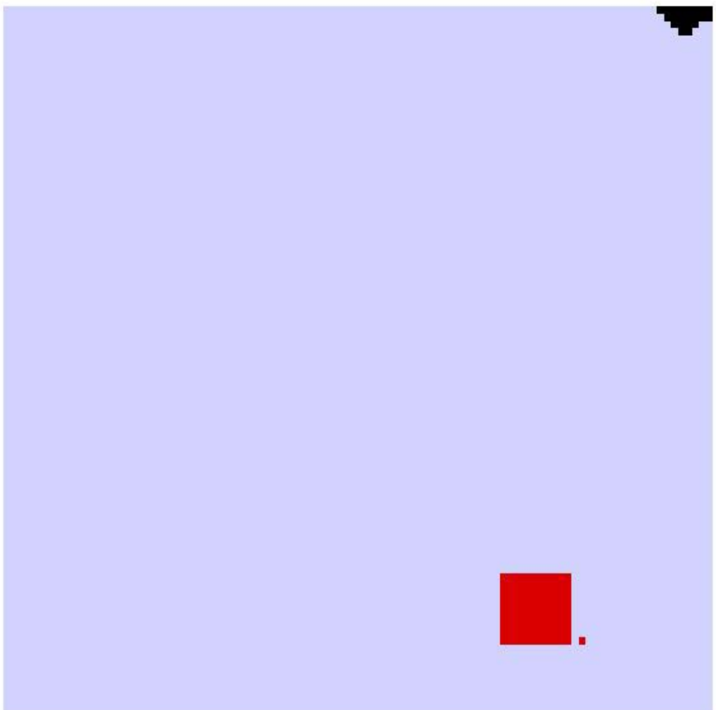
CuPm (Currency Post Medieval).



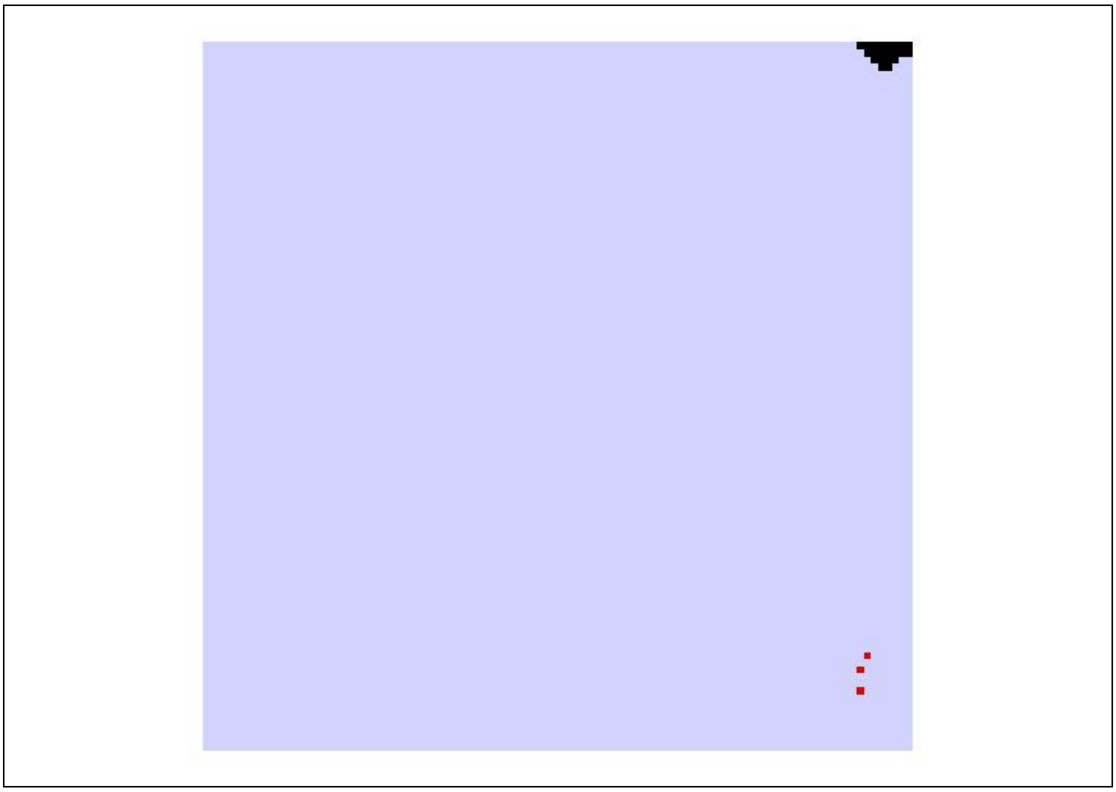
CuRo (Currency Roman).



