The Spatial Distribution of Iron Age Hillforts in the British Isles

M Simon Maddison

Dissertation submitted in partial fulfilment of
the requirements for the degree of
MSc in GIS and Spatial Analysis in Archaeology
of University College London in 2016

UCL INSTITUTE OF ARCHAEOLOGY

Abstract

Although amongst the most iconic and clearly visible of prehistoric remains in Britain, Hillforts are generally poorly investigated and understood. There are some 3000 known sites in Britain and Ireland.

A project run by Oxford and Edinburgh Universities is creating a definitive Atlas of Iron Age Hillforts in Britain and Ireland; this database is used to identify patterns in the distribution of Hillforts in the British Isles, using modern spatial analysis and Geographical Information Systems (GIS) methods and tools.

Initial analysis focuses on identifying groupings of Hillforts using percolation analysis, based purely on the Euclidean distance between them, a technique applied in geography. This has produced results showing clusters in Britain that have distinctive regional characteristics. It has also been applied in Ireland, showing quite a different pattern.

This study continues with analyses to identify possible territorial and hierarchical relationships between Hillforts within selected clusters, based on the sites' enclosed area, and these are explored for possible explanations. Spatial comparisons are also applied with other classes of monument, boundary data and with finds from the Portable Antiquities Scheme in order to identify spatial relationships and possible patterns of continuity through time.

Acknowledgements

I would like to thank Mark Lake for his guidance, enthusiasm and support for the duration of my Master's course at the Institute, and throughout my work on this dissertation. It was Mark in particular who suggested using Percolation Analysis. I also need to thank Mark along with Andy Bevan for their excellent, stimulating and inspiring teaching on the GIS MSc, and Andy for giving me the impetus to actually join the course in the first place.

Elsa Arcaute from the Centre for Advanced Spatial Analysis at UCL has been particularly helpful in explaining Percolation Analysis as used in Geography, and in providing 'R' code kernels that got me started. Stuart Brookes from the Institute of Archaeology at UCL has also been very helpful in discussions concerning the analysis of historical spatial datasets, and for interesting ideas on other parameters that might be drawn in for enriching the study in future.

It was Gary Lock's excellent course on Iron Age Hillforts at Oxford University that first engaged me academically with Hillforts, and it was then that I first heard of the Atlas project. He has been enthusiastically supportive of my working with the Atlas data and provided useful suggestions and guidance; he has also given me the wonderful opportunity of getting down and dirty on the excavation of the Hillfort at Bodfari in north Wales for the past three seasons. Paula Levick also of the Atlas project has been extremely helpful in providing not only extracts from the Atlas database over the past year, but latterly a full export, and patiently answering my questions.

Damon Ortega in the GIS lab provided me with his database of Stone Circles in the British Isles, extracted from Burl. My fellow students this past year, Marta Krzyzanska and Alex Sammut provided me with data on the Ringforts of Ireland, generated as part of their group project.

Thanks to my fellow students in my first year, particularly Kit Jones, Mads Jorgensen, Lokwalo Thabeng and Giorgia Amici for their good fellowship, enthusiastic collaboration and shared learning.

Lastly and not least, thanks to my wife Barbara for her patient indulgence of my passions and interests, long hours and absences over the years and for her enthusiasm and enormous support, particularly in my difficult times.

Table of Contents

List of Figures

Preface

Hillforts have fired my imagination as far back as I can remember, tramping over the rough Cornish countryside on childhood summer holidays, whether in breezy sunshine or driving wind and rain. Poring over the treasure trove of Ordnance Survey maps in the summer twilight was a further inspiration, seeking out new sites to explore.

With spectacular views, but located in seemingly wild and often inhospitable locations, and each site clearly representing a significant commitment in terms of physical effort, what drove their creators to construct these sites, and how do they relate to the landscape and to each other? Studying Hillforts on Gary Lock's course at Oxford, consequently engaging with the Atlas of Iron Age Hillforts project and taking the Masters in GIS and Spatial Analysis in Archaeology at UCL has enabled me to realise the dreams of a lifetime, and start to address these questions on a systematic and analytical basis. Inevitably many questions have been raised, but it has opened more avenues of enquiry to explore.

This dissertation was prepared as part of the requirements for the MSc in GIS and Spatial Analysis in Archaeology at UCL, the course was started part time in September 2013. Studies were interrupted in 2014-15, and recommenced in September 2015.

Introduction

Iron Age Hillforts

Iron Age Hillforts (which will be referred to simply as 'hillforts') are amongst the most visible and iconic of prehistoric remains in the British Isles. They have been the subject of active investigation and study particularly since the 19th Century. Nonetheless despite continuing and extensive research, in many ways they are still poorly understood (Brown 2008).

In terms of description, as opposed to function, Hogg (1979, 1) defines a hillfort as "an enclosure with substantial defences, usually on high ground and probably built between about 1000BC and AD700 but showing no significant Roman influence". Harding (2012, 1) identifies the key element being "enclosure, physically or conceptually demarcating an area to which access is restricted or controlled", incorporating both humanly constructed features and natural topographic elements such as cliff-top, marsh or slope. Both recognise hillforts as neither necessarily restricted to being on hills, nor primarily being for defence, but serving a range of purposes which vary both geographically and temporally.

Smaller sites are generally excluded by both authors as well as by Lock for the Atlas of Iron Age Hillforts Project (pers.comm., April 2016) , including many of those known as 'duns' or 'brochs', in order to consider only those large enough to be regarded as "community sites, serving a social unit larger than a single family or household" (Harding 2012, 12). This is discussed further later.

Hillforts are very diverse in form and location, and many are monumental in scale. A few examples are presented to illustrate this point and provide some visual references.

Maiden Castle in Dorset (Figure 1), arguably the most iconic of hillforts, is monumental in scale, with multiple ramparts and ditches and imposing and complex entrances, representing a huge amount of manual effort. It is hard not to see defence being a major factor in its design. Like Uffington in Berkshire (Figure 1) it is both dominant and prominent in the landscape. Pen-y-cloddiau in the Clwydian range in north Wales (Figure 2), and the largest hillfort in that region, is both enormous and dominant with spectacular views for many miles in all directions, but it is high up and hard work to reach, suggesting some difference in primary function, perhaps more concerned with stock management and transhumance, and used only seasonally.

Not all hillforts are on hills, as with Cherbury in Oxfordshire (Figure 3), which is located on an island in (now reclaimed) marshland, but is nonetheless dominant in a relatively flat landscape, and with multiple ramparts and ditches. As with Pen-y-cloddiau a primary stock management and protection function seems plausible. Common in south-west England, Wales and Ireland are coastal promontory hillforts, taking advantage of the natural topography to build a formidably defined enclosure, with proportionally much reduced labour compared to a landlocked site. One of the best in south-west England is The Rumps (Figure 4), which has ready access to agricultural land and grazing in its hinterland. Whilst very attractive to visit on a sunny summer's day, experiencing it during storms and much of the winter it proves to be exposed and inhospitable, suggesting seasonal use and a mix of functions.

Figure 1 - Hillforts: Imposing in scale and dominant in the landscape
Maiden Castle, Dorset (top); Uffington, Berkshire (bottom)
(English Heritage; S Maddison; S Maddison)

Figure 2 - Pen-y-cloddiau, north Wales (www.archaeology.org)

Figure 3 - Not all on hills, Cherbury, Oxfordshire (S Maddison)

Figure 4 - On coastal promontories; The Rumps, Cornwall (Cornwall HER, S Maddison)

The sheer size and scale of hillforts creates particular challenges to their investigation, in particular to any kind of systematic excavation. Attention tends to focus on ramparts, ditches and entrances. Interiors, particularly for larger sites, have often been ploughed, reducing the evidence of occupation (Hogg 1975, 34), for example Cherbury (Bradford 1940) and Segsbury (Lock, Gosden and Daly 2005). Only in exceptional cases such as at Danebury (Cunliffe and Poole 1991; Cunliffe 1993) have extensive excavations taken place over large parts of the interior, or as in the Wessex Hillforts Project (Payne, Corney and Cunliffe 2006) where the interiors been comprehensively subject to geophysical survey. The level of information available is however beginning to change with the widening availability of Lidar images and aerial photographs, through programmes of aerial survey, for example *Airborne Remote Sensing* , but of course this still needs to be applied in detail to individual sites.

Dating of hillforts also presents particular challenges, as discussed later, with sites that may have gone through many stages of development over many hundreds of years. Personal experience of rampart excavation at Moel y Gaer Bodfari (*Moel-y-Gaer, Bodfari, Denbighshire*) over three seasons has shown the multiple phases of construction, the variation of construction style even within a trench 5m wide, and the paucity of evidence to establish any kind of dating, let alone the timeframes of construction phases.

Aims and Objectives

This study has collaborated with the Atlas of Iron Age Hillforts project (*An Atlas of Hillforts in Britain and Ireland*), henceforth called the 'Atlas', and uses the data generated for a series of GIS based and spatial analyses investigating various aspects of the distribution and location of Iron Age hillforts in Britain.

The objective of the dissertation is to identify patterns of hillfort distribution in Britain and any regional variations, then focus on a few areas in detail, with the aim of generating explanations as to these patterns, and further questions and directions of research. It also aims to deliver a set of practical and openly available tools which can not only be applied to the remaining areas, but which can be taken up, used and developed further in the future by others for analysis on other datasets.

Geographical Groupings

Is it possible to identify 'natural' groupings of Iron Age hillforts throughout Britain that may have existed in the Iron Age?

The Atlas survey is built up on the basis of modern counties and nations. These counties do not necessarily align with any regional or geographical divisions that may have existed within the Iron Age. It is therefore important to establish, in the first instance, suitable intrinsically coherent geographical regions that can reasonably be used for analysis. Initial thoughts were to base this selection on simplistic geographic characterizations (e.g. central mountainous Wales, south-west England). However in exploratory discussions Mark Lake suggested a more analytical approach using Percolation Analysis, a technique very recently applied to Domesday sites by Arcaute, Ferguson, Brookes et al. (2014). This had generated very promising results in identifying groupings of sites that match known medieval administrative regions and helped to unpick the complexity of their origins.

It was therefore decided to apply this technique, using simple coordinate locations of all Iron Age hillfort sites, to attempt to establish any 'natural' or intrinsic groupings that might exist. This did not initially rely on detailed completion of the Atlas, but simply required coordinates of Atlas adopted sites (and where not finalised, candidate Atlas sites). The process could then be re-run as data was updated or refined. This analysis was based simply on the site coordinates, using Euclidean distances. In the earlier stages of the project, insufficient data was available from the Atlas, and the hillfort index created in the 1970's by Hogg (1979) was used to develop the programs and analysis techniques, and offered some useful initial results.

Once identified, these groupings were then compared with topography, geographical regions, as well as counties and other historical data sets. Comparison with cultural evidence would have been interesting (e.g. pottery styles) but was outside the scope of this project.

Hierarchical Relationships

Is it possible to determine putative hierarchical relationships between hillforts, on a regional scale, based on known hillfort attributes?

Working on a regional-scale, spatial analysis has been applied to currently observed/ known hillfort attributes drawn from the Atlas, to the extent that they are available, such as site area, vallation and number of entrances, to identify possible spatial distributions or implied hierarchy on this basis. The working assumption has been that the hillfort is a point with attributes. Using the enclosed area of the hillforts readily offered the most interesting results, and this has been pursued for the British Isles as a whole and also for the particular regional areas selected for detailed study.

Topography

Is it possible to establish whether there are topographical characteristics that are significant determinants in hillfort location, on a wide area, regional or local scale?

Some analyses have been conducted on hillfort and terrain elevations, both at a Britain-wide and regional cluster basis, to identify whether hillforts really are on hills, how they utilise the regional topography, and whether these patterns vary in different localities.

Comparisons

What can be learned by comparing hillfort distribution with other historical datasets, in terms of cultural continuity and interaction through pre-history and into recorded historical times?

Four comparisons have been carried out with other data sets to establish the potential for seeking evidence of cultural continuity and evolution over long periods of time. The data used is:

- The locations for Iron Age finds from the Portable Antiquities Scheme, which applies to modern England only
- The sites of Ringforts in Ireland
- The sites of Stone Circles, across the British Isles
- Domesday County boundaries, in England only

Terminology

For brevity and convenience, the following terminology is used for the analyses and discussions within this document:

British Isles – The whole of mainland Britain, the whole island of Ireland, the Isle of Man, also the Shetlands, Orkney Islands, Western Islands , Inner and Outer Hebrides, Anglesey, Isle of Wight, the Scillies and other small islands. The Channel Islands are NOT included in this study.

Britain –Britain incorporating Scotland, England and Wales, with their islands as described above, and the Isle of Man. *Mainland Britain* is the same but excluding the Isle of Man.

Ireland – The whole island of Ireland, incorporating the Republic of Ireland (Eire) and Northern Ireland, including the Irish and Northern Irish coastal islands.

Isle of Man – occasionally referred to distinctly, but most analyses and discussion incorporate it within *Britain.*

Modern national and political entities (England, Scotland, Northern Ireland, Eire, Wales and Isle of Man) and modern county names are used for regional descriptive purposes, for convenience. Unless otherwise explicitly stated, there is no implication in using any of these terms in either a political or national sense, whether modern or historic.

Hillforts are frequently also described as ‘sites’, particularly in terms of physical and spatial location.

A review of previous studies

The earliest scholarly studies that included descriptions of hillforts go back many centuries (Cunliffe 1991, 1-20; Harding 2012, 29-35) and cannot be covered here, but in terms of spatial analysis Cyril Fox's oft cited 'Personality of Britain' (1952), first published in 1932, is a notably comprehensive work from several points of view. Published by the National Museum of Wales, it sets out a context not just of Britain, but the British Isles including Ireland, and that within a wider context of continental Europe. It provides many maps, which must have been painstaking and time consuming to create, of many different classes of objects and sites, including: megalithic monuments (page 12); Iron Age 'A' and 'B' cultures (page 18, page 30), including classification of hillforts based on in-turned entrance type; hillforts in north Wales, showing their relations to mountains (page 76); a physical map of the British Isles (opposite page 28) showing the 'lowland and highland zones' of Britain, which will be referred to again later. These maps provide powerful images for qualitative analysis, covering a very wide range of evidence and environmental factors. Substantially the work of Lily Chitty (Lynch and Burgess 1972),they set an impressive benchmark for later researchers to follow.

Nash-Williams (1933), writing shortly after, included an appendix on the distribution of hillforts and other earthworks in Wales (pages 311-315) including a beautifully detailed map, covering some 822 sites. These were also tabulated by county and type (e.g. promontory, hillfort, hut-circle) and also by elevation; their relationship to Roman sites was noted, and a tentative attempt to assign them to tribal territories made, with much discussion about the routes of migration that resulted in their construction.

The conference held at the Institute of Archaeology in 1958 (Frere 1958a) on the Iron Age in Southern Britain includes a paper by Rivet (1958) covering the preparation of the OS Map of Southern Britain in the Iron Age (Ordnance Survey 1962) and the choices made for classifying sites in it. Entrances were excluded because of the complexity and frequent difficulty of identification (without excavation); even vallation had its complexities depending on the specific topography and methods of construction. Size however was less ambiguously ascertained, and was split into three categories at thresholds of 3 and 15 acres (1.2 and 6 ha), figures for which he does not claim any 'particular magic', but were arrived at after some experimentation. He also discussed smaller regional groups such as within the Cotswolds, based on details of type, which could not be drawn out in the map. Highly pertinently, he ended with a plea for a national survey of hillforts to be carried out, something now happening over 50 years later under Lock and Ralston (*An Atlas of Hillforts in Britain and Ireland*). In the same volume Aileen Fox (1958) goes into considerable detail of the 'south-western type', found in Cornwall, Devon, parts of Somerset and south Wales, including those with widely spaced ramparts, the outer enclosure probably used for livestock. Her analyses include good outline plans of many sites, as well as a distribution map (page 36), and show the importance of very detailed work, site by site. The OS map itself (1962) includes in the introduction a number of coin distribution maps and a gazetteer. Like the conference and Frere's own paper (1958b), it perhaps inevitably dwells on the coin evidence, Roman records and then current theories of invasion from the continent, with rather compressed dating within the latter Iron Age. Despite its visual quality and great aesthetic appeal, it is perhaps as a detailed gazetteer it is most valuable.

Rivet also implies that an OS map of northern Britain was planned, however never published, but one for Scotland is included in Rivet (1966), along with distribution maps for other classes of object. In the same volume Piggott (1966, 4) proposes a scheme of provinces and regions in Scotland and Feachem (1966, 85-86) goes into some detail as to the background for this, based on the different classes of Iron Age monument.

One of the early numeric analytical studies of hillfort distribution was by Newcomb (1970) who manually applied nearest neighbour analysis to the hillforts of Penwith in Cornwall, following an earlier discussion of the potential of quantitative spatial analysis in the same area (1968). His conclusions were however strongly criticised by Hodder (1971) because of possible edge effects, as Newcomb acknowledged (1971), demonstrating both the challenges of computational analysis, but also its strength in that it can be questioned and revisited and revised in a methodical way.

Prior to this, the earlier approaches were what might be called intuitive. Hogg (1971), whilst not deprecating those efforts, noted their weakness, in particular their subjectivity and hence poor repeatability. He argued in favour of numeric approaches and introduced methods by example, including coin distributions around Bagendon, and a variety of interacting methods to assess the territory and population associated with hillforts.

Further drawing on Clarke (1968, 508-509), Hogg (1971) drew Thiessen polygons around hillforts south of the Thames, weighted according to the size of the hillfort, and providing a quite nuanced picture of possible territories and relationships, whilst recognising its limitations. Cunliffe (1971, 59-60) similarly drew Thiessen polygons around hillforts in Wiltshire to establish possible territories and relationships. Building on this Grant (1986) applied a more powerful tool (Xtent) that could incorporate different characteristics with the aim of revealing relative dominance and evolutionary patterns; this tool would be interesting to explore further. Collis (1981) looked theoretically at hillfort distribution patterns and relationships within the wider context of other types of settlement, and developed models of how relative importance and roles change over time. Although illustrative, this was however not applied to real examples.

