MEK 10/11/06

NIAN GIS - Technical summary (in progress)

Base GIS

The GIS work for this aspect of the National Ice Age Network is being carried out using ESRI's ArcGIS 9.0. The Ordnance Survey mapping for each county was initially downloaded from Edina Digimap and five base GIS projects set up. These preliminary GIS projects included the 1:50k raster maps, the 1:50k panorama contour data and the 1:50k Meridian vector data. The rivers within the counties were isolated from within the Meridian dataset using the OS feature codes, before being exported as new shapefiles. Each county boundary was also isolated in this manner and added as a new theme to the GIS projects. The county boundary theme was then used to clip the relevant OS data to the extent of the administrative border (Figure 1).

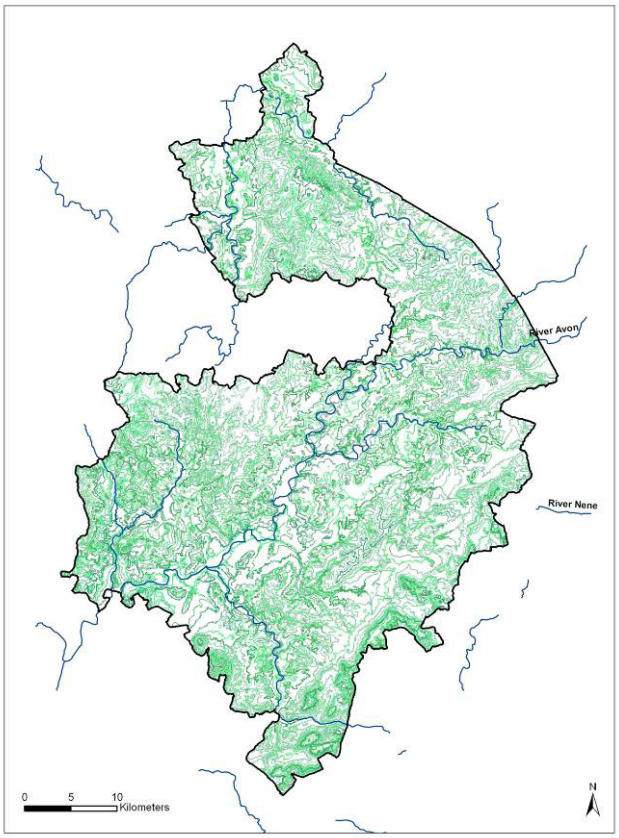

Figure 1: Warwickshire Base GIS

Drift Geology Data

The geological map data was obtained from the British Geological Survey (BGS) on a one year academic licence (provided in ESRI shapefile format). A separate GIS project was set up and all of the individual digital drift geology map sheets added as themes to the display. These map sheets were all merged together to form a single overall drift geology map for all of the five counties under investigation. The county boundary themes were added to the same

project and then each boundary used to clip the overall drift map into five separate county based geology maps. Each of these maps was next added to the relevant county GIS project.

Due to the different sheets being surveyed at various dates and with differing levels of detail, there were obvious discrepancies along the sheet edges. This problem was exacerbated by changes in the nomenclature of identical deposits between sheets, which the GIS software viewed as completely different polygons. Where possible, these issues were solved by altering the colour palette for the shapefile to match similar deposits together. Other than these aesthetic changes, it was decided to keep the BGS polygons in their original format in order to maintain the maximum amount of primary information in the GIS.

The drift geology data provided by the BGS contained numerous attribute fields which were not deemed necessary for the desired research outcomes of the project. The table fields were therefore looked at in turn and those which were not directly relevant were removed from each geology shapefile. New fields were also added to each theme to cover the 'Terrace', 'Date Range' and 'Marine Oxygen Isotope Stage (MIS)' for each geological deposit, which were all missing from the raw BGS data (see Appendix 1).