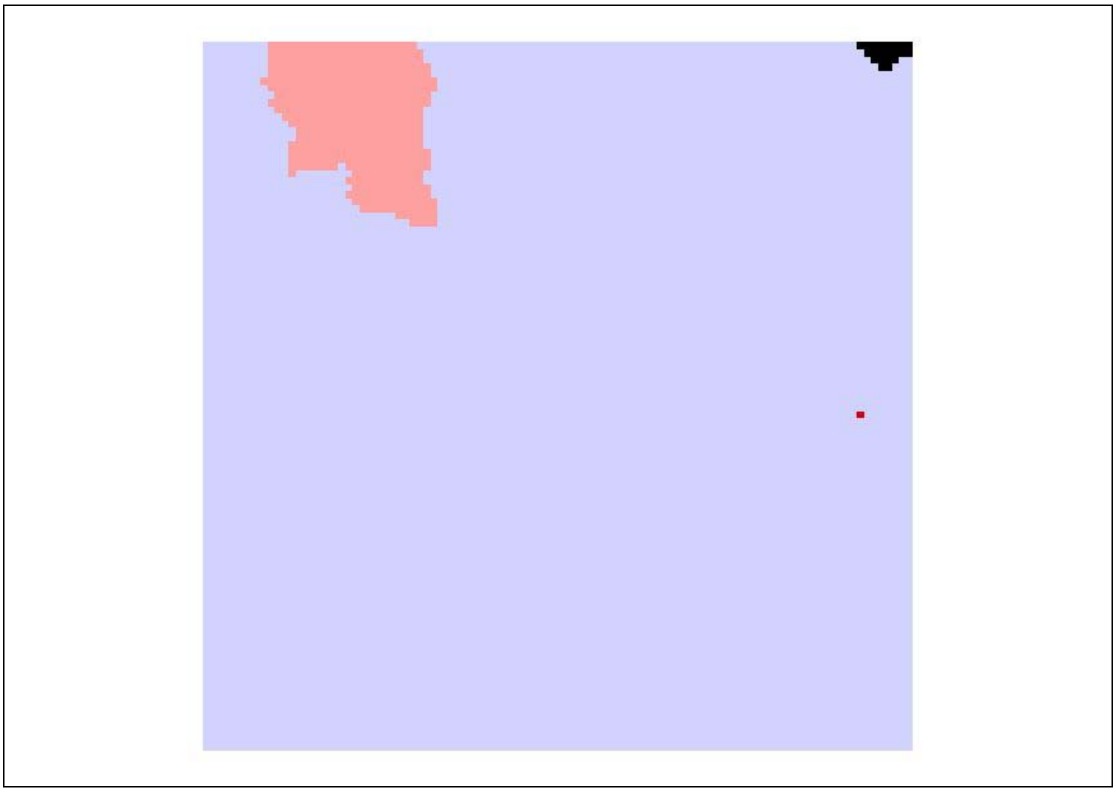
DpEm (Dress and Personal accessories Early Medieval).



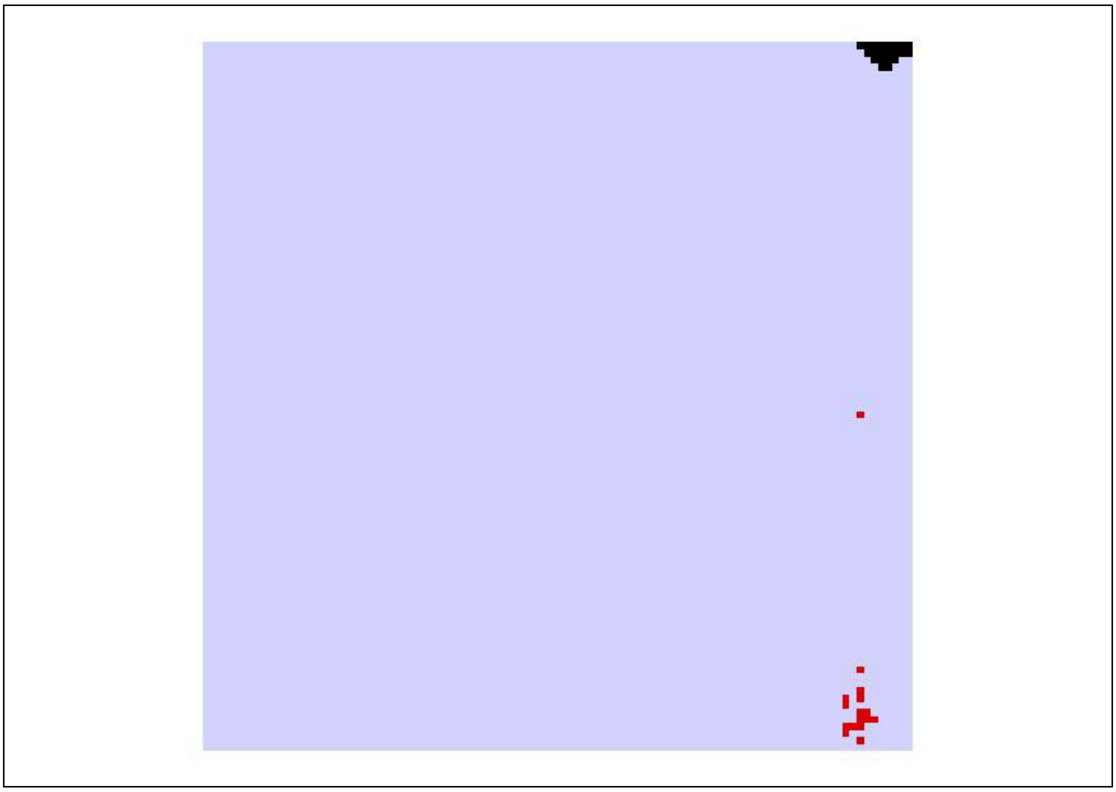
DpMe (Dress and Personal accessories Medieval).