Earlier (Hogg 1972) had developed a very broad ranging analysis of hillforts in Wales and the Marches based on their size and vallation, and manually established groupings using circles around specific sites, as well as different settlement forms. His statistical rank test based on size is hard to understand, but something similar is presented as a log-log rank size plot in Hodder and Orton (1976, 72) (credited to Pierson-Jones (1973)), with a computation of the best fit power law. Hogg's work was probably ahead of its time, and as with other work of this period the impact of Carbon-14 dating had not been absorbed; it is admirable considering it was done without computers, and it would be worth revisiting and investigating his ideas further.

Gent and Dean (1986) analysed the size of hillforts and plotted them with respect to the productivity of radially determined catchment areas in Wessex, and obtained some quite good correlations, generally also tying in with other hillfort characteristics.

Acknowledging the impact of Carbon-14 dating was by then beginning to have on hillfort studies and chronologies, Hogg (1975, 37-45) worked further on maps with symbols based on hillfort size, arguing that size was their most important single attribute. He added two further thresholds of 0.24 and 30ha to those previously used (1.2 and 6 ha) and originally set by Rivet (1958). He used these to identify seven or eight distinct regions in Britain, with some tentative explanations of function, whilst emphasizing the consideration of both size and function in local context. He also noted that whilst hillforts are a characteristic of Iron Age culture in Britain, they are not an essential one and this will be returned to later.

Lynch, Aldhouse-Green and Davies (2000, 147), harking back to 1972 work (Lynch and Burgess), also presented a map of hillforts in Wales with symbols based on the same size ranks as the OS Map (1962) but attempted no analysis.

Harding (1976) includes work by Mackie (1976) on Scotland and Savory (1976) on Wales, with distribution maps, but there is no analysis.

Cunliffe (1991) looked at hillforts within the context of Iron Age communities as a whole, with much work on distribution patterns of pottery types and other artefacts. He was very cautious in his interpretation of hillforts because of his perception of poor quality and very limited evidence from previous excavations, and the fact that in proportion so few sites had been excavated. He focused most on hillforts in the area where there was the most excavation evidence available, viz. central southern England. Hogg (1971) by contrast argued strongly for using the observable evidence of hillforts i.e. from surface survey of what was visible, recognising that this in itself represented a lot of work still to be done. Cunliffe (1982; 1991) did discuss regional zones e.g. (1991, 527) and how they evolved through the period; along with pottery distributions these are interesting for future comparisons and further study.

There is too much in Sherratt's "Why Wessex?" (1996) to adequately summarise here, but discussions of the role and location of hillforts in the context of trade, transport, portages and river systems provide valuable ideas which are taken up later. In particular these include their role in the transfer of goods across catchments, their domination of passes through obstacles, and control of routes along coastlines and into the hinterland. Other notable points include observations on the nature of long distance trade being in compact high value goods, the coincidence of the hillfort zone (Cunliffe 1991, 369; Cunliffe 2013, 304) with the Avon catchments, and the changing importance of particular geographical areas through the impact of technological developments and shifts in trade routes and patterns. These ideas continue to be developed in ArchAtlas, but unfortunately for this study not for Britain's Iron Age and hillforts (*ArchAtlas, Version 4.1*).

Brown (2008, 196-204) takes up this theme and develops it in further detail, with a particular focus on Wales and the Marches. In support of arguments concerning the Severn and Wye Valleys (pages 198-199) he undertook detailed numeric analysis of hillfort distances from the rivers and their spacing, taken from his unpublished thesis (2002), and discussed many individual sites, which will be returned to later.

Bradley (2007) suggests that despite the huge amount of effort involved in their building, hillforts may have been rather on the margin of a much more prosperous region orientated towards the North Sea with a strong relationship to northern Europe, with different styles of settlement (pages 261-262). This is not dissimilar to the observation of Hogg (1975) noted above that hillforts were not an essential element of Iron Age Britain. This is revisited later, in the comparison with Iron Age finds from the Portable Antiquities Scheme.

Batardy, Buchsenschutz, Gruel et al. (2008) undertook a spatial analysis of some 11,000 Iron Age sites in France, (excluding the Mediterranean maritime provinces and Corsica), albeit not specifically hillforts. These were plotted based on the number per department rather than by individual coordinates, reflecting stylistic and cultural characteristics and for different periods.

The use of departments may have been for ease and speed of analysis, but it does give a modern political overtone to the work, which may have been intended.

Llobera (2001) developed tools for analysing the prominence of locations in the landscape, which he continues to teach at the University of Washington. There has not been time to pursue these within this study, but it is potentially highly relevant for the topographic analysis which has been touched upon later, and would likely be a fruitful area for further research in the future.

In summary, early work on the spatial distribution of hillforts (and other types of evidence) was implemented through often beautifully drawn maps, but perhaps of necessity was qualitative and intuitive in interpretation. However from the late 1960's numeric/ computational techniques started being applied, most notably by Hogg (1972). What comes out of nearly all these later studies is the importance of drawing on detailed data and the wider context and understanding of the sites analysed in order to complement quantitative analysis. Computation itself is no silver bullet to providing a comprehensive picture of the past; it is a tool, a means to an end, and needs to be used in a context of broader knowledge. It is also clear that as compared with subjective studies, the strength of such approaches is that they are generally repeatable, accessible for review, revision, and further development.

Data Sources

Fundamental to this study was the database created as part of the Atlas of Iron Age Hillforts Project (*An Atlas of Hillforts in Britain and Ireland*). This four year research programme was started in September 2012 and proceeded generally on a county by county basis. Although some counties were available quite early, this did not give the data necessary for an analysis covering either Britain or the whole of the British Isles. In fact the database was not in a sufficiently complete state to start working with until April 2016.

In order that development of the analyses could start in advance of this, a database of Iron Age Hillforts was created using A H A Hogg's index from the 1970's (1979). This both allowed development and testing of the initial analysis programs, provided a benchmark for comparison with the Atlas, and also provided an insurance policy in the event that the Atlas data would not be fully available for this dissertation to be completed within the deadline.

Hogg's 'British Hill-Forts'

This comprehensive book (Hogg 1979) was prepared in the early 1970's and declared itself to be complete up to 1976, acknowledging that anything published after mid-1976 was not included. The material is manually typed and tabulated, and represents a commendably major commitment of effort, particularly considering it was done in a pre-personal computer age.

The key data provided for each site within this index includes: national OS grid reference; the name of the site; a simple classification and one or more primary references. An area field is also provided, giving the main enclosure area but is far from complete. This work covers Britain, and does not include Ireland or the Isle of Man. UK grid XY coordinates of the sites were generated from the index and used for initial development of the percolation analysis programs. The data extraction and validation process is described in Appendix A.

A summary of Hogg's classification is included in Appendix A. However there are a few points to note. Many sites are classified as 'Duns', whether they be in Scotland or not, being sites below 0.1ha in enclosed area. Some sites are classified as excluded, as although originally claimed as a hillfort, subsequent work rejected this. The index includes some sites where multiple entries cover different phases of the same hillfort, represented by different forms and classification. For the purpose of analysis those sites were described by a single site location. Apart from these duplicates, all sites were included in analysis of the dataset, as the primary purpose was to develop analytical techniques.

The index generated includes a total of 3588 unique sites, and is maintained within a Microsoft Excel spreadsheet, along with all the derived coordinate data described above and core KML code for use in Google Earth.

The Atlas Project Data

As described earlier, although some individual county data became available in 2015, the complete data for Britain was not ready until around April 2016, and this was still being reviewed. A final 'confirmed' database was provided on 6/5/16, in the form of a Microsoft Excel spreadsheet, exported directly by the Atlas team from the source database implemented in FileMaker Pro.

This database contains some 93 fields, with a very wide range of attributes, comments and some textual descriptions. Key information was exported from this spreadsheet for the analyses; principally the coordinate data, site name and the unique Atlas index number, as well as a select set of numeric attribute data. There were of course practical issues, and it took approximately one month to clean up the data. See Appendix A for more details.

The confirmed sites database comprised a total of 2869 sites in UK and Ireland, including 390 in Ireland, 22 in the Isle of Man and 2457 in mainland Britain. This was actually several hundred sites smaller than some earlier draft versions of the database, but the project directors were clear that they only wished to include confirmed sites.

It is interesting to note that the Hogg database contains 3588 sites for the British mainland, compared to 2457 for the Atlas, as many of the sites in Hogg's index had been rejected. In some cases this is because they have been categorised as fortified farmsteads rather than hillforts. However it also interesting to note that some sites in the Atlas are not in Hogg's work; this is generally those that have been subsequently recognised through Aerial Photography and were unknown in the 1970's.

The database is maintained as a Microsoft Excel file, as for the Hogg data. This was then used to export key parameters and data in csv files for analysis as described below, whether computationally or within the selected GIS package.

Note: where hillforts are mentioned by name, the Atlas index is included in brackets immediately after, to aid referencing, although the Atlas data is not included within this document.

Qualification regarding Hillfort Data

Before moving on to the way in which these data sets have been used for analysis, it is worth highlighting some qualifying issues regarding the data, and the possible impact this has on the outcomes and conclusions of the study.

Firstly there is what has or has not been included in the datasets. For Hogg, an unqualified assumption has been made that all sites are regarded as hillforts, including those that he classified as excluded, but the primary aim of generating and using this data was to develop techniques and generate provisional results.

For the Atlas, the strict criteria applied by the Atlas project team of only including sites that are confirmed has been followed. Draft versions of the database included other sites where the evidence is either not sufficiently robust to confirm, or they otherwise fall outside the project's criteria. Both datasets of course omit unknown sites that have been lost and which may never be found. This study therefore is focused specifically on what can robustly be classified as a hillfort, according to the criteria of the Atlas Project. In summary, the selection criteria for the Atlas are:

- The site's topographical position in the landscape, should be dominant
- Scale of the enclosing works, should be imposing
- Enclosed area should be greater than 0.2ha

Two out of three of these are required.

Additionally there are two reliability categories:

- Reliability of data
- Reliability of interpretation

These can be either confirmed, unconfirmed or irreconcilable. Confirmed sites are where at least the reliability of interpretation is confirmed. Sometimes the data is poor, with for example unconfirmed features, but where this site is nonetheless obviously a hillfort (Gary Lock, Paula Levick, pers. comm. 4/4/2016).

Further studies could valuably include a broader spectrum of Iron Age sites, including those less robustly classified as hillforts, potentially offering a finer grained picture of patterns of Iron Age activity, and also confirming or otherwise the patterns identified in this study, as attempted by Hogg (1972) for example.

A final qualification on the data regards dating, or more specifically the question of which particular sites were in use at any one period of time. Whilst the datasets used comprise all sites as defined above, this does not in any way imply that all sites were in use at the same time, as highlighted by many studies, for example Cunliffe (1991) and Hogg (1971). Unlike the analysis of Domesday Vills (Arcaute, Ferguson, Brookes et al. 2014) which uses a remarkably precise snapshot of England in one specific year, a similar claim cannot be made for this study of hillforts.

Whilst the database does include dating information where known, this only covers quite a small proportion of sites. 2365 are classified as unknown date. Certainly as more information becomes available this can be refined, but the weakness remains. What can be said is that sites are generally highly visible as part of their very nature, and would therefore have been recognised as such even if they had passed out of regular use. In conclusion it can be said that hillfort sites reflect an aggregate of Iron Age activity over the period, but can point to possible socio-political entities, for which other evidence might be sought in the future.

Other Data

Other data used in the study were principally map outlines and topographic Digital Elevation Models (DEM), for Britain, the Isle of Man and Ireland. The map outlines are used for display and also for defining the bounds of spatial analyses, and the topographic data has been used for rendering within ArcGIS. The DEM and high water coastline for the British mainland were downloaded from Edina (*Digimap Resource Centre*) and for the Isle of Man, the Government Mapping Survey (*Isle of Man Government - Maps*). The Irish DEM was provided from fellow students, derived from the CGIAR (*Consortium for Spatial Information*); the coastline from merged administration county outlines for Ireland and Northern Ireland is from Irish Townlands, based on OpenStreetMap data (*Irish Townlands*). The river data for Britain was obtained from the Ordnance Survey Open Data website (*OS Open Rivers*). The map of Domesday counties was provided by Stuart Brookes (pers.comm., April 2016).

The Portable Antiquities Scheme data base was used to generate a set of coordinates for finds classified as Iron Age (*The Portable Antiquties Scheme*). Fellow students provided similar data for the Stone Circles of the British Isles, extracted from Burl (2000) (Damon Ortega, pers. comm., May 2016), and for Irish Ringforts, extracted from the historic and environment records of Ireland (*National Monuments Service*) and Northern Ireland (*Northern Ireland Sites and Monuments Record*), (Alex Sammut, Marta Krzyzanska, pers.comm., May 2016).

Mapping the data

As described above the data from both Hogg and the Atlas were used to generate kml files, for display in Google Earth (Figure 5), to allow both datasets to be visually inspected, validated and compared.

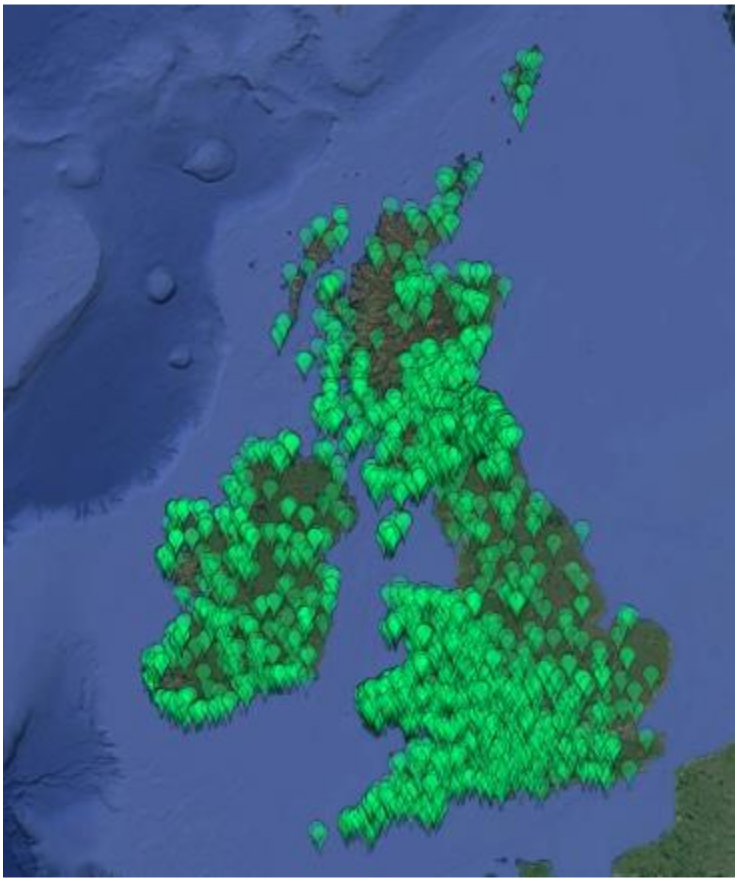


Figure 5 - kml display of Atlas sites in Google Earth

The specific kml code was designed to allow easy visualization (e.g. Figure 6) and to avoid the confusing visual clutter that is so common with kml displays.

Display within Google Earth immediately made clear the highly non-uniform distribution of sites, as well as their regional variation.

Figure 6 - Sites from both Hogg (red) and the Atlas (green)

This was a simple, effective and powerful tool to identify obvious coordinate problems, for example sites located in the sea, inspection of sites where other problems came up whilst analysing site specific data, or during computational processing of the coordinates (Figure 8). Comparison between Hogg and the Atlas was also quick and easy. Figure 6 shows sites in Central Cornwall, with Hogg only sites in red, Atlas only sites in green, and sites which are in both, albeit with minor differences in coordinates, as a combination or closely overlapping. It also continues to be very valuable as an on line gazetteer for reference purposes.

Figure 7 - Example site in both Hogg and the Atlas highlighted, and a site only in Hogg (above and right)

Figure 8 - Example close up in Google Earth showing both Atlas and Hogg markers The Hogg marker highlighted (Castle-an-Dinas, St Columb, in Cornwall)

Percolation Analysis

Background

Percolation Theory was originally developed in the 1940's as a means of describing gelation as small branching molecules bonded to become macro-molecules, but has since become much further developed and widely applied (Stauffer and Aharony 1991). Essentially it is a way of mathematically describing and analysing clusters of spatially arranged points. Conceptually the definition of a cluster is simple. A distance threshold is defined, then for any given point all neighbouring points falling within this threshold are part of the cluster. The test is then re-applied for each of these neighbours, and any further points meeting this criterion are also part of the cluster. The test is based on the Euclidean distance between points, and can be applied at any scale (Stauffer and Aharony 1991). The term percolation applies to movements between points determined by the threshold parameter. A good example is the propagation of fire in a forest, where a burning tree may or may not ignite a neighbour, depending on how far away that neighbour is (and other conditions of course). Other examples include the diffusion of hydrogen through solids, and the flow of gas and oil through porous rocks (Stauffer and Aharony 1991).

Extending this widely diverse range of applications, percolation theory is being used in geography for example, in identifying urban systems and hierarchies through the density of street interconnections (Arcaute, Molinero, Hatna et al. 2016). This reveals patterns that have evolved over millennia through the influence of culture, politics, administration and trade, expressed as the modern patterns of streets and roads. This is not purely academic and has the potential for directly influencing regional development policy, for example.