In order to display the drift geology data in varying formats, the properties of each shapefile were edited in order to alter the 'primary display field' for the file. The primary field was changed to cover each of the main classificatory groupings for geological deposits in turn, with a new exported shapefile being created at each stage. The result of this process was a series of drift BGS themes for each county project, each displaying a different geological characteristic (Figure 2). The classificatory groupings employed are shown below:

- BGS Lexicon Name
- Formation Name
- Member Name
- Terrace Name
- Date Range
- Suggested minimum Date
- Maximum Date
- Probable Marine Oxygen Isotope Stage (MIS)
- Period

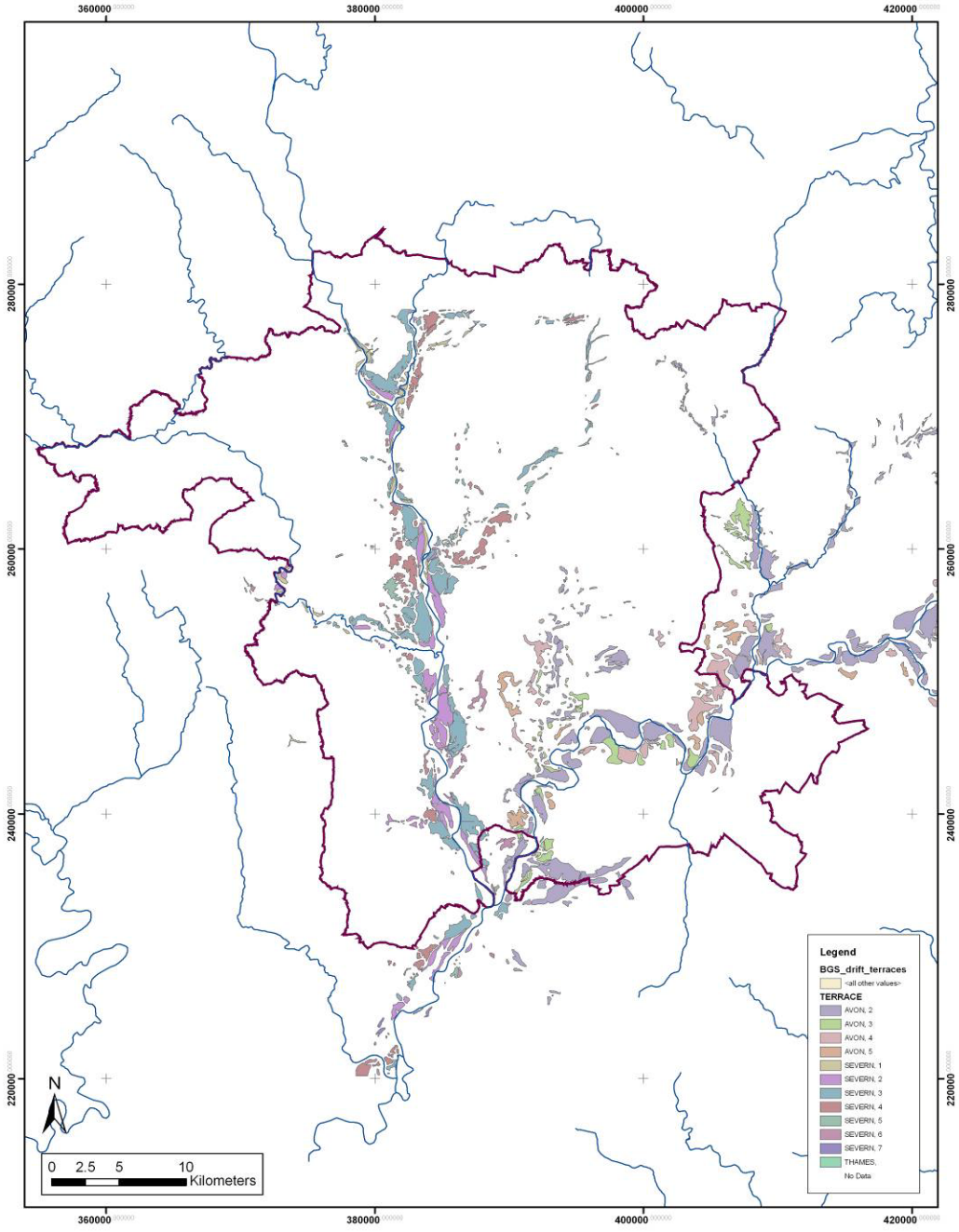

Figure 2: Worcestershire BGS drift data (by terrace)