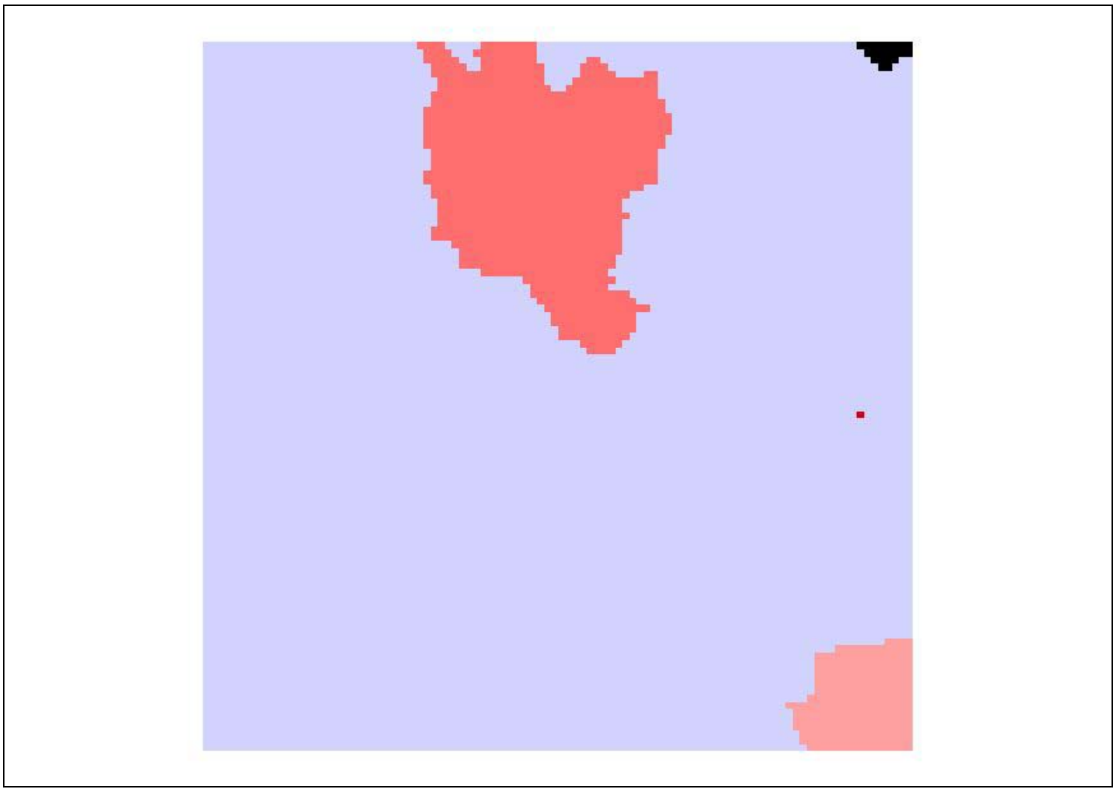
DpPm (Dress and Personal accessories Post Medieval).