Arcaute, Ferguson, Brookes et al. (2014) have collaborated in applying the same technique to Domesday Vills and the administrative territories of hundreds, wapentakes and shires in 1086AD to peel back the palimpsest of regions, territories and administrative boundaries, reveal Domesday administrative organisation, and relate it to modern Britain. Building on earlier work on the Anglo-Saxon state by Brookes and Reynolds (2011) and the Landscapes of Governance Project, they apply percolation analysis to sites extracted from the Domesday Book (e.g. *The Domesday Book Online*).

It was this work that led to the application of the technique to Iron Age Hillfort sites, and both Elsa Arcaute and Stuart Brookes provided advice, data and source code in support (pers.comm., 2015-2016).

Methodology

A suite of programs to perform percolation analysis has been developed in the R statistical programming language (*The R project for Statistical Computing*), based on core code provided by Elsa Arcaute. This uses an algorithm very similar to the City Clustering Algorithm developed by Rozenfeld, Rybski, Gabaix et al. (2011), and as described qualitatively above. Clusters are identified by creating a 'graph' of nodes based on distances within a defined radius. More detail on this and the programs are provided in Apppendix C. Note that this works satisfactorily for datasets of a few thousand points, but would not manage the 10's of thousands of points of the street intersection analysis (Arcaute, Molinero, Hatna et al. 2016), where programs written in the C programming language were used. Further code has been developed to compute and generate the percolation transition graphs (Figure 13, Figure 17 below) as presented in that work, which are improved from earlier forms used in the Domesday study (Arcaute, Ferguson, Brookes et al. 2014).

This program suite has been applied to the Hogg and Atlas datasets for mainland Britain, and to the Atlas dataset for Ireland and the Isle of Man. It has also been applied to randomly located sites in these geographies. Full details and code are provided in Appendix C.

Note that this method of identifying clusters within a set of spatial points is quite similar to DBSCAN (*DBSCAN; fpc*, 38-40). DBSCAN provides two levels of cluster membership, and defines clusters on the basis of a minimum number of points that must be within the given radius. Points that satisfy this condition are 'core', and those that are not but are reachable via core points are 'density-reachable' or 'non-core'. There is also a category of 'outliers' or 'noise' which are neither. Depending on the value of the minimum number of points, this reduces the effect of single points or thin chains of points linking clusters. Percolation analysis is essentially a reduced case of this with the minimum number of points being 1, and therefore having only a single category of cluster membership. Comparison of results using the two different techniques would be worthy of future investigation, particularly in terms of how clusters change through different radii.

Results and discussion

Figure 9 shows a simple plot of Atlas Hillfort sites in Britain. It is immediately apparent that the distribution is highly non-uniform, with a significant concentration of sites in south and south-east Scotland, and south-west Wales. There is a more even spread through south and south-west England, with a very low density north and east of a line roughly between the Thames Estuary and the Dee up to Northumberland and the Scottish Borders. Figure 10 shows the same number of sites plotted using a random process (see Appendix C for more details), reinforcing the nonuniformity of hillfort locations. Figure 11 shows the distribution for Ireland, and Figure 12 the same number of sites randomly plotted. At first it is not quite so evident, but closer inspection shows that there is a marked concentration of sites close to the coast, particularly in the south-east, south and west, and it contrasts significantly with the distributions in Britain.

The figures following show the outputs of the percolation analysis of data from the Atlas project. Note that the legend on each map shows the rank order of clusters by size. They have a standard colour coding with the largest cluster in red, then dark blue, then green and so forth. The number of the cluster is as generated by the clustering program, and will not necessarily be the same for each radius value. Clusters will change colour as their rank alters at different percolation radii.

Figure 13 shows for Britain a plot of the maximum cluster size vs. the percolation radius. Note that as the percolation radius gets larger, all sites will eventually merge into a single cluster. The importance of the plots of maximum cluster size vs. the percolation radius (Figure 13 for Britain and Figure 17 for Ireland) is in highlighting potentially interesting transitions and critical thresholds as the radius value is increased and new larger clusters appear; specific examples are picked out below. This is an effective tool in identifying clusters that are potentially worthy of further investigation as they appear at these transitions. A second factor in selecting clusters for further analysis is the persistency of a cluster through a range of radius values, suggesting that it is more strongly spatially distinct, and is less sensitive for example to the omission of a small number of peripheral sites, whether through survival or categorisation. This is achieved through inspection of successive cluster plots.

Figure 13 very clearly shows that at 36km there is a major transition, which is where Scotland and northern England (excluding some remoter parts) join the cluster for the rest of England and Wales, as shown in Figure 14. There are a series of transitions at lower radii, the most interesting of which are in the ranges 7-9 km and 11-14km.

The 36km plot in Figure 14 also shows that the predominantly Scottish cluster (blue) includes sites in England as far south as a line roughly between Morecambe and Flamborough Head, with sites further south largely in the red cluster. Looking at England as the radius values reduce, sites in the Pennines and the east progressively break out of the bigger cluster, and at 14km, the south-west peninsula forms its own cluster in Cornwall and part of Devon, and other clusters appear in the south-east.

The plot for 12km (Figure 15) shows for example Cornwall (yellow) and Devon/part of Somerset (green) as individual clusters, and a cluster along the Chilterns (mauve). The plot for 9km (Figure

16) shows north-west Wales (mauve), the Clwydian Range (pink), south-west Wales (blue), the Gower (light blue), central Wales and the Marches (yellow) and two clusters on the north-west and the south-east of the Severn Valley (turquoise and pale orange), the latter being the Cotswolds. Some of these clusters have been selected for more detailed analysis, see 'Discussion of Specific Regional Clusters' later.

Comparison with the same percolation analyses run on the Hogg data showed broadly similar results, although there are some differences and some distinctive clusters particularly in south-west England, but this is unsurprising given that it has over 1000 more sites, and is noticeably denser in this part of England. This does highlight the importance of what is or is not categorised as a site, but also suggests that repeating the analysis for a broader range of sites, including known occupation sites, would potentially give a more fine-grained picture of clusters.

Looking at Ireland for 12km (Figure 18) the clusters are predominantly along the coast, and for 19km (Figure 19) some of these clusters have merged and some have linked in to inland clusters. The differences in Ireland are partly due to the relatively low density of sites as compared to Britain. These have not been investigated further in this study.

Note that there is no sense in which percolation is being posited as a cause of these clusters forming, rather that it is a descriptive process that inductively highlights underlying patterns. These are based purely on the Euclidian distance between the sites, without any reference to topography or other data. The question then is whether these clusters are in some sense a relict (or 'footprint' as the term Arcaute, Ferguson, Brookes et al. (2014) use) of former socio-political entities, and explanations for these are sought, as is discussed in more detail later, in 'Discussion of Specific Regional Clusters'.

A second important point is that there is nothing necessarily special about particular values of the percolation radius, and that transitions will not necessarily occur for the same values in different regions or countries. This is clear in comparing Britain and Ireland in Figure 14 and Figure 17, where there is a significant difference in the patterns, although interesting transitions do occur in the 12-14km radius range for both. This latter distance is arguably a comfortable day's walk. How this might vary in different topographies and how that compares with specific clusters would be an interesting question to explore in the future.

Figure 21 shows a plot of the maximum cluster size vs. radius for the random plot of sites in Britain (Figure 10). This shows just a single transition, between 10-15km radius values. A cluster map for this data at 14km (Figure 22) shows this to be the point where Scotland north of the Firth and Cornwall both appear as clusters; at 13km southern Scotland and the far north-east of England form a separate cluster, partly along the line of the Tyne, Tees and Eden, and in the south-west much of Devon as well. This clearly illustrates the effect of natural boundaries, in these cases the consequence of being on a peninsula, in northern Scotland's case separated by an isthmus. Analysis of repeated runs of the random generation process would demonstrate how robust this is in terms of the specific radius value, but the principle is quite clear. At the risk of generating a circular argument, it does quantitatively illustrate the geographical factors that help regions such as Cornwall and northern Scotland to develop their own clear distinct identities. It does of course become self-evident the more extreme the geographic isolation is, ultimately when it becomes an island.

One of the principal objectives of applying percolation analysis was to identify groups of sites for more detailed analysis in other ways, as opposed say to using some potentially arbitrary criterion such as a modern or historic county boundary, or a topographically defined region. To this end, four clusters in Britain have been selected, and are investigated in more detail, namely Cornwall, the Cotswolds and the lower Severn Valley, central Wales and the Marches, and the Chilterns. These have been identified from the transitions in Figure 13 and as described above. These clusters are covered in more detail later, in 'Discussion of Specific Regional Clusters'.