HER/SMR data

The HER/SMR data for the counties was sent through in varying formats depending on the county office. Where the data was sent through in Word document format, the relevant information was inputted into a new Excel spreadsheet. Where the original data was sent through as an Excel file, this table was simply edited in this format. The scope and capabilities of the HER/SMR database search engines varied considerably between the

different counties. Some county offices provided the data reduced to only those records relating to a Palaeolithic date, whereas others could only provide all 'prehistoric' records for the county. Since it was decided that only finds of definite Palaeolithic date were to be included in the GIS, the data therefore initially had to be checked and refined where appropriate, to remove any non-Palaeolithic or undated records from the lists. This was done by cross-checking the Excel tables with the full monument reports which were also sent through from the HER/SMRs, in an attempt to verify the origin of the records.

The fields included in the Excel tables for the Palaeolithic records were largely based on the style and nomenclature of each individual county (for an example see Appendix 2). Although one overall standardised format would have been preferable, and more useful in terms of querying data across arbitrary modern boundaries, the format for each theme was designed to fit the specific county office it addressed. This decision was taken as the data is to be ultimately integrated into the county HER/SMRs and therefore needs to be in a format which can be readily assimilated into the pre-existing database systems. Without a single standardised nationwide approach to HER/SMR systems, this appeared to be the most suitable approach to the creation of the Palaeolithic GIS themes.

The National Grid references (NGR) for the records were provided as a centroid point with a minimum bounding rectangle (MBR) measurement. As the NGRs were in the conventional Ordnance Survey grid format (i.e. including the 100km grid square letters), they had to subsequently be converted into absolute grid references. Although the 'Manual and Data Standard for Monument Inventories (MIDAS 2003:86)' warns against adding zeros to the end of grid references to produce a location of 'spurious precision', this step was unavoidable in order for the GIS to accept the points. To account for this issue, in some cases it was necessary to create a new field in the point attribute tables called 'Coord_Precision'. This table was based on the recommended INSCRIPTION 'Coordinate Set Precision Wordlist' maintained by the Forum on Information Standards in Heritage (FISH) (www.fish-forum.info/i_csp_e.htm). Each coordinate in this field was therefore allocated a number between 1 and 1000 to record the precision of the point created in the GIS. However, in other cases the MBR measurement already provided a readily accessible gauge of the precision of each point in the files and therefore it was not necessary to create a separate coordinate precision field. There were several records that were missing NGRs altogether and it was decided that these should be left out of the GIS due to the complete lack of spatial positioning. The completed tables were saved as (.CSV) files and imported into ArcGIS. The HER records were provided with X and Y coordinates based on the absolute grid references and the resulting event theme saved as a new shapefile.

Although the HER/SMR points were now integrated into the relevant GIS projects and could be viewed above the underlying drift geology, there was no record in the point attributes themselves about what the geological context of the finds were. A spatial join was therefore performed between the HER/SMR point file and the BGS polygon file. This process used the spatial location of the point data as a means to attach attribute information from the geology maps, before creating a new point shapefile containing the combined attributes of the two datasets. The spatial join appended all of the attribute fields from the BGS data to the HER/SMR theme, whereas only the 'BGS lexicon', 'member', 'terrace' and 'formation' fields were actually desired. Each new shapefile was therefore edited to remove the newly joined fields in the point file, which were not of direct relevance. Maintaining the 'date range' field from the BGS data would have provided a date for the Palaeolithic finds based solely on their geological context, whether they were in their original stratified context, re-worked or from an unstratified surface provenance. The BGS date range field was therefore also removed from the spatially joined file to leave only the date range based on the HER/SMR record itself, even if this simply listed the date as '500000-10001' (i.e. Palaeolithic) (figure 3).