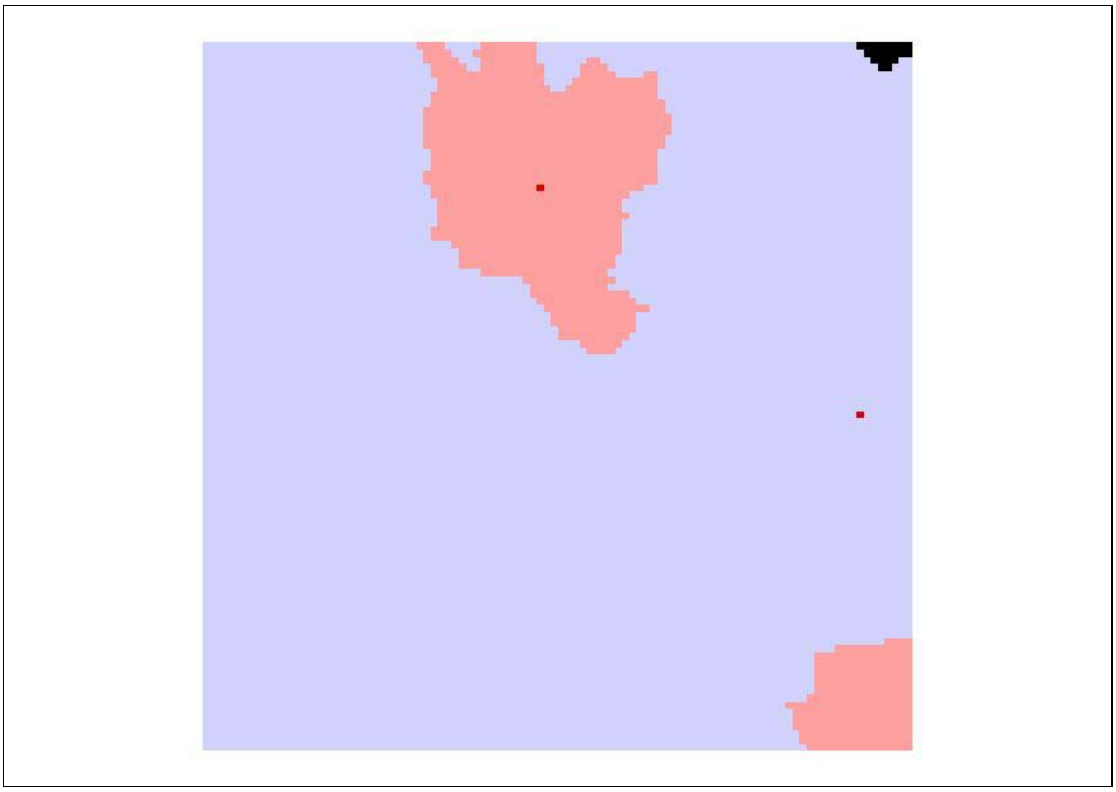
DpRo (Dress and Personal accessories Roman).



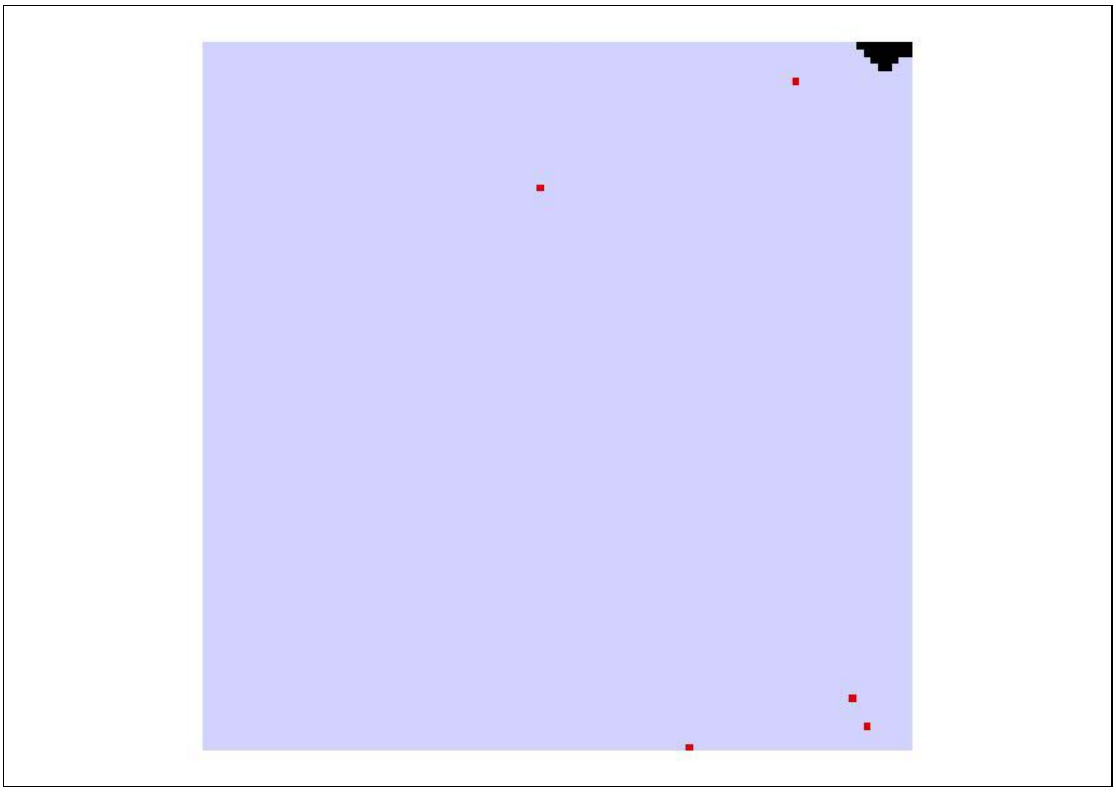
MmMe (Measurement Medieval).