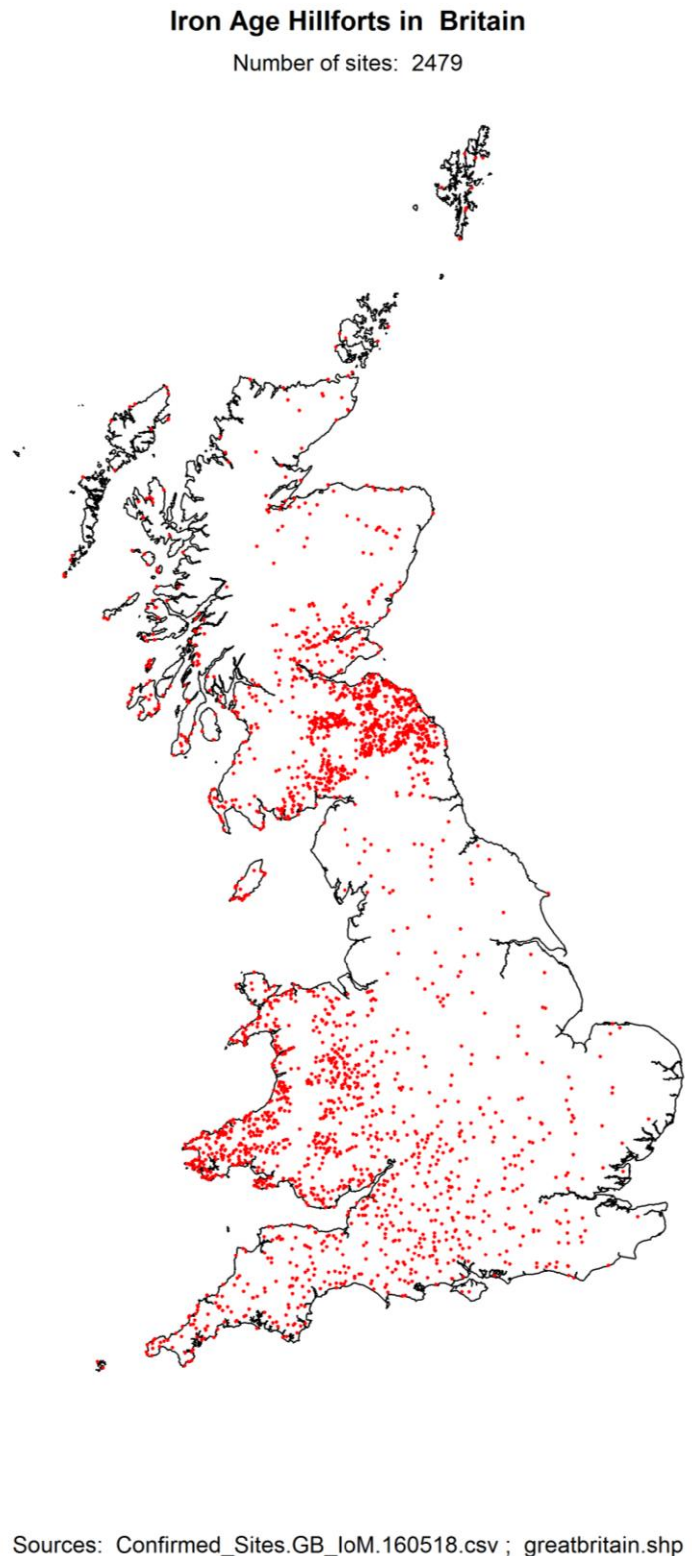


Figure 9 - The Distribution of Hillforts in Britain

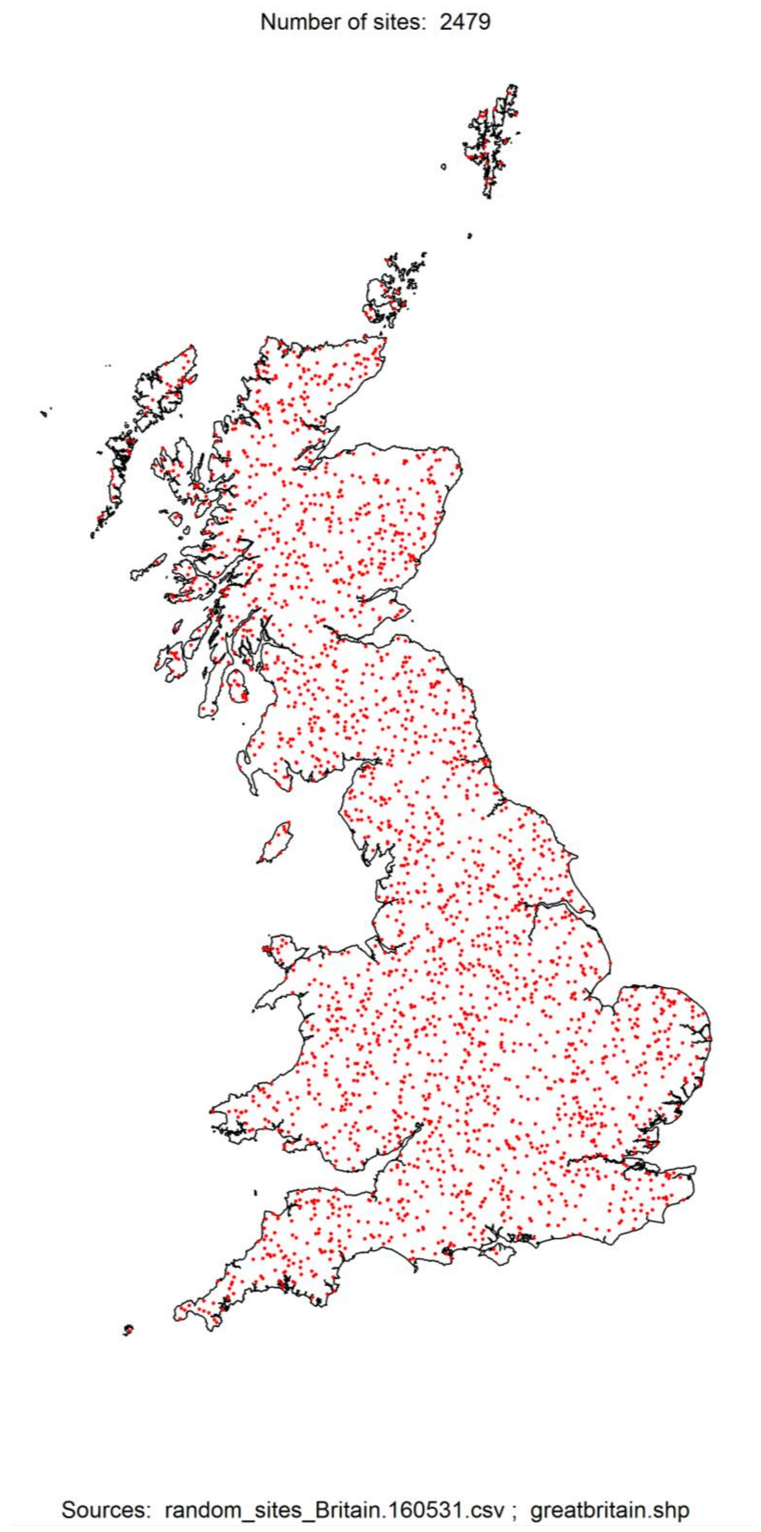


Figure 10 - Random distribution of sites in Britain

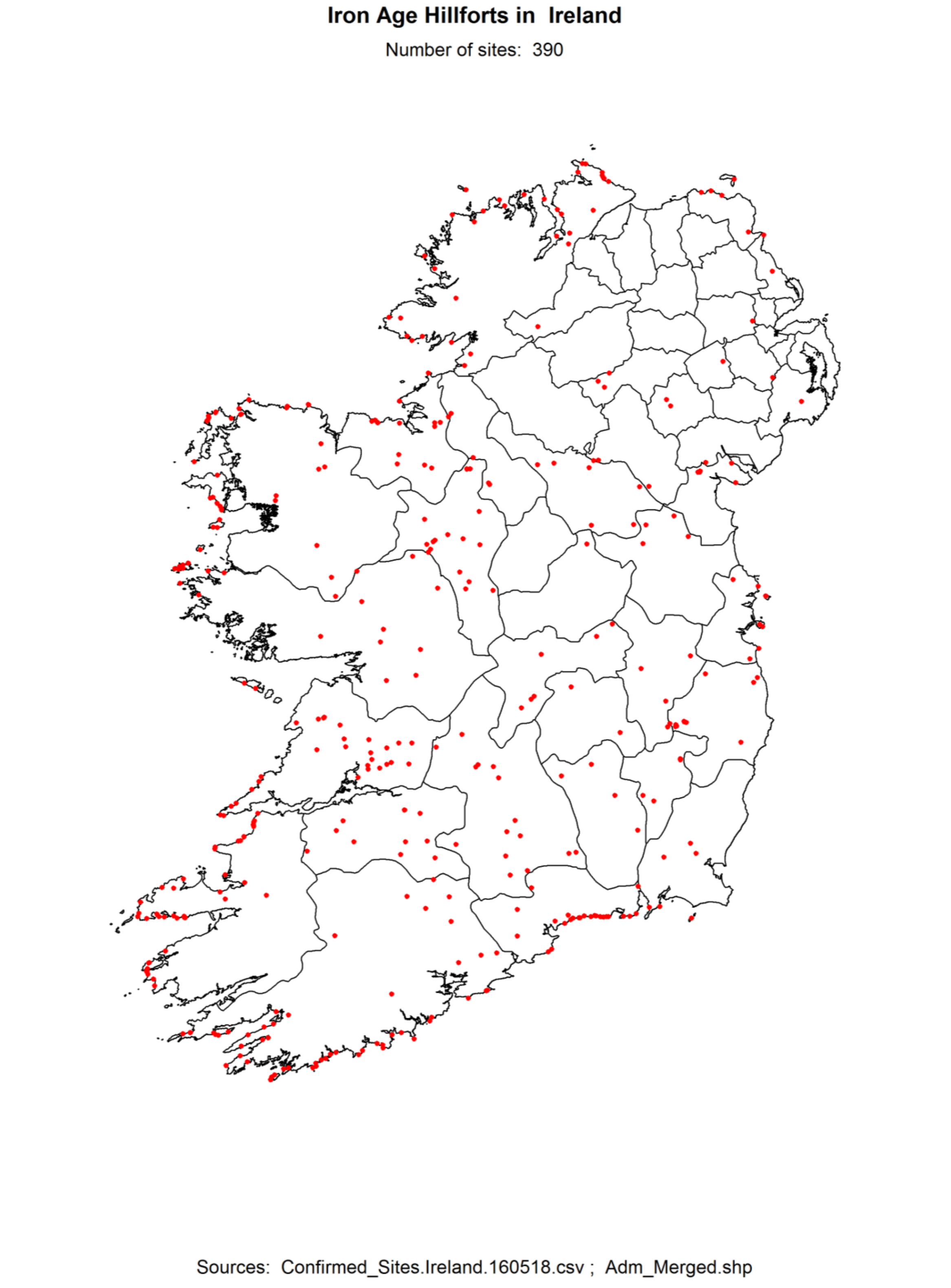


Figure 11 - The Distribution of Hillforts in Ireland

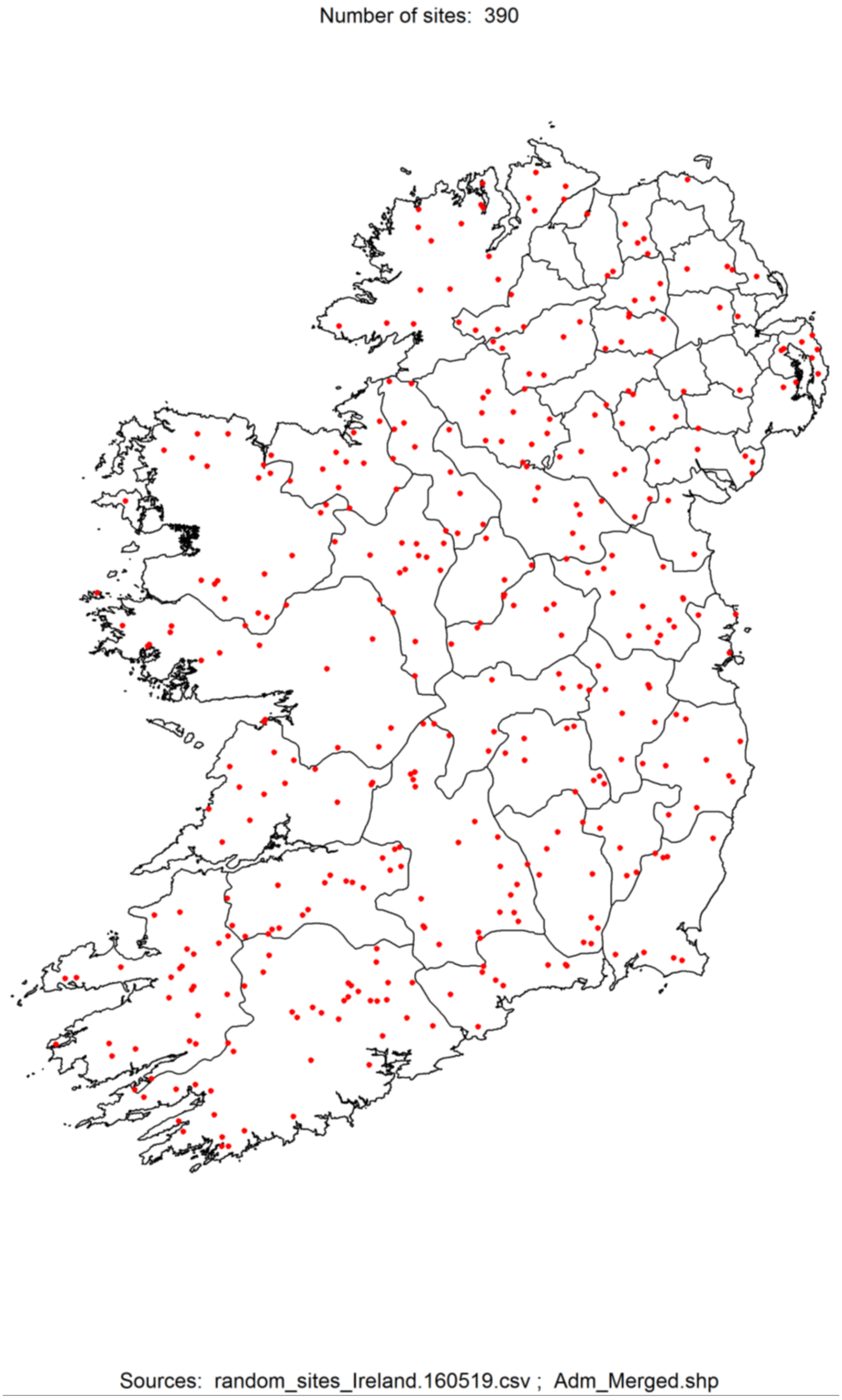


Figure 12 - Random distribution of sites in Ireland

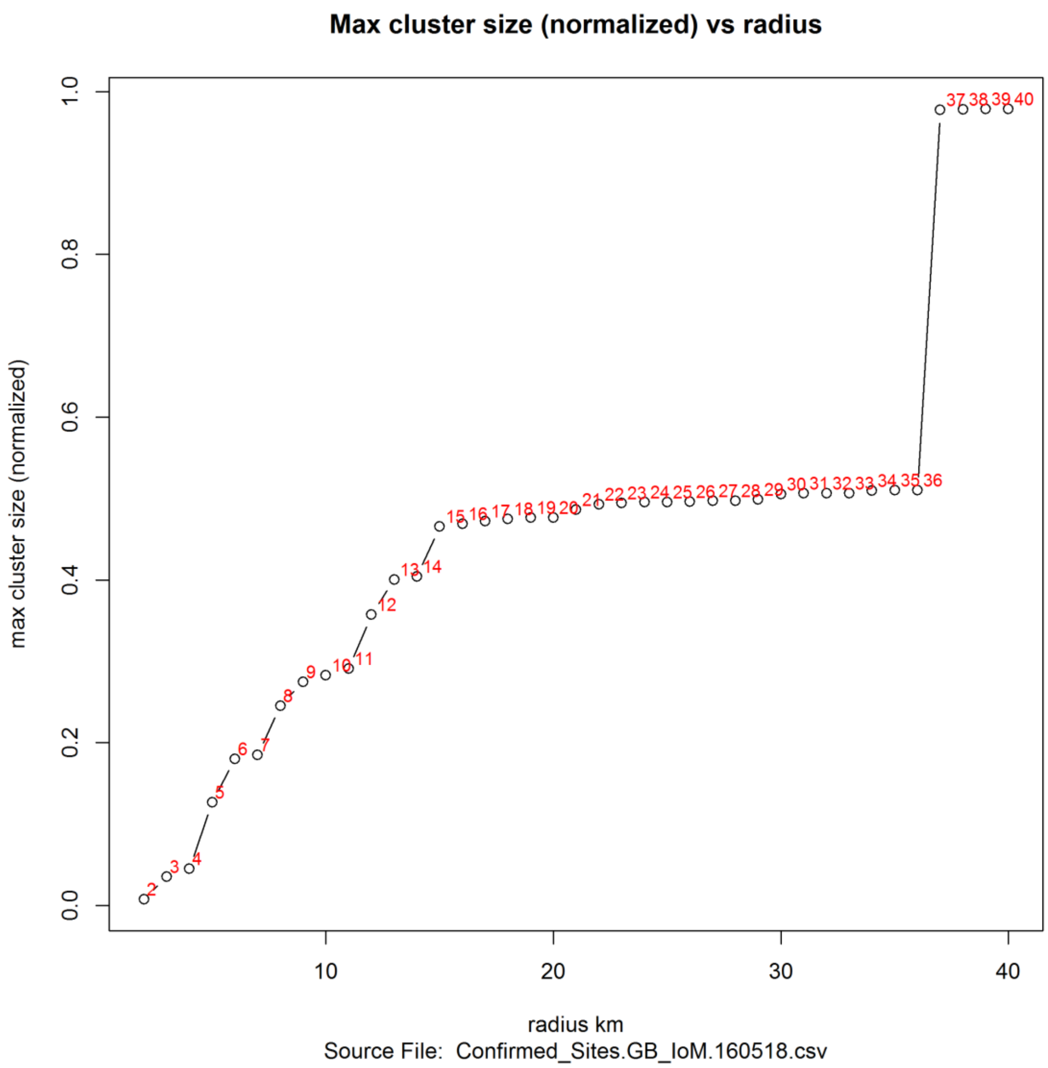


Figure 13 - Maximum cluster size vs. percolation radius, Britain
The radius value is plotted in red adjacent to each point

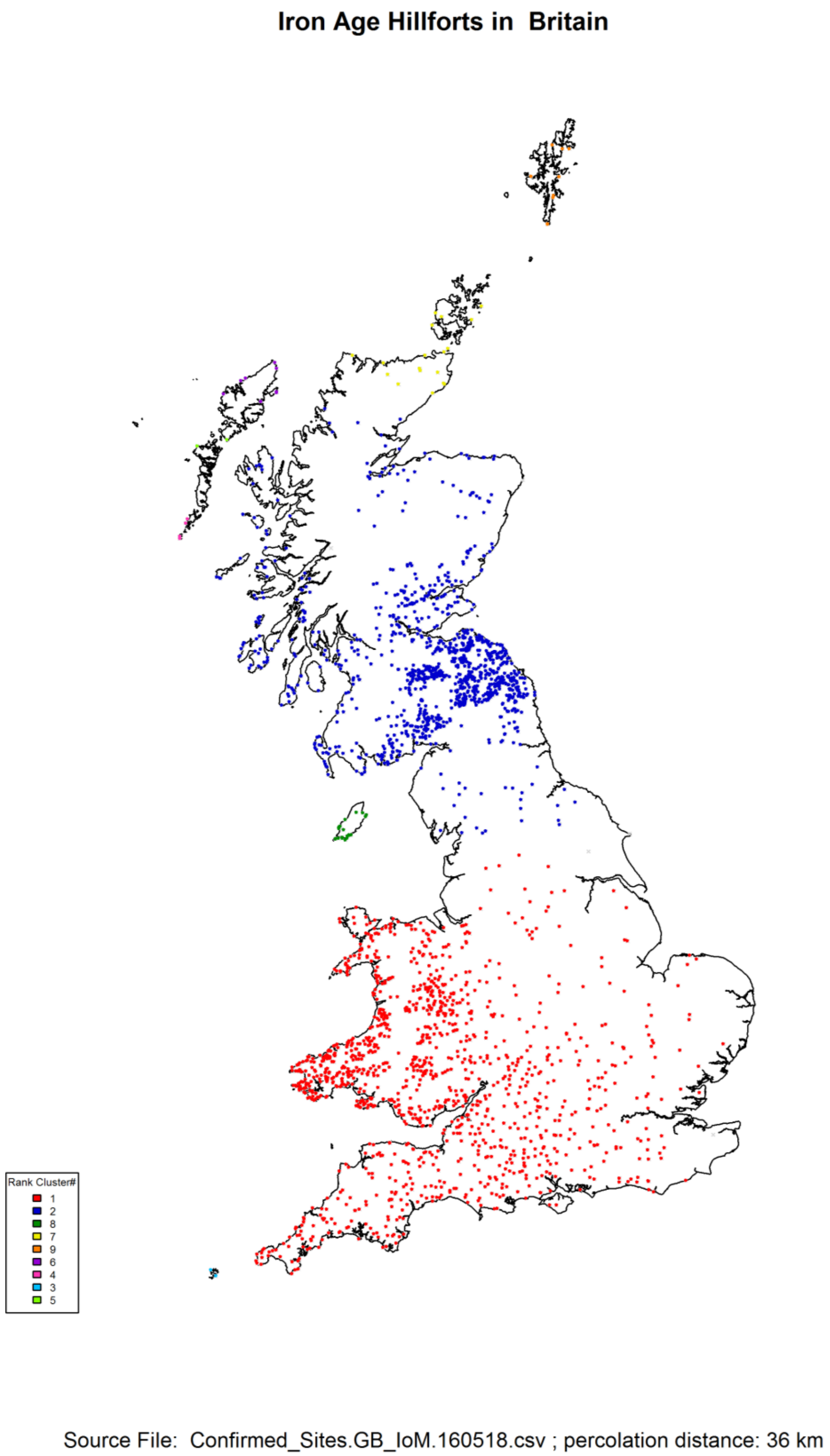


Figure 14 - Hillfort clusters in Britain; 36km percolation radius

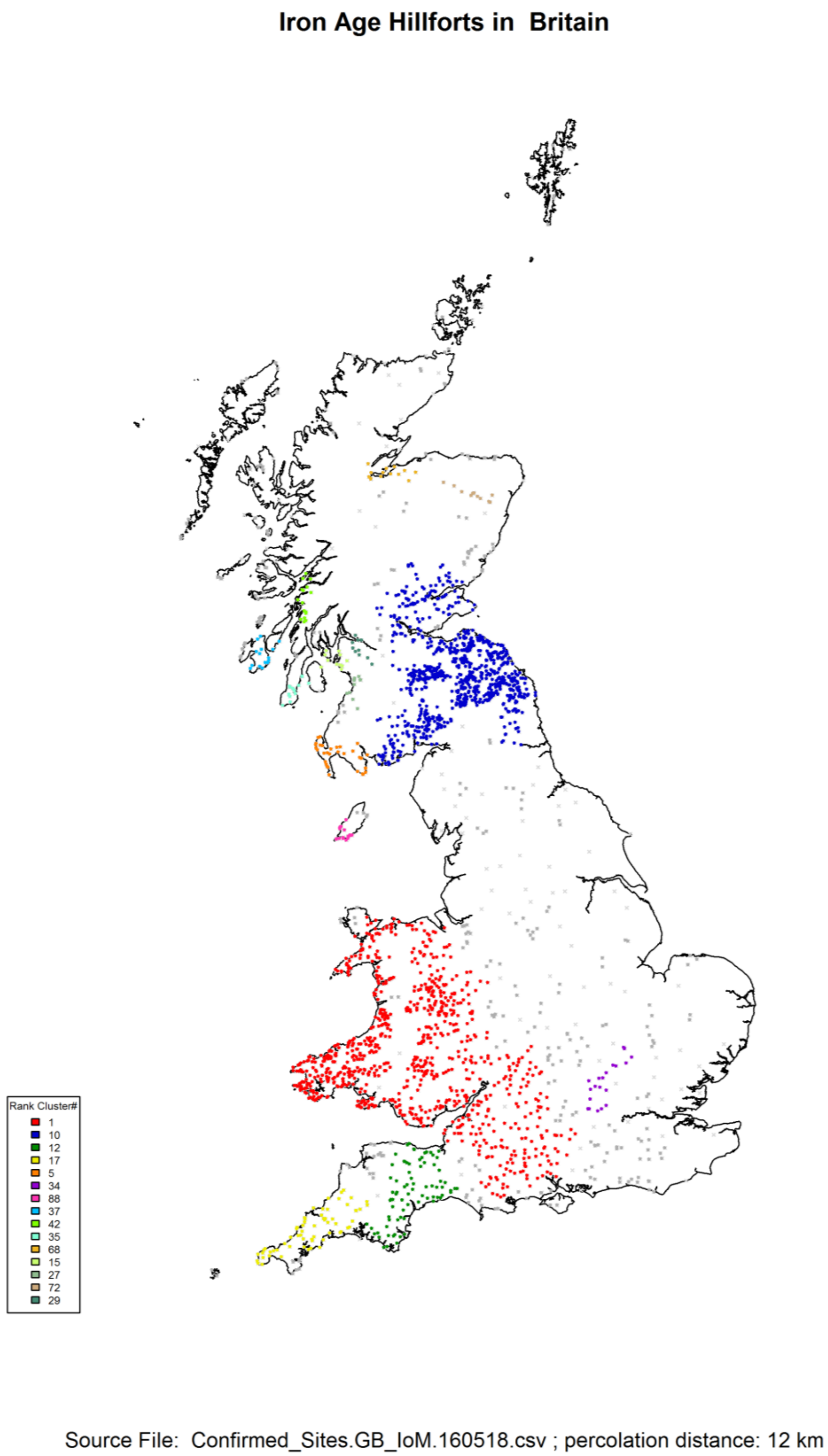
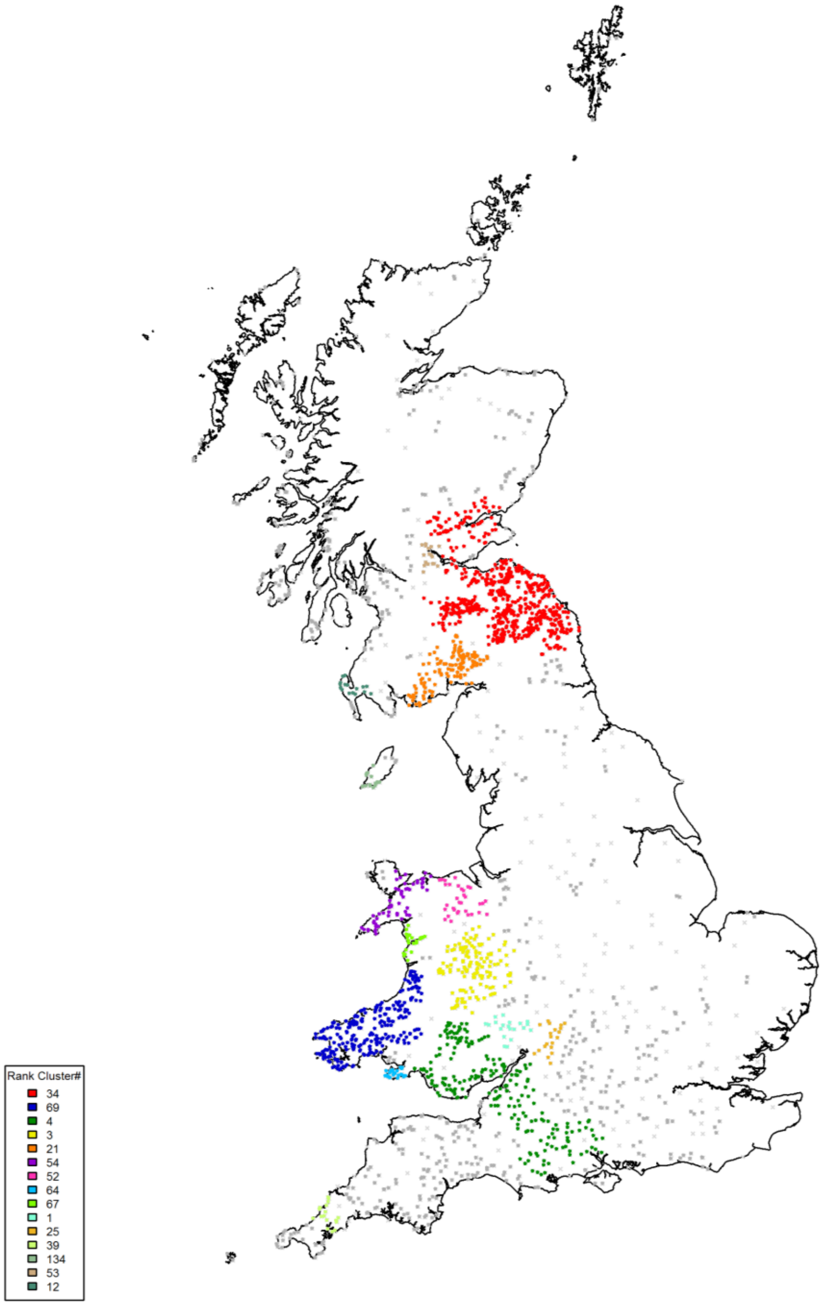