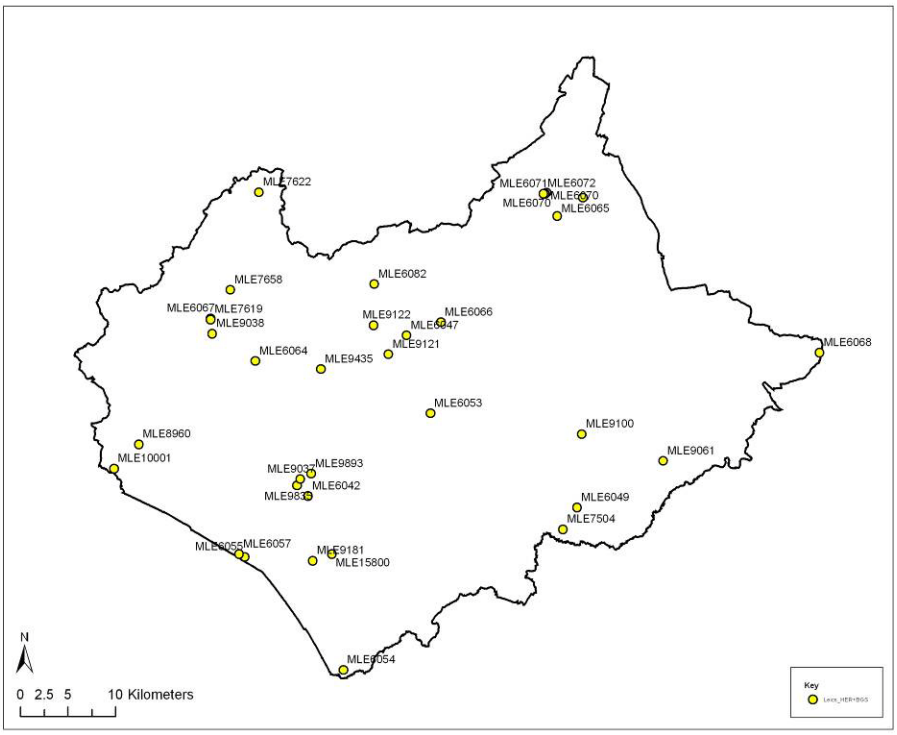

Figure 3: Leicestershire SMR points following spatial join to BGS drift attributes

Waite Collection data

The Waite collection is a private assemblage of Palaeolithic finds, primarily quartzite, from a local enthusiast based in Warwickshire and Leicestershire. The Waite archive is in the process of being analysed by NIAN staff and the resulting interim data was provided in Excel table format by Dr Mark Stevens of Leicester University. The tables initially needed a certain amount of editing prior to their integration into the Leicestershire and Warwickshire GIS projects. Single ID codes had been assigned to each site but with several find entries often coming from the same site. The GIS requires individual point IDs for each entry and so the site numbers were extended to include letters as well as numbers (e.g. 1a, 1b, 1c…). The grid references were again provided in the conventional Ordnance Survey format and therefore had to be converted into absolute grid references. As mentioned above, adding zeros to the end of grid coordinates can create a false idea of the precision of a specific point. A new field was therefore added to the tables to list the coordinate precision for each set of Eastings and Northings, as recommended by the INSCRIPTION wordlists (www.fish-forum.info/i_csp_e.htm).

Since there were numerous finds coming from the same sites and with the same grid references, a new 'quantity' field had to be created to show how many records were coming from each grid reference. This was because if the GIS has more than one point feature for a given location then it will simply stack them above one another, leaving one or more features invisible. Once completed, the table was again saved in (.CSV) format before being opened in the GIS and saved as a new shapefile. The properties of the resulting point theme were edited so that the records displayed as graduated symbols based on the 'quantity' field created in Excel, providing a clearer idea of the distribution and density of the Waite archive finds. A spatial join was again applied to the point data in order to attach the attributes of the

underlying drift geology data to the Waite attribute tables. The newly created shapefiles with the combined attributes were then edited to only leave the 'BGS lexicon', 'member', 'terrace' and 'formation' fields from the original BGS tables (Appendix 3).