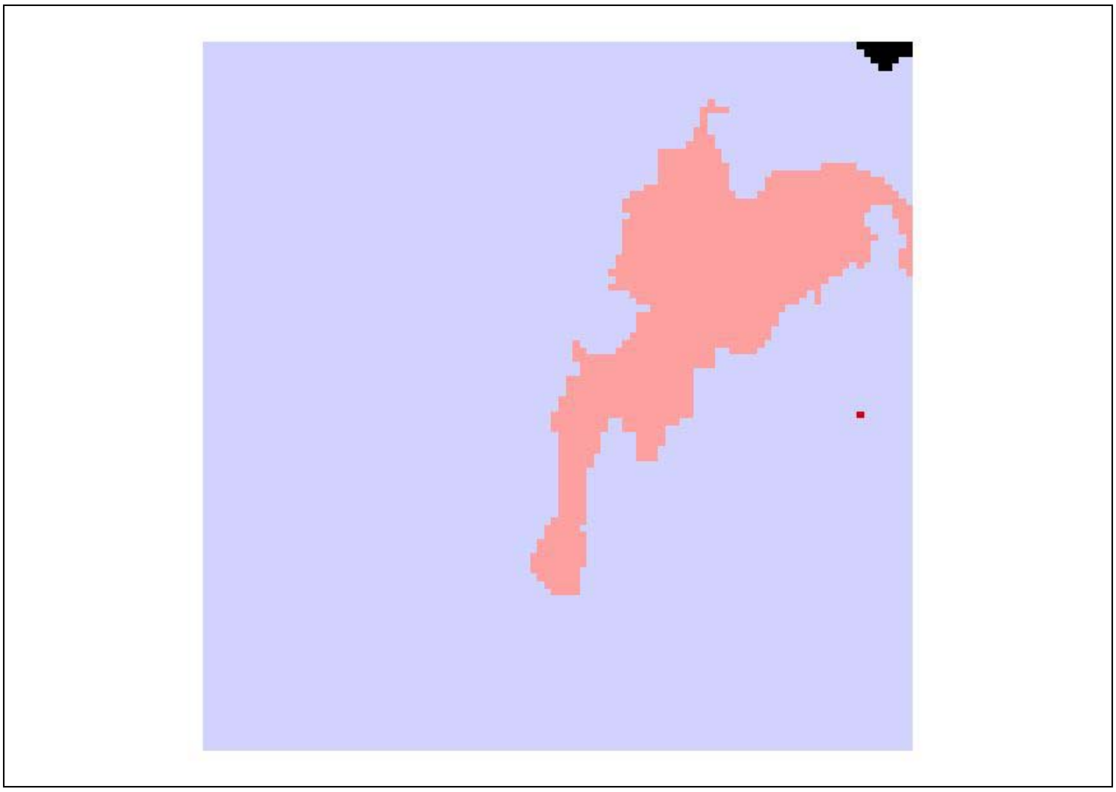
MmPm (Measurement Post Medieval).



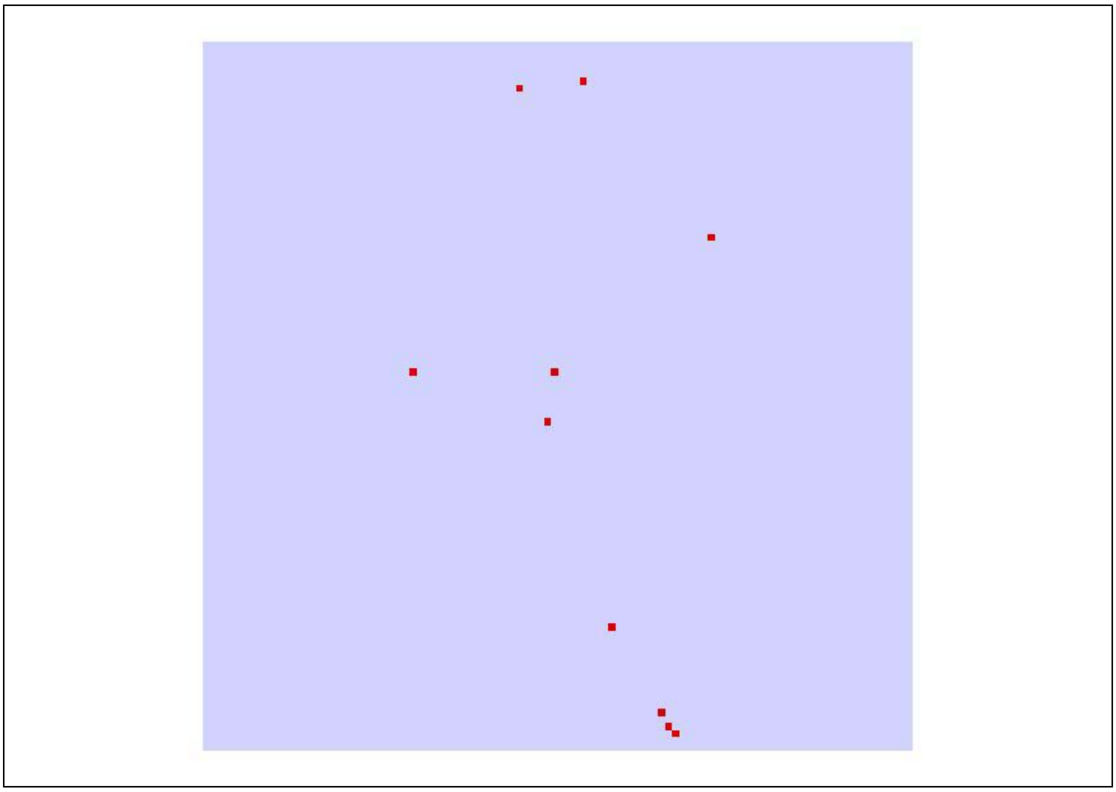
TePh (Tools and Equipment Prehistoric).