Figure 15 - Hillfort clusters in Britain; 12km percolation radius

Iron Age Hillforts in Britain



Source File: Confirmed_Sites.GB_IoM.160518.csv ; percolation distance: 9 km

Figure 16 - Hillfort clusters in Britain; 9km percolation radius

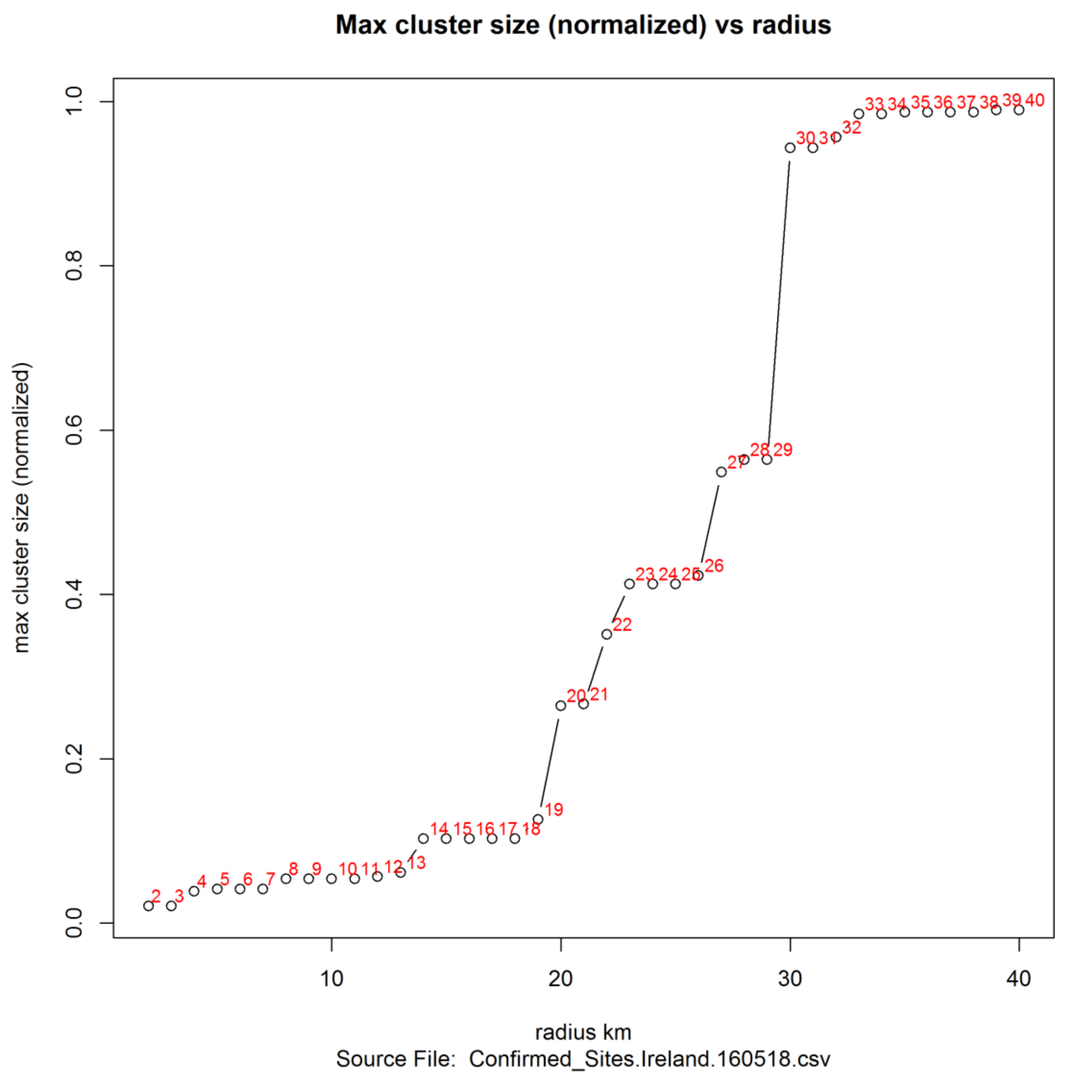


Figure 17 - Maximum cluster size vs. percolation radius, Ireland
The radius value is plotted in red adjacent to each point