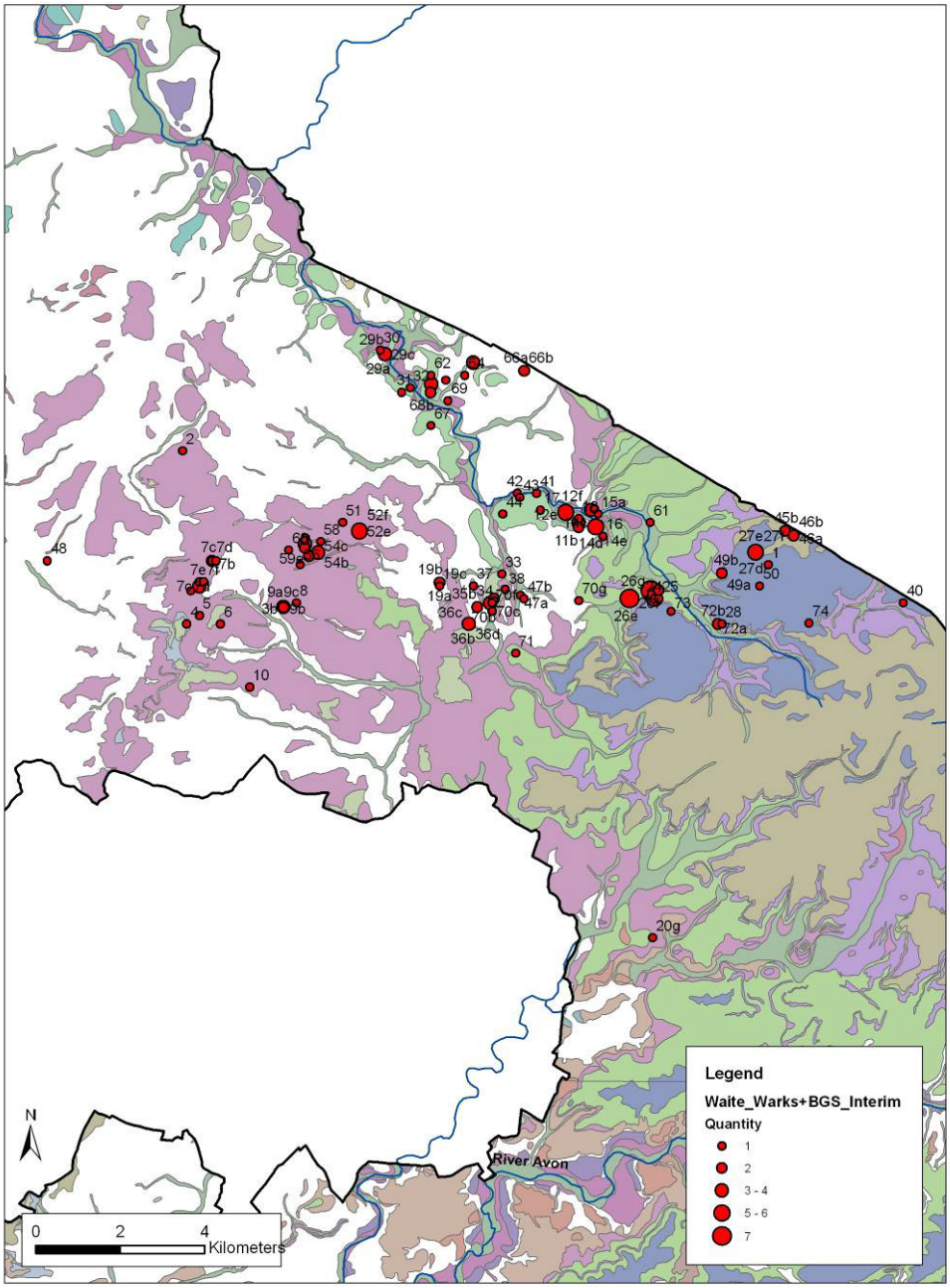

Figure 4: Warwickshire Waite collection finds following spatial join to BGS drift attributes.
The point file is displayed with graduated symbols based on the quantity of finds from each site.

Mammalian Finds

The NIAN GIS is intended to be a spatial database of known Palaeolithic remains, both archaeological and environmental (flora + fauna). To this end, a database of faunal remains from Leicestershire and Derbyshire was obtained by Leicester University from its original source at the New Walk Museum. This database was filtered by Dr Mark Stevens so as to include only the records relating to mammals that are known to be Pleistocene. The filtered database was then provided for integration into the relevant GIS projects for the two counties.