TePm (Tools and Equipment Post Medieval).



UaMe (Unassigned Medieval).



Appendix 5

Instructions for recreating the Project

The main text of the dissertation has given information about the overview of the project. This appendix is aimed at detailing how to recreate the ArcView Project from the data on the CD, and how to run the various models.

With this information, and the data provided, it should be possible to re-create each separate phase of the Project independently, or to work through the whole project and recreate the whole thing.

Project Data

The various models were created to run from, and store data in, a directory c:\00data, so the first step would be to copy the Folder 00Data from the CD to the C Drive.

The key data for the Project is within the folder zProjectData and consists of the shape files 100mGrd; proj1k; proj5k, and projParish. These files contain the raw data which was manipulated to create the various models.

Also needed are the Extensions Spatial Analyst, Model Builder, Database Access and Geoprocessing to be loaded.

Running the ADI Models

The Models have combined the Vector Conversion and ADI creation Phases into one model.

To run the Model it first needs to be opened in Model Builder. The Models are stored in the Folder 00Data\ADIData – with each model being stored in the appropriate sub folders. For example the AeMe model lives in the folder 00Data\ADIData\AeMe\AeMeADI, and so on.

Once the model is opened it can be run form Model builder and should create the appropriate data.

Once the ADI has been created it can be saved in the format <ModelName>ADIGrd within the 00Data\ADIData folder, where Model name is AeMe or equivalent.

These created Grids have also been provided to enable the later stages to be re-created without needing to go through the above stage.

Creating the Development Control ADI

The different DCADI's have different methods of creation, but they share some features. The models and data for the supplied data are all stored in the folder 00Data\DCADIData\DCADIV# for the appropriate numbered DCADI.

These processes also need the Extensions Spatial Analyst, Model Builder, Database Access and Geoprocessing to be loaded.

To run the first three ADI models, all the ADI Grid themes need to be added as Grid Data Sources to a Project, found in the folder 00Data\ADIData. Once this has been done, for the first three DCADI's it is merely a matter of running the appropriate model to generate the DCADI.

Development Control ADI Version 1 (DCADIV1)

This data is best displayed using the Standard deviation Classification in the Legend Editor Dialog box, as the default display makes it appear as if there areas with no data.

Development Control ADI Version 2 (DCADIV2)

Experimentation has shown that the optimal way to display this layer appears to be using the Equal Interval classification, with the number of classes set to one more than the maximum value. E.g. in the example provided the maximum value is 7, so the Number of classes has been set to 8.

Development Control ADI Version 3 (DCADIV3)

There were some problems with this DCADI, as detailed in the main text. However, a cut down model version 3a is included to show how the process should work, and it is possible to see the errors created if version 3 is run.

This data is best displayed using the Standard deviation Classification in the Legend Editor Dialog box.

Development Control ADI Version 4 (DCADIV4)

To produce this DCADI, an additional extension Grid Machine Version 6.53 is needed. This is needed to convert the raster grid into Point themes.

The Point themes have been provided and are stored in the folder 00Data\pt.

Each of these themes then needs to be exported into an Access database. This was done by exporting the Attribute table of the Shape files as Delimited Text files (also provided in 00Data\Databases\ADITxt).

These files were then imported in an Access database (provided at 00Data\Databases) and linked into a single table (1DCADIV4Original – does not include AeMe data). However, it was noted on re-importing this data into

ArcView that the data did not appear as expected, as the area of no data in the top right corner had vanished, and data was missing from the bottom row of the data. It was realised that the fact there was missing data needed to be accounted for. The missing grids were identified (see Gap.txt and Gap Table in database).

The data was then exported into Excel and the correct Gaps placed in the data. This Excel data was re-imported into Access (as table 2DCADIV4Excell) and then the UaMe theme was added into the data (as table 3DCADIV4PreEx2). This data was added separately as it contained no missing grids.

This table was re-exported to Excel (see 0DCADIV4Export2) to allow the creation of the extra fields i.e. DCADIV1; Medieval, DP. (Note – this is probably achievable within Access, but was done in Excel as it is simpler). This data was finally re-imported into Access as Table 0DCADIV4.

This table was exported from Access as Dbase V file called 0DCADIV4.dbf. At this point, a copy of the Vector 100mGrd Shape file was made, and renamed DCADIV4. This file was edited in ArcView to remove all the fields except the ID field. Both these files are provided in 00Data\DCADIData\DCADIV4.

Finally, it was possible to bring the 0DCADIV4.dbf table into ArcView and link it to the Shape File DCADIV4, using the Join Command. To do this the attribute table of the Theme DCADIV4.shp needed to be opened, and the 0DCADIV4.DBF file added to the tables in the project. These could then be joined using the ID field as the linking field.