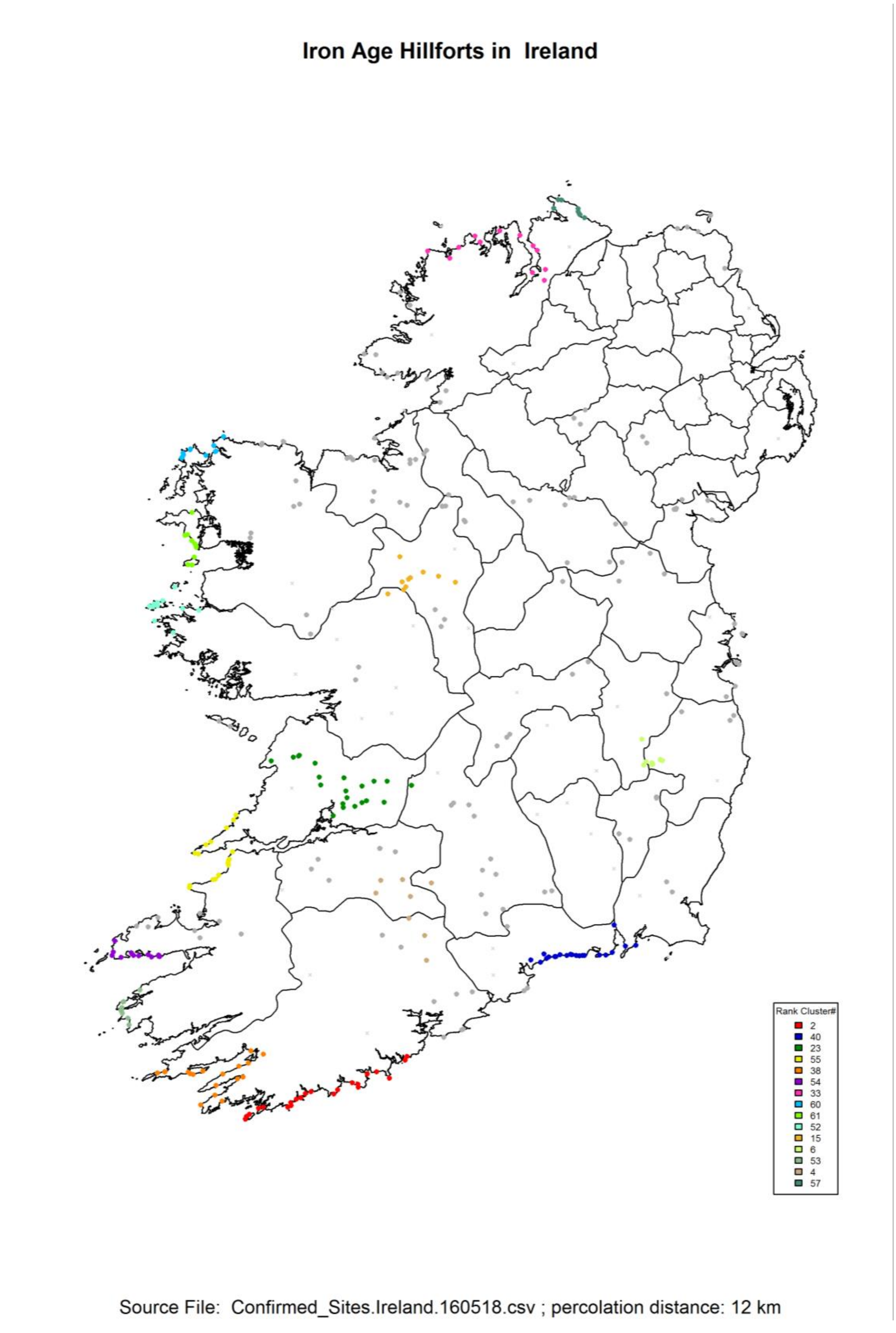


Figure 18 - Hillfort clusters in Ireland; 12 km percolation radius

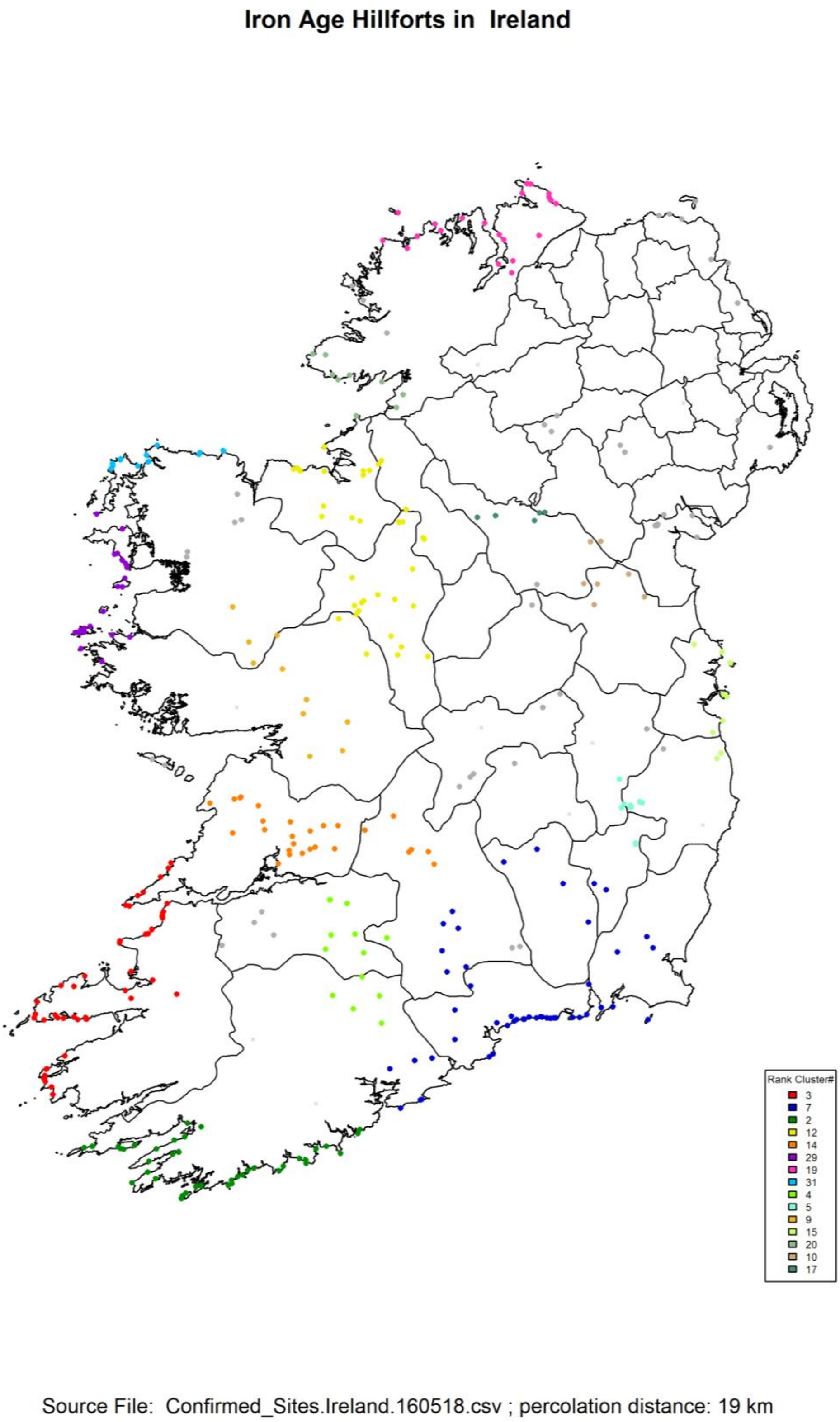


Figure 19 - Hillfort clusters in Ireland; 19km percolation radius

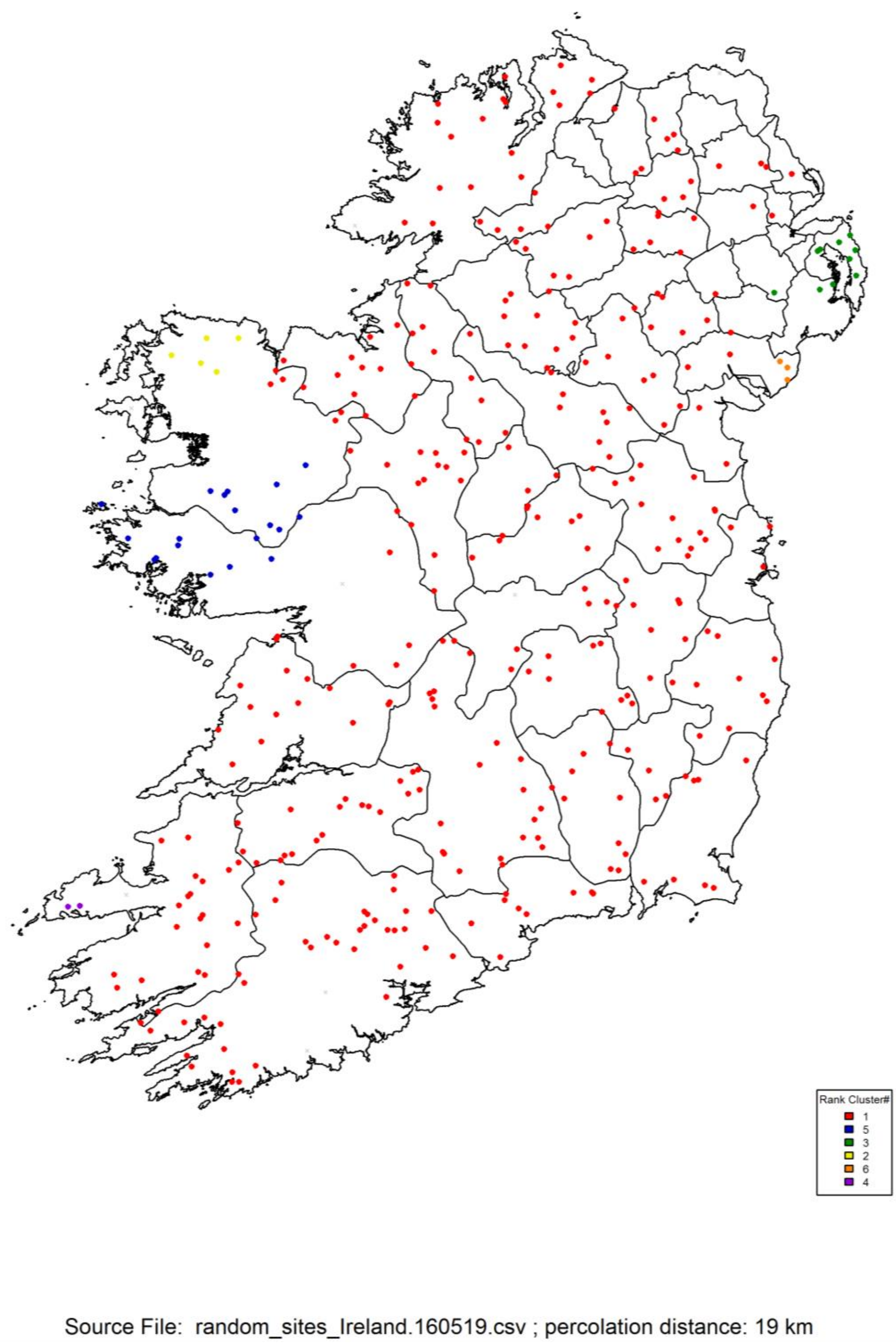


Figure 20 - Random sites in Ireland; 19km percolation radius

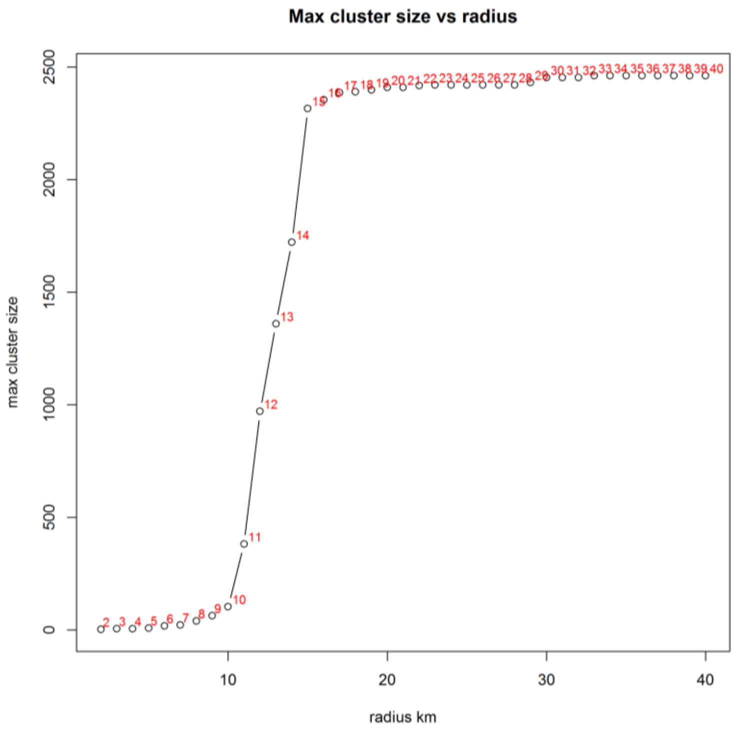


Figure 21 - Max cluster size vs. percolation radius, random sites in Britain

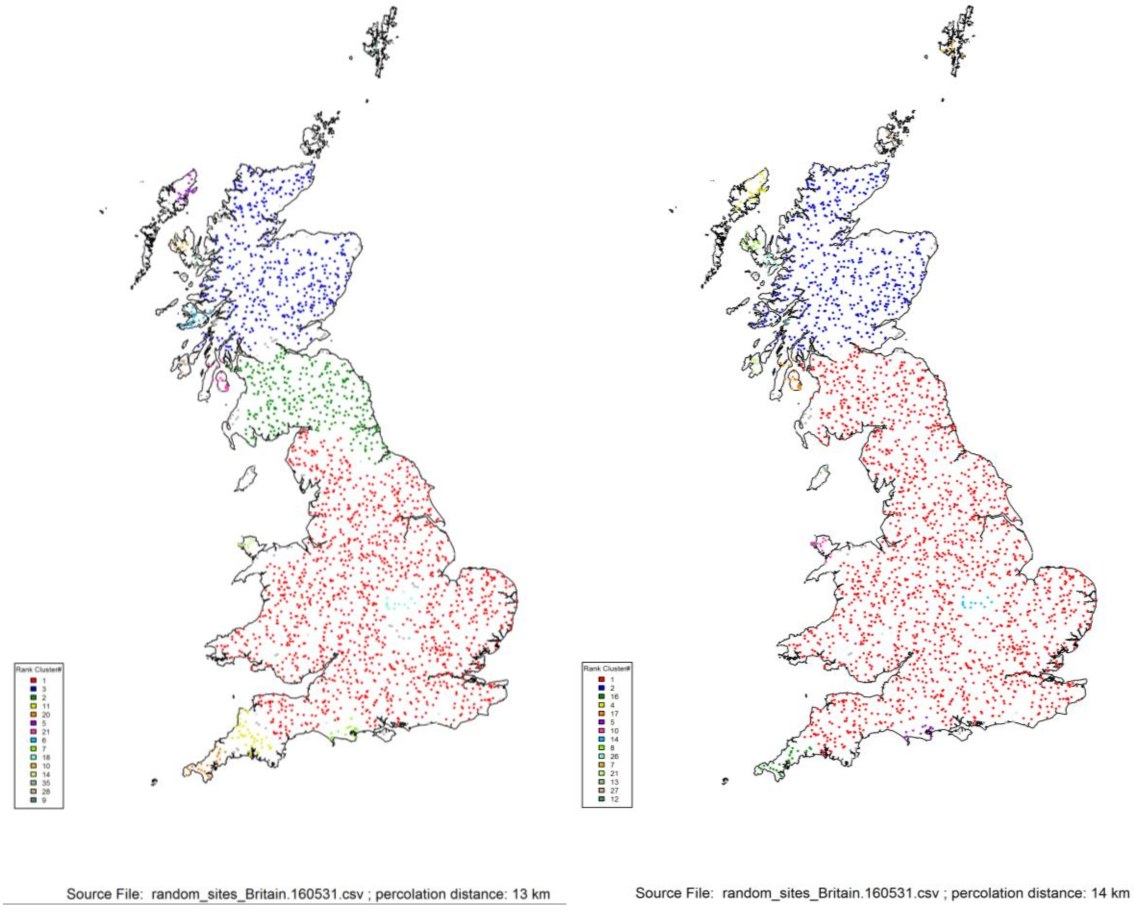


Figure 22 - Random sites in Britain; 13 and 14km percolation radii

Topographical Analysis

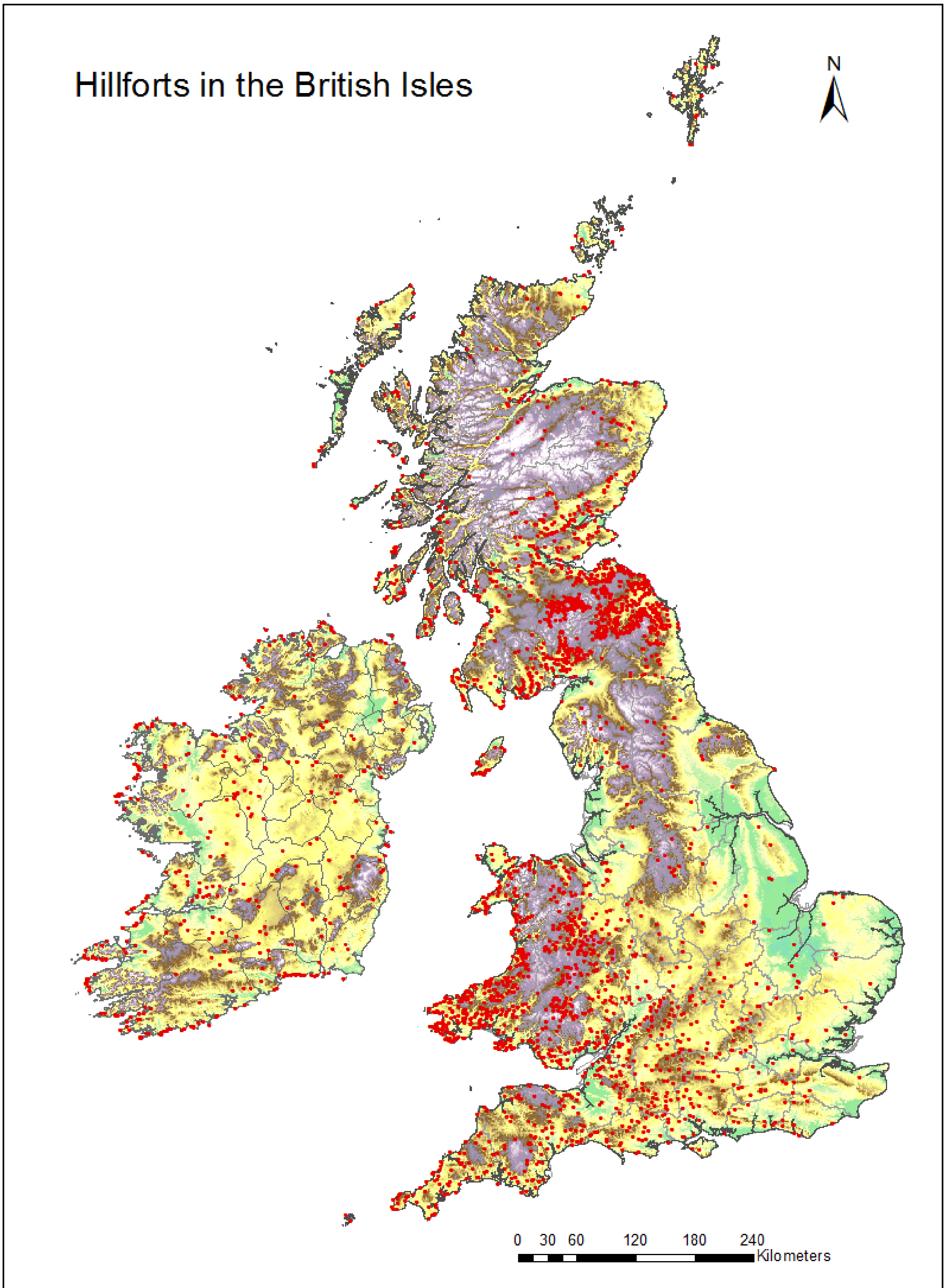


Figure 23 - Hillforts and the topography of the British Isles

Figure 23 shows hillforts in the British Isles, overlaid on a topographical map (regarding the colour scaling, see note in the following section). At a global level hillforts are generally located on higher ground, but not necessarily the highest, as can be seen in south-west England and Wales for example. It is also clear that there are significant hilly regions of Britain where there are few hillforts, notably the Pennines and the Scottish Highlands. In both these regions, the few hillforts are distributed around the hill zone periphery.

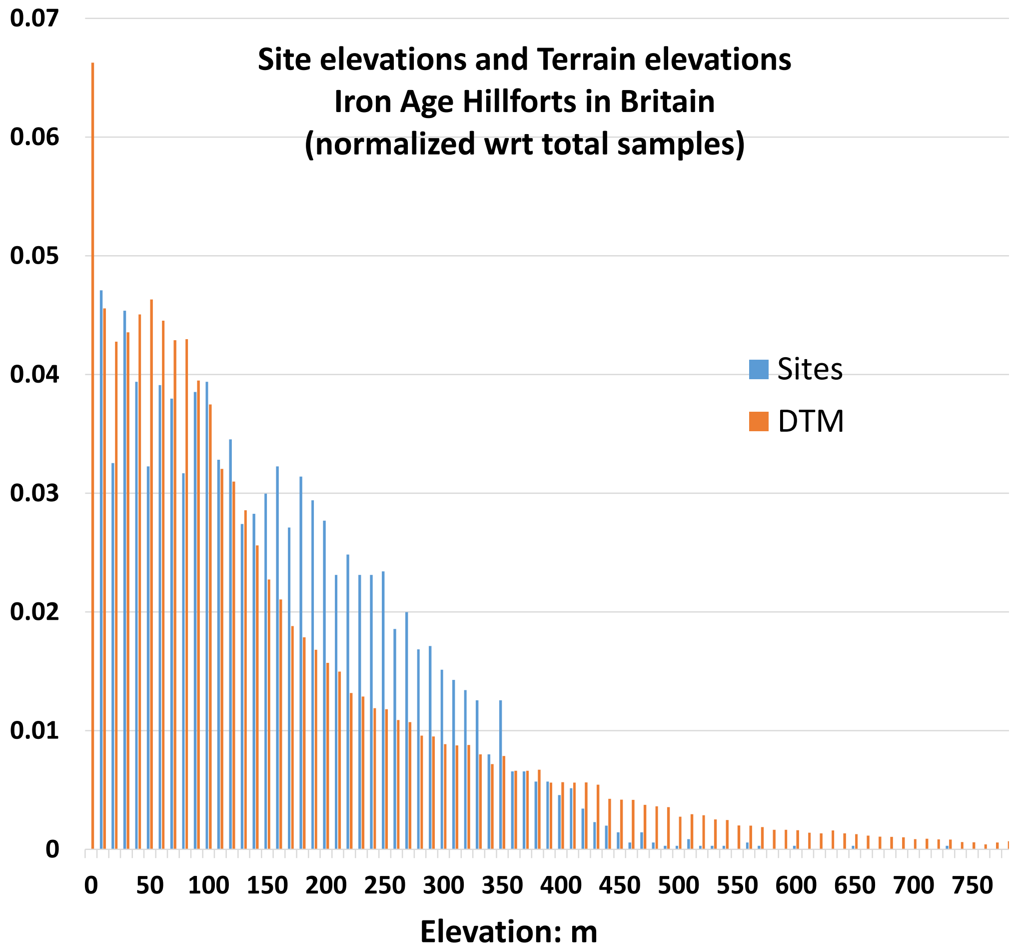


Figure 24 - Histogram of hillfort and terrain elevations in Britain (maximum site elevation is 722m, and terrain 1230m)

This global observation is reinforced by Figure 24, which is a histogram of hillfort and terrain elevations for Britain as a whole, truncated at an elevation of 750m. This shows that at the lowest elevations there are no hillforts, then up to about 150m they are at a similar if slightly lower density than the terrain at these levels. From 150m to 350m elevation, hillforts are above average elevation than the terrain profile; there is then a rapid tail off in sites from ca. 400m in altitude. The maximum elevation of a hillfort is 722m, whilst the maximum terrain elevation is 1230m. This apparent elevation limit for hillforts may be a first order factor in the paucity of sites in the

Highlands and the Pennines, although there must be other factors at work, possibly cultural, otherwise one would expect many more hillforts in the nearby valleys and foothills. Of course the upper elevation limit for hillforts itself requires explanation; likely causes would include the habitability of the sites, the agricultural value of the surrounding land and accessibility of the site to the wider community living in the surrounding terrain.

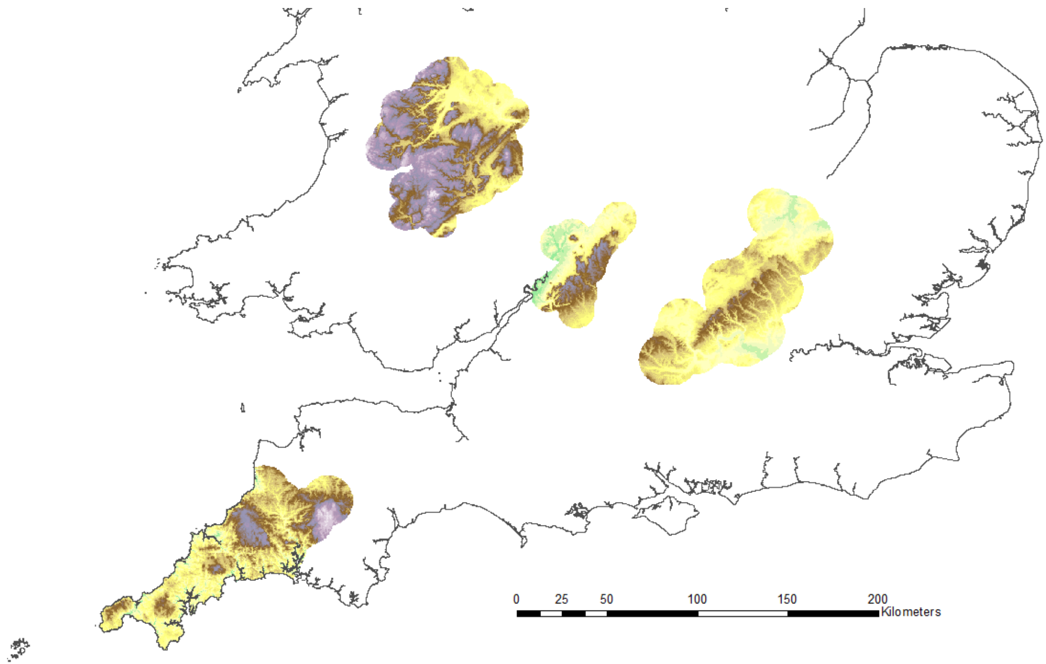


Figure 25 - Extracted cluster rasters

This analysis has been repeated at a regional scale for the four percolation clusters selected for the 'Discussion of Specific Regional Clusters' later. The terrain for each cluster was defined in ArcGIS using the relevant percolation radius to draw a buffer around each site in the cluster; these were then dissolved into a single polygon. The polygon was then used to cut a raster from the overall terrain model (see Figure 25). The elevation data for each cluster were then exported, and statistics applied for both terrain and hillforts using Microsoft Excel.

Figure 26 shows elevation histograms for Central Wales and the Marches, and Figure 27 for the Cotswolds. The terrain profiles reflect the regional character and the hillforts show a propensity to be located at higher elevations. In Figure 28 the Chilterns histogram is very similar to that for the Cotswolds, but Cornwall (Figure 29) is notably different. Whilst hillforts do not occur at lower elevations, they otherwise approximately follow the terrain profile, with a cut off at 260m. This may be partly explained by the large number of promontory forts, which are in their own way dominant in the landscape, but through their positioning rather than by elevation. The elevation cut off at 260m is not specifically limited by the terrain, as it is for the Chilterns (and the other two, albeit higher), but it seems likely that the particular local environment is different in its impact than for Britain globally and these other regions. As mentioned below, further investigation at a local level would be valuable to identify possible first and second order location factors.