Although the mammalian finds were provided as an Excel file, the format had to be altered to allow direct import into ArcGIS. The table fields were largely left in the form in which they were provided, but several additional fields were necessary. The National Grid References (NGRs) provided with the tables were in conventional OS format and therefore had to be converted into absolute OS grid references. Again, this involved the addition of zeros to the end of the grid references and so a separate coordinate precision field was created, based on the INSCRIPTION wordlist as before (www.fish-forum.info/i_csp_e.htm).

As with the Waite collection archives, there were often several finds attributed to the same grid reference. A 'quantity' field was therefore created to show how many records were coming from each grid reference and to allow the proportional display of all of the records on the GIS. Once completed, these tables were saved in CSV format and imported into the relevant GIS projects. The properties of the themes were again altered to display the quantities of finds as graduated symbols. A spatial join was then applied to the mammalian themes in order to attach the drift geology attributes to the point files, again maintaining only the 'BGS lexicon', 'member', 'terrace' and 'formation' fields (Figure 5) (Appendix 4).

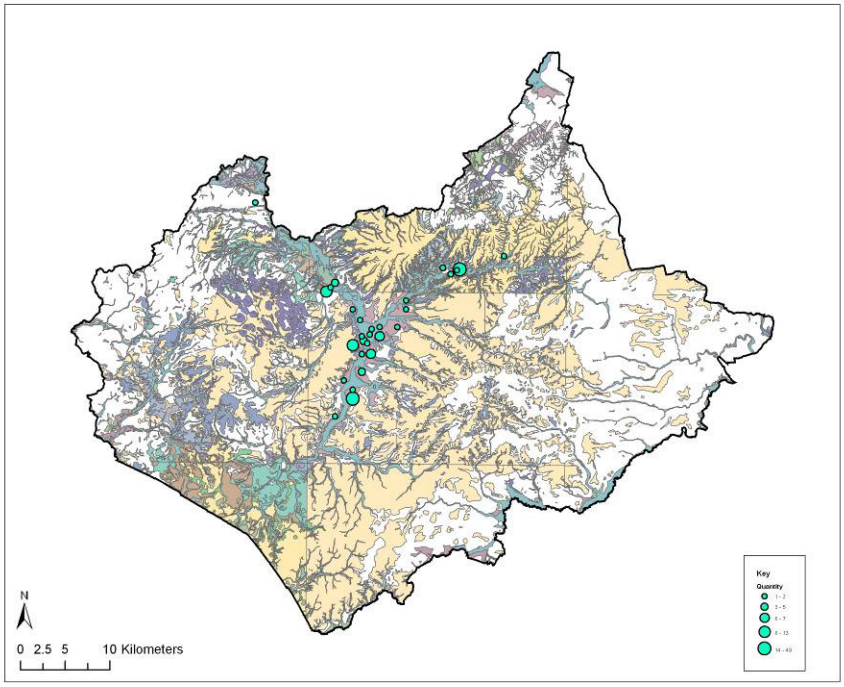

Figure 5: Leicestershire Mammalian finds overlying drift geology map.
Points are displayed with graduated symbols based on the quantity of finds from each site.

Quarry Locations

An important consideration in terms of the distribution of Palaeolithic finds within these Midlands counties was deemed to be the location of aggregate quarries. The 'Directory of Mines and Quarries' (BGS 2005) was used as the basis to gain information on the names and locations of working sand and gravel quarries in the areas covered by the GIS. This information was entered into an Excel table along with the operator of the quarry and the basic geology of the area. The coordinates again had to be converted into absolute OS grid references and so a coordinate precision field was also created. This table was then saved as a CSV file and added to the GIS projects, before being converted into shapefile format (Figure 6) (Appendix 5).