Appendix 6

Dissertation Project Diary

January 2004

I had planned to meet Julian Richards and discuss possible dissertation topics. However, on the day of the meeting, my fiancé was taken into hospital.

February 2004

I have managed to meet Julian and agree on a possible Dissertation topic, the incorporation of digital excavation archive data from the West Heslerton Project into the north Yorkshire County Council Historic Environment Record. The aim is to try and prepare this project as an assessed lecture in May as a way of determining if the project is suitable. A possible alternative topic was also briefly mentioned, looking at the mapping of finds across North Yorkshire.

Nettie, my Fiancé, has been taken seriously ill, and has been in hospital for three weeks, so in between work, keeping up with lectures and visiting/looking after her have not had much time to start on my dissertation.

March 2004

I have met Dominic Powlesland, head of the Landscape Research Centre, and director of the West Heslerton Project, to discuss my proposal. He seems happy enough to support the project and has given me some data on DVD to look over to get an idea of the background of the project, and an idea of the data structure of the archive.

April/May/June 2004

I have been catching up with coursework, and had to postpone the assessed lecture due to Nettie's illness and the aftermath. I have spent these months finishing off the taught part of my course, finishing formative assessments and completing my final Summative assessment.

July 2004

I am now well behind where I hoped to be, and as I had planned on things working differently, am still trying to catch up, as well as fit in various holidays I had planned on the basis of having finished my course work. Basically, I need to have gone through the process of determining if my project is feasible by September at the latest, so I can crack on with it over the Autumn/winter.

Whilst getting ready to commit to the West Heslerton Import Project, have done some thinking and decided to change tack. There are a number of reasons for this, the main one being that it is not certain that this project will be feasible from an academic viewpoint, and I can't afford to waste time on it if that is the case. Also, I may be able to persuade my employers that is worth doing as a work project anyway, which means I'll still be able to do a version of it.

August 2004

Have discussed the possibility of changing my dissertation with Julian, and he says it is ok, and appears interested in the alternative one proposed – trying to incorporate imprecisely located stray finds data into the DC process. We met on 18 August to discuss possible options and came up with a rough timescale for the dissertation to proceed.

I have started the feasibility study – basically contacting Steve Dobson to ascertain if the project is doable in ArcView. He seems to think it is, the only real question is whether the extension I need will be available for me to use at home, or if I will need to come into York to work on these aspects.

I have been having major PC hassle for a while, lock ups etc, so have also been trying to sort these out. Have upgraded to XP and this seems to have solved the problems.

September 2004

I have contacted the Portable Antiquities Scheme to see about obtaining data from them and this is progressing – I have signed the licence and am now just waiting for the data.

In the meantime have been roughly planning the Dissertation with Microsoft Project. I have also started trying to think of which area of North Yorkshire to focus on, have been asking work colleagues if they have any ideas, and the Vale of Pickering seems to be emerging as a likely contender.

The XP fix didn't last, and my PC has still been playing up, more drastic remedies are required. Went to visit Stewart Waller (colleague form the AIS course now working at the ADS) and he basically rebuilt my PC (with some token help from myself) with a new (to me) motherboard and processor, based on the old PC he was using for his server (but his old PC was still higher spec than my hardware).

October 2004

I am writing this about three weeks since my upgrade, and no problems, so hopefully I now have what I need to complete my dissertation, a stable PC I can rely on not to crash or freeze up.

Most of the work this early this Month has been focussed on what I need to prepare for my next meeting with Julian.

I have started doing some preliminary background reading for the academic side – this has proved quite fruitful in terms of ideas already as well as academic angles to go on.

I have also started on approaching some other HER's to see how they deal with finds so I can compare and contrast – so far Greater London, Cambridgeshire, South Yorkshire, and Norfolk have agreed. Norfolk in fact phoned me up to answer my questions, which was a bit surprising, as I hadn't decided on them exactly! Still, I got some useful info from the chat which helped me frame my ideas more, and their HERO, Alice Cattermole, is happy for me to go back to her for clarification if needed.

I have been thinking about a project area, and whilst I know the size of area I want to do, haven't picked the exact area yet, as I want to get the PAS data before I make a final decision. However I have determined some criteria for determining it, and with this in mind, have been looking at the data for North Yorkshire. This has also involved me bringing the data home and importing it from MapInfo into ArcView, and determining that this works, and familiarising myself with ArcView. It's ok, but difficult as I am so used to MapInfo, and find aspects of it less user friendly. It is only 3.3, so some of these things (I would hope!) would have been addressed in ArcView 8. For example, the whole importing procedure of Mif's was long winded compared to the ease with which you can import Shape files into MapInfo.

I met with Julian on 12th October and agreed how to proceed, so have been working on my first two chapters. I have also got the PAS data now and have begun looking at this, and getting to grips with it.