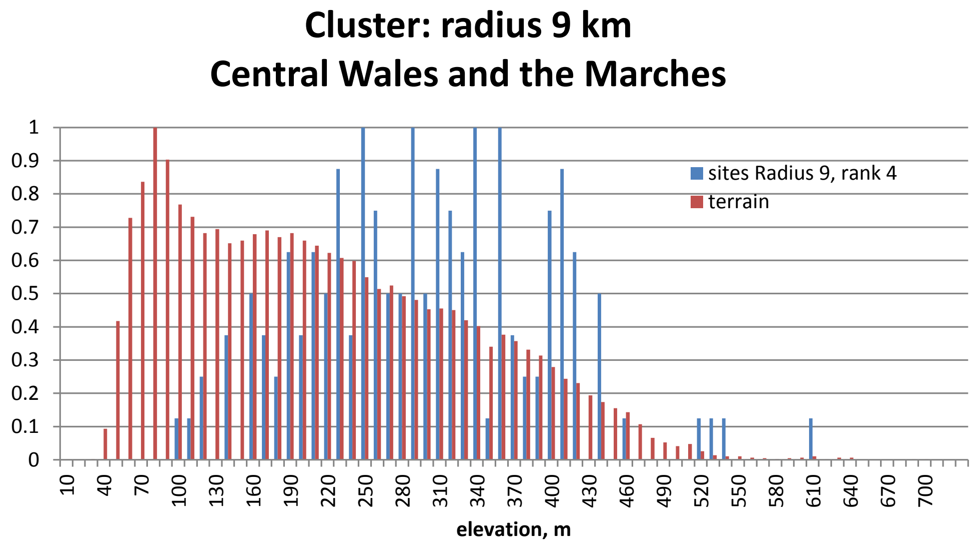


Figure 26 - Elevation histogram; Wales/Marches (normalised wrt total samples)

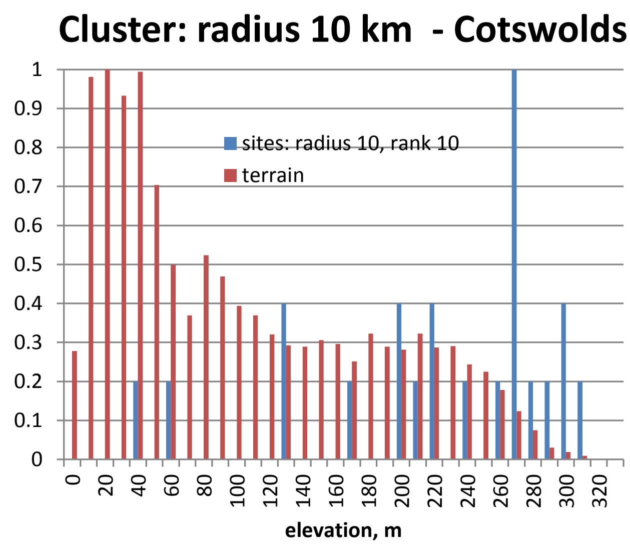


Figure 27 - Elevation histogram: Cotswolds (normalised wrt total samples)

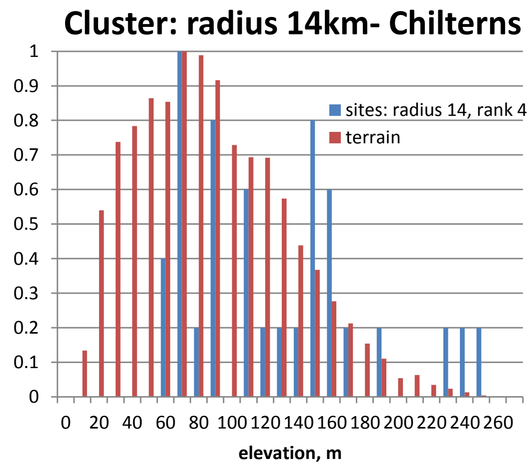


Figure 28 - Elevation histogram: Chilterns (normalised wrt total samples)

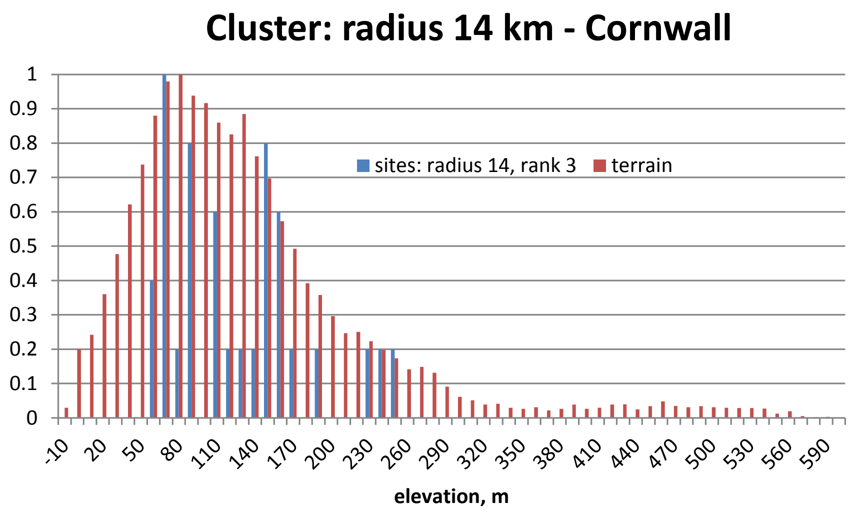


Figure 29 - Elevation histogram: Cornwall (normalised wrt total samples)

Whilst illuminating at a regional level, these histograms do not show the relation of hillforts at the most local level. A program was created in R to perform a comparison of hillfort elevation with respect to the surrounding terrain, intended to be used over a radial distance up to 1-2 kilometres. However whilst working effectively for the Isle of Man, it was far too inefficient for the very large file size of the terrain map of Britain, and the work was put to one side for lack of time. This would be worth pursuing, as well as experimenting with the 'prominence' approach and tools developed by Llobera (2001) to gain a better understanding of how hillforts are located in their immediate terrain.

Spatial Analysis by Attributes

Methodology

The Atlas database contains a wealth of information about its sites, as given in more detail in Appendix A. In terms of a more detailed analysis of selected clusters, a handful of key attributes were selected to see what they might offer in the way of providing material for further detailed study. The attributes chosen were those that could potentially offer a relatively simple way of classifying and comparing sites quantitatively:

- ***Enclosed area***: Defined as being the area enclosed within the innermost rampart. Several fields are available (mostly not completed) that also record whole site area, and the area between up to four widely spaced ramparts (as seen in the south-west for example). Data with respect to Annexes has been ignored. Enclosed area is arguably the simplest proxy for the importance of a site, without needing to determine specific function, as this reflects most clearly the amount of effort involved in building the Hillfort. It also gives a relative indication as to the number of people it can accommodate: in terms of living there permanently, semi-permanently or seasonally; for festivals either secular or religious, or some combination of both; for military rallies or assemblies; as a place of refuge or defence in times of conflict. Larger sites can also accommodate larger numbers of animals, either for seasonal bloodstock exchange, for protection against rustling or in times of conflict, or as part of a transhumance system.
- ***Ramparts:*** The number of ramparts is recorded for each of the four compass quadrants, along with information about the nature of construction, condition, preservation and the use of natural features. The challenge that this raised was generating a single numerical index that would reasonably reflect the vallation of the site, and this was not attempted.
- ***Ditches:*** The number of ditches is recorded as a single figure, unlike ramparts above, with a comment field. This enabled a simple attribute to be used that reasonably reflected the complexity and vallation of the enclosing structures of the site.
- ***Entrances:*** The number of original entrances. The database also provides for the number of extant entrances and breaks, and comments on these, as well as a field for the number of original entrances and a comment on this too. Entrances might reflect the level of activity of the site, in terms of people, livestock and goods entering and exiting. However identifying original entrances is not always possible without detailed survey and/or excavation, and there may be considerable uncertainty in distinguishing between modern and original entrances as well as the effects of later damage.

All of the above attributes reflect the final state of site construction of course, and will have likely altered through time.

Other key data was also extracted, to provide a readily accessible source of information without recourse to the principal database during analyses and mapping. These parameters were the Atlas index, the site name, country and county, the description of the site, date and aspect. Fundamental of course was the coordinate information.

This data was exported from the primary spreadsheet as a csv file, which was used within analysis programs in R, and also imported into the ArcGIS program package, version 10. The DEM models for Ireland, mainland Britain and the Isle of Man were imported and configured with a uniform elevation colour scheme to illustrate the topography. The scale of colour scheme was the same for all three, with a common elevation maximum. The colour scheme was deliberately selected to highlight the terrain, and used 32 levels with default quantile scaling, i.e. non linear. Shape files of the coastline were also imported to define the outlines. Counties are included in Ireland as a coastline only graphic could not be found. Counties are also included in some plots for Britain where it was a useful aid to regional identification. These outlines were the same as those used in the computational analyses written in R.

Site locations were exported as shape files from the R program suite, and imported into ArcGIS. Files identifying cluster membership and various rankings (see Appendix C) were also exported as csv files from the R programs. These were then imported into ArcGIS and associated with the site locations. It was then possible to selectively plot sites, on the basis of cluster membership for example, and to apply symbology scaled using numerical attributes.

Initial experimentation was conducted using ArcGIS to plot sites with a symbology using respectively the enclosed area, the number of ditches and the number of original entrances. The aim of this empirical approach was to generate patterns with potential for detailed study. Examples of these plots are shown in Figure 30 to Figure 35, in each case for the British Isles as a whole, and for more detail, Wales and south-west England.

For ditches and entrances a simple scaling of three categories was used based on the attribute value in the Atlas data. For ditches the thresholds were 2, 3 and >3, and for entrances 1, 2 and >2. For enclosed area the range of values was a large continuum, up to a maximum of over 80ha in Britain and 130ha in Ireland. Some experimentation was done, using the various symbology options available in ArcGIS (e.g. Jenks, equal area), also referring to the histograms of enclosed area (Figure 36, Figure 37) for guidance in terms of ‘natural breaks’ in the size. The same symbol levels were configured for all sites, and the values that worked best in terms of highlighting the pattern of bigger sites overall was to use thresholds at 4ha and 16ha, with the choice of the upper threshold being the more important in terms of patterning. As noted in 'A review of previous studies' above, the levels in published distribution maps, e.g. Hogg (1975, 38, 40, 42) are derived from the scale used by Rivet in the Iron Age OS map (1962), and revisiting these for comparison would be an interesting area for future work.

It is worth considering for a moment what these values mean in terms of size and scale of a site. Assuming the sites are circular, then the length of just one rampart and ditch for a small site of 0.2ha and sites at the two threshold values are as follows:

- 0.2ha – 159 metres
- 4ha – 709 metres
- 16ha – 1418 metres

The largest site in Britain is Ham Hill (Atlas: 0448) at 84ha with three ramparts and a perimeter of 3.3 km and in Ireland the largest enclosure is at Spinans Hill of 131ha with two ramparts and a perimeter of 4 km (Atlas: 0727). Even for a single rampart the larger sites represent massive works of construction.

As discussed earlier for percolation clusters, these maps are not in any sense an explanation of particular distribution patterns but rather they are descriptive and used to highlight regions which can then be explored in terms of other data sources for possible explanation.

Comparing the maps for enclosed area (Figure 30, Figure 31), ditches (Figure 32, Figure 33) and entrances (Figure 34, Figure 35), in all cases Ireland shows a distributed pattern for the largest classification. Britain shows regional patterns for both ditches and area, but the plot for entrances is not particularly clear. For the reasons described above regarding the challenges of the data, entrances were excluded, and although ditch based patterns could be interesting to explore in the future, it was decided to focus on enclosed area as showing not only distinctive regional patterns, but interesting patterns within selected clusters.