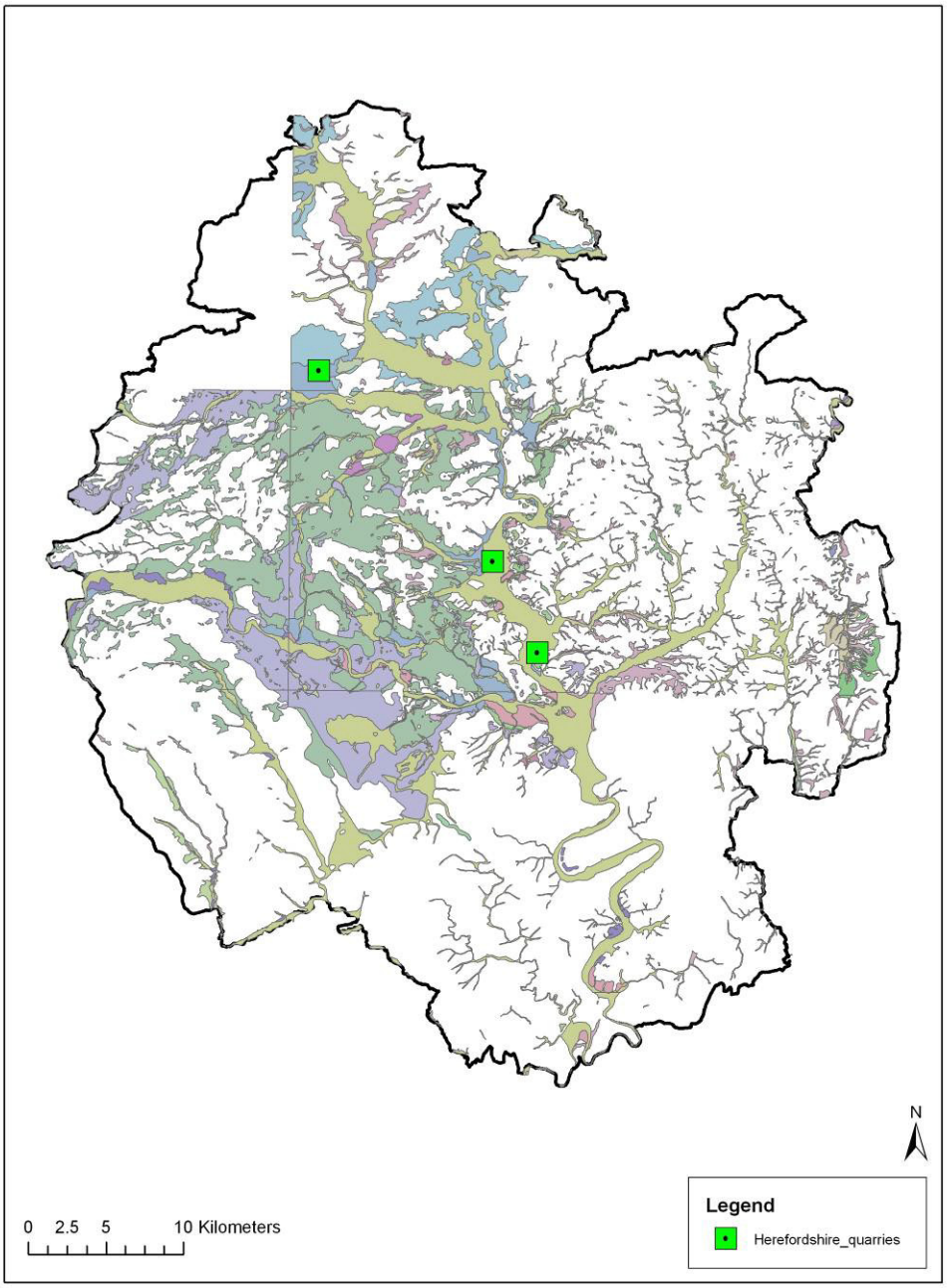

Figure 6: Location of sand and gravel quarries in Herefordshire

Appendix 1 - List of all the reduced BGS drift geology fields in GIS

Table Field	Example Entry
Lex	AILT
Rock	SAGR
Lex_D	Ailstone Member (Warks Avon)
Rock_D	Sand and Gravel
Lex_Rock	AILT-SAGR
Rank	Member
BED_EQ	Not Applicable
BED_EQ_D	Not Applicable
MB_EQ	AILT
MB_EQ_D	Ailstone Member
Terrace	River Terrace 1
FM_EQ	Avon
FM_EQ_D	Avon Valley Formation
Marine Isotope Stage	2
Max_Time_Y	1800000
Min_Time_Y	0
Date Range	
Max_Age	Anglian
Min_Age	Anglian
Max_Epoch	Pleistocene
Min_Epoch	Pleistocene
Max_Period	Quaternary
Min_Period	Quaternary
Sheet	EW184_warwick_v2
Version	2.11
Released	14-4-2005
Nom_Scale	50000
Nom_BGS_Yr	1984