November 2004

Since meeting Julian I have been doing a lot of background reading regarding my first two chapters, but also in preparation for later phases. I have also managed to get most of the feedback from different HER's; have determined my area, based on various criteria, and have written my first two chapters. I have begun some preliminary work on the next chapter, as well as continuing to get more familiar with ARCVIEW, in particular the Spatial Analyst and Model Builder extensions which I will need for this project.

I have arranged to meet Julian in early December, so am focussing on what needs to be ready for this meeting, as well as more general progress.

December 2004

Met with Julian on 13th December 2004 to discuss the first two chapters. Following discussion, it is obvious that they need extensive re-writing to focus them more tightly. We also discussed the model I am working on and what is involved, and these discussions were useful in helping me decide how to actually implement the model.

January 2005

Over Christmas and the New Year have been doing some more background reading, and working on re-writing my first two chapters, as well as working on the third chapter. I have arranged to meet William Kilbride in February (as Julian is on Sabbatical) to discuss these chapters. Work has begun on the preparation of the data for the area chosen, involving checking to see if there is additional data which could digitised in the NYCC system, and preparatory work on the design of the model in ArcView was also undertaken.

February/March/April 2005

Only limited work carried out on the dissertation for most of this period, as I was organising my wedding for 2nd April.

The main work carried out prior to my wedding was to ensure that all the relevant data was input to the NYCC system and exporting this data. Work was begun in mid/late April in preparing this data and inputting the NYCC and PAS data into ArcView, in preparation for the Assessed lecture. A subset of the data (coins) was chosen to input and run through the model, to ensure there was something to demo at the lecture. It is hoped that there will be time to do the same for at least one other dataset (pottery or flint) in time for the assessed lecture. Work on the assessed lecture (mainly in determining format – i.e. what needs to be included, and gathering images etc) was also begun in late April.

May 2005

I have taken the first week of May off work to prepare some data and work on the general model, in preparation for the Assessed lecture.

The API (Archaeological Potential Index) has been renamed the ADI (Artefact Density Index), following comments from correspondents - this change was decided on a while ago, but it took time to come up with the right sort of name.

The entire ADI Process was carried out for one data set (coins data) and ADI's generated to illustrate the lecture.

The assessed was lecture was given on 16th May and received quite well.

Following this some work on other data preparation was carried out (for lithics and pottery).

June 2005

Following comments at the lecture, and discussion of issues with Julian on 1st June, I have decided to take a different tack with the data. The issue of classifying the Finds by materials (which lead to problems with metal objects), had lead to a proposal to classify metal finds by the Class type in the MDA Thesaurus. This approach has been taken for all finds now.

This means all the data needs to be recast and the model recreated, and time has been lost working on this, but the exploration of different methodologies is all good stuff to go into the dissertation text.

I have taken a week off in June to try and finish the majority of the model work, but there are still some issues to resolve - mainly how to get ADI's into one layer format for ease of DC use. I have discussed some options with Julian, which are proving ok. I am having trouble contacting Steve Dobson to discuss this though.

Meet with Julian to discuss work on June 24th and he seems pleased with progress. It is agreed that rather than demo one final display DCADI, a number of options (at least 2, probably a third, and potentially maybe a fourth) will be included.

I manage to make contact with Steve who has been doing the field school, and have an email discussion with him, but arrange a face to face meeting for early July to discuss the issues/details in more detail.

July 2005

I have a deadline to try and write up as much as possible for 15th July, so I can meet Julian to discuss it on the 20th of July. This has proceeded ok, though there is still some project work to do (some more data preparation, and the finalising of all the ADI's and DCADI's) I have done enough to do some rough drafts of most of the dissertation.

I meet Julian on the 20th July to discuss progress so far and what needs to happen next. He seems happy with how I am getting on with the project and the dissertation, which is re-assuring.

I also manage to meet Steve Dobson for a chat on 20th July to see about trying to export the raster ADI's into a vector format. He suggests a method which involves using ARCGIS, so I will need to travel to York to do this. His suggested method involves using 3d Analyst to link a Z value, provided by the ADI, to a point, and then using this point data to link to the appropriate vector square.

August 2005

The week beginning the 8th of August I have taken off work to finish my dissertation. The weekend before I spend finishing preparing the data as I need to go to York to export the data into a vector format, and want to do that all in one day in possible.

On the Monday, I eventually manage to get the data out of the Raster ADI and work out and trial successfully a method to get this exported data into a vector layer. The method Steve Dobson suggested didn't work – for some reason the values didn't transfer but couldn't figure out why. Luckily, Ben Gourley had an extension on his laptop which allows the export of raster to point data. Thanks Ben and Phew!

The Tuesday is spent finishing the project off at home, and the rest of the week is spent finishing writing the dissertation, checking for typos and fine tuning.

This final (?) draft will be handed in to Julian on 15^{th} August to discuss on the 20^{th} August. Hopefully not much will need to be done then except printing and binding!

Appendix 7

Project Planning Information

To fully compile a hard copy version of this document, Microsoft Project Document DissertationPlan.mpp (created in Microsoft Office Project Professional 2003) should be printed out as formatted and bound into the Dissertation after this page.