Appendix 2 - List of table fields for HER record theme (Warwickshire)

Table Field	Description
ID No.	HER record identification number.
Name	Name on HER record sheet, e.g. 'Palaeolithic handaxe'.
Record Type	Type of record, e.g. 'Findspot'.
Type	Abbreviated form of Record type field, e.g. 'FS'.
Location	Written description of location, including site + parish.
Context	Context of find, e.g. whether 'Stratified' or 'Unstratified'.
Date Range	Where applicable.
Grid Reference	Centroid reference with MBR in conventional OS format, e.g. 'SP 1820 9820 (MBR: 10m by 10m)'.
Easting	Easting coordinate in absolute grid reference format, e.g. '418200'.
Northing	Northing coordinate in absolute grid reference format, e.g. '298200'.
BGS Lexicon	The rock lexicon assigned by the BGS
Member	The geological member (as assigned by the BGS)
Terrace	The geological terrace (as assigned by the BGS)
Formation	The geological formation (as assigned by the BGS)

Appendix 3 - List of table fields for the Waite collection archive theme

Table Field	Description
Site No.	Allocated NIAN site number (e.g. '31, 32a, 32b').
Parish	e.g. 'Wolvey'.
Grid Ref	Ordnance Survey 100km grid square letters (e.g. 'SP')
X	Conventional Ordnance Survey Easting value (e.g. '433').
Y	Conventional Ordnance Survey Northing value (e.g. '898').
Easting	Easting coordinate in absolute grid reference format (e.g. '443300').
Northing	Northing coordinate in absolute grid reference format (e.g. '289800').
Quantity	Number of finds with the same grid reference (for GIS display purposes).
Period	e.g. 'L-M Palaeolithic'.
Material	e.g. 'Quartzite'.
Type	Type of find (e.g. 'Pointed handaxe').
Max_dimension_mm	Maximum dimension in millimetres (e.g. '128').
Mus_Coll_Ref	Museum/Private Collection reference (e.g. 'Waite Coll'n Box 2').
Waite_ID	Waite Object Label Identification (e.g. 'Astley Field 16').
References	Publication references (e.g. 'Saville and Shotton 1974').
BGS Lexicon	The rock lexicon assigned by the BGS
Member	The geological member (as assigned by the BGS)
Terrace	The geological terrace (as assigned by the BGS)
Formation	The geological formation (as assigned by the BGS)

Appendix 4 - List of table fields for the Mammalian finds theme

Table Field	Example Entry
Box	Box 1
Accession no.	G604.1953
Card	Yes
ID	*Elephas primigenius*
ID Card	*Elephas primigenius*
Description	Part of tusk of mammoth
Description B	[any additional information]
Size (mm)	410x70x70
Condition	Split, coated
Authority	Leicestershire
Locality	Wanlip gravel pits
Locality card index	[Any additional information on the card index]
NGR	SK 606102
Easting	460600
Northing	310200
Coordinate Precision	100
Quantity	2
Horizon	Weichselian, Upper Pleistocene terrace
Horizon card index	[Any further information on the horizon]
Collected	Mr Samarkzuk
Acquisition type	Donation
Acquisition source	Loughborough Museum
Cited	Yes
Notes	Card not in box
Comments	Needs new label
Comments_b	[Any additional comments]
BGS Lexicon	[The rock lexicon assigned by the BGS]
Member	[The geological member (as assigned by the BGS)]
Terrace	[The geological terrace (as assigned by the BGS)]
Formation	[The geological formation (as assigned by the BGS)]

Appendix 5 - List of table fields for the quarry location themes

Table Field	Example Entry
Quarry Name	Attenborough
County	Derbyshire
Grid Reference	SK 506317
Easting	450600
Northing	331700
Coordinate Precision	100
Mineral Planning Authority (MPA)	Attenborough
Operator	Cemex UK Materials
Basic Geology	Quaternary, Hemington Terrace Deposits (River Trent Gravels)