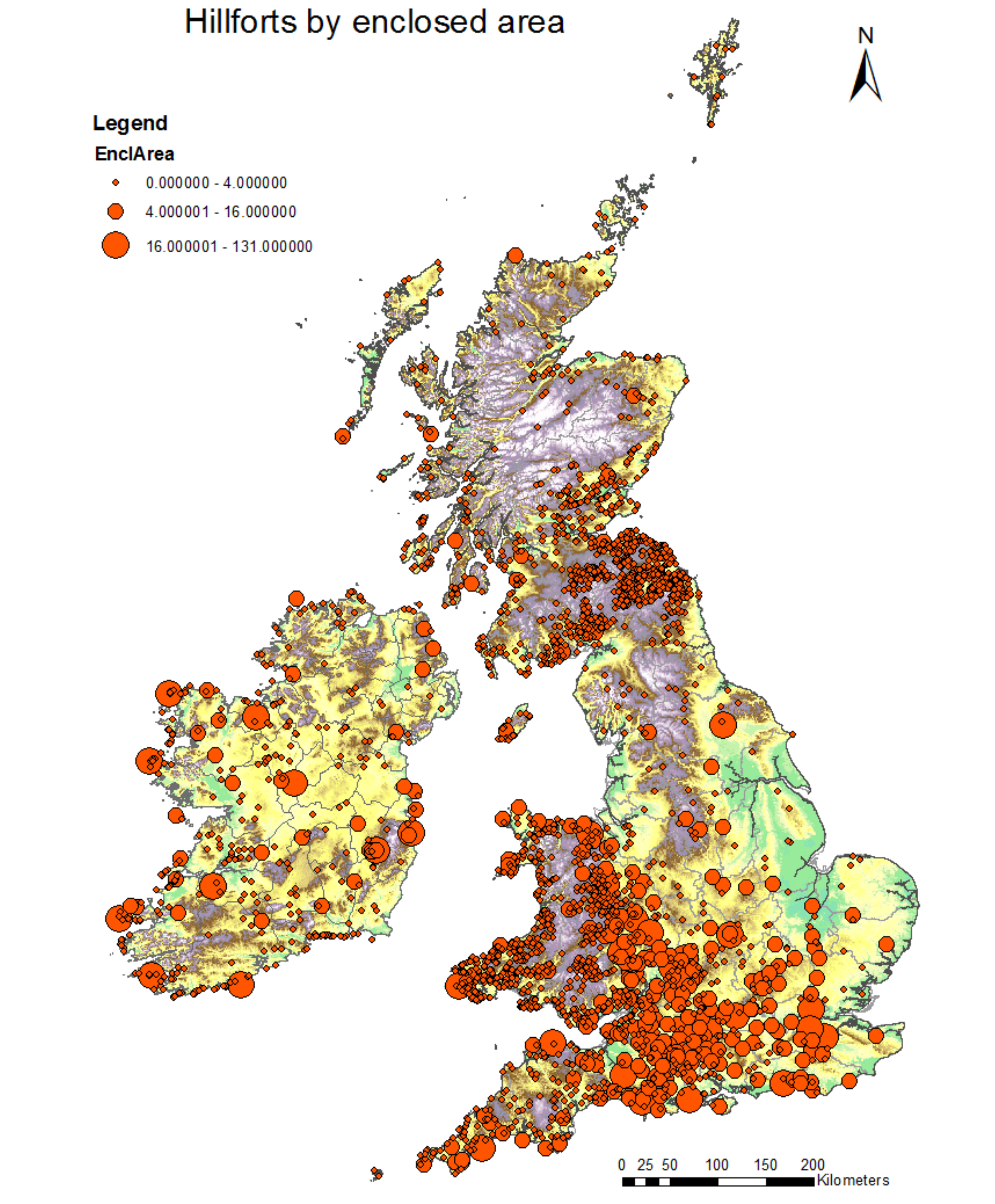


Figure 30 - Hillforts in the British Isles, by Enclosed Area

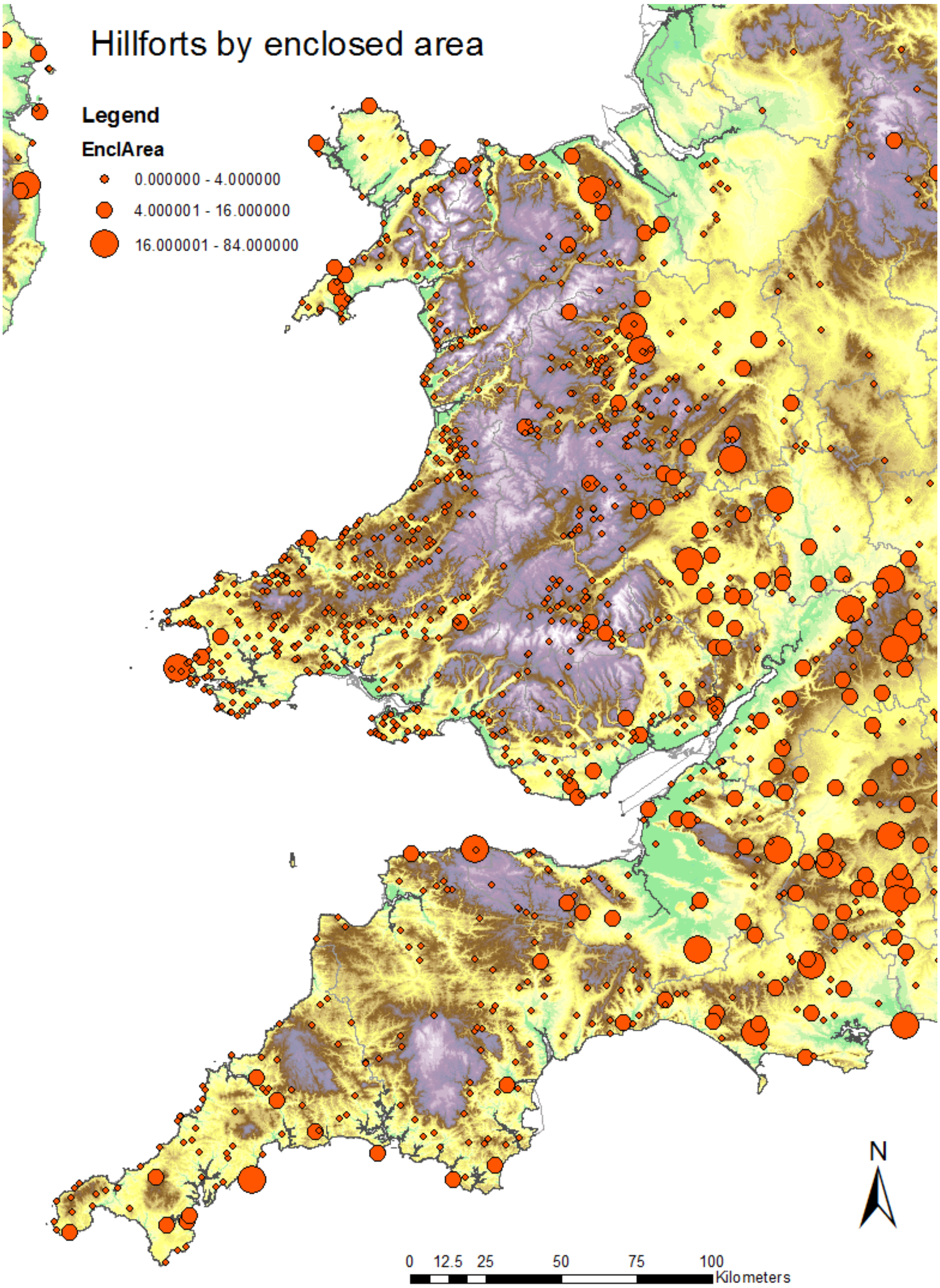


Figure 31 - Wales and south-west England by Enclosed Area

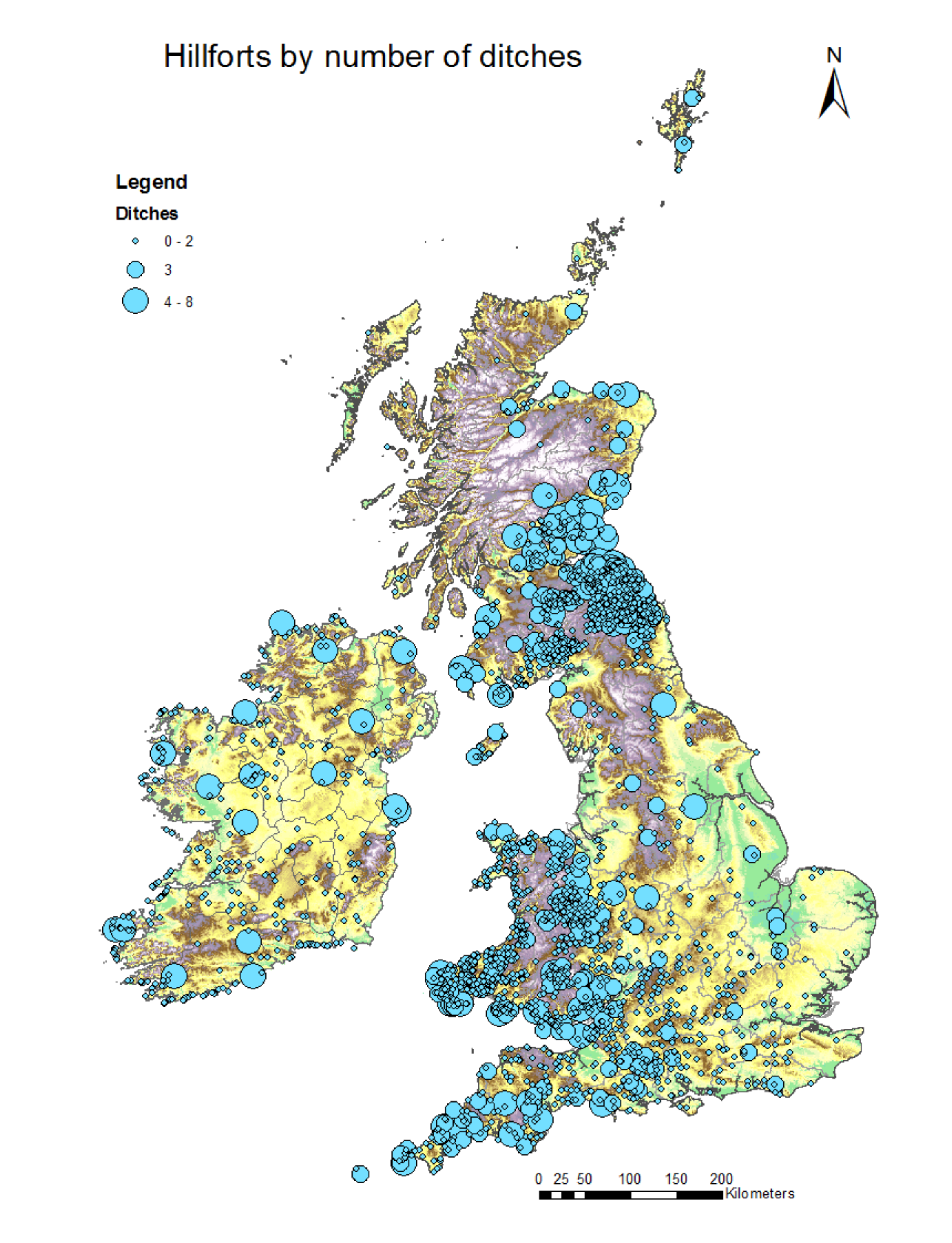


Figure 32 - Hillforts in the British Isles, by number of Ditches

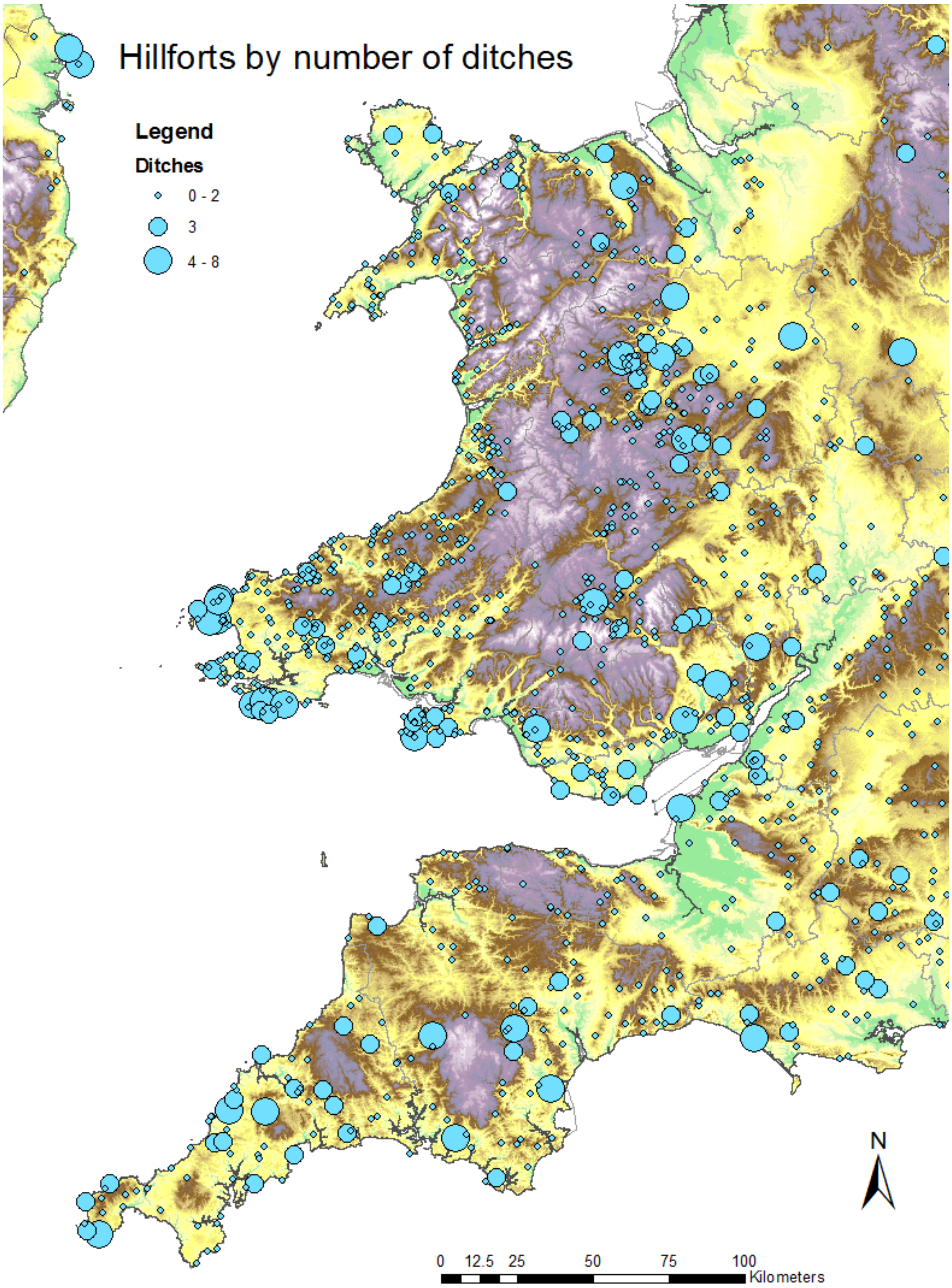


Figure 33 - Wales and south-west England by number of Ditches

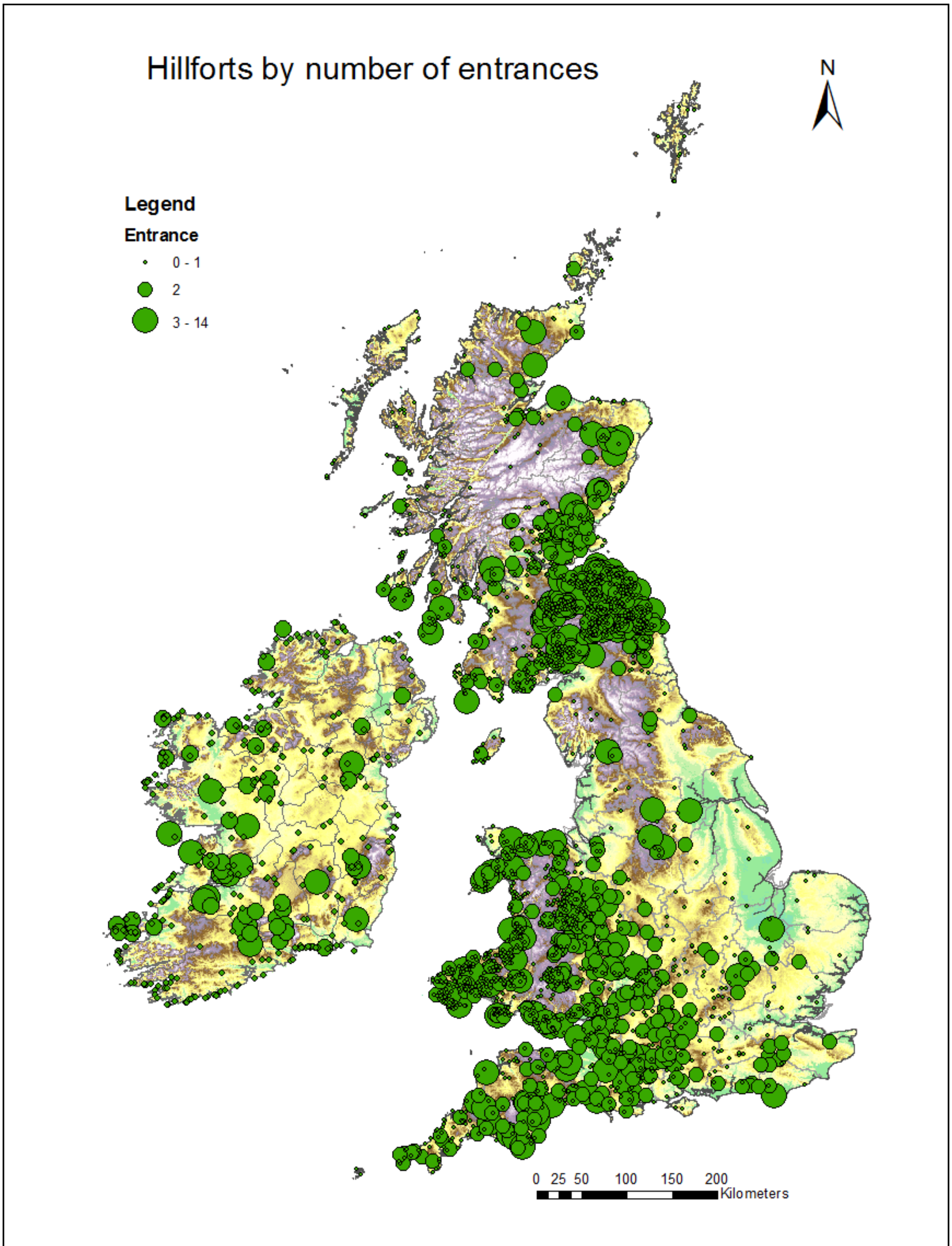


Figure 34 - Hillforts in the British Isles, by original Entrances

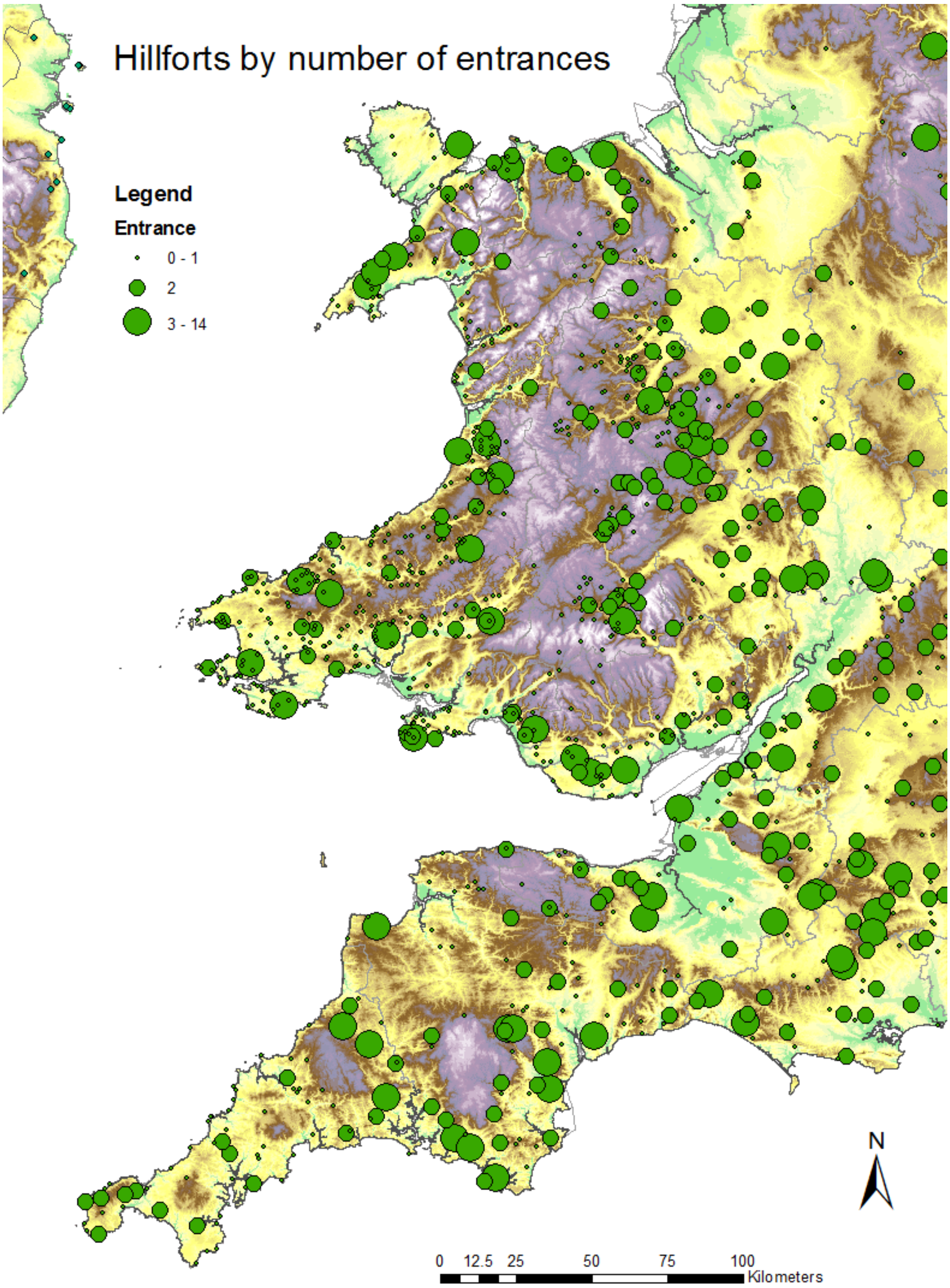


Figure 35 - Wales and south-west England by original Entrances

Analysis by Enclosed Area

Figure 30 shows Hillforts in the British Isles by enclosed area, with symbols for three ranges of size, with thresholds at areas of 4 and 16ha. The same scheme is applied uniformly in Britain and Ireland. As noted earlier, these thresholds are different from those used by Rivet in the OS Map of Southern Britain in the Iron Age (1962) and subsequently by many others, for example Hogg (1975) and Lynch, Aldhouse-Green and Davies (2000), but as Rivet (1958) himself said, those values were arrived at after some experimentation, and that is how the values used here have been established, probably with considerably less effort.

Histograms of enclosed area for sites in England, Ireland, Scotland and Wales are shown in Figure 36 and Figure 37, based on current national boundaries, as these parameters are readily extracted from the database; the numbers of sites for each are shown in Table 1. The histograms have been computed and rendered graphically in Microsoft Excel. The histogram in Figure 36 uses a logarithmic vertical scale in order that the sites on the 'long tail' are more apparent. The plot in Figure 37 uses a linear scale, and only goes up to 31ha to show more detail.

Country (modern)	Number of Hillforts
Scotland	1027
England	809
Wales	621
Ireland	390
Isle of Man	22

Table 1- Number of Hillforts by Country

Rank-size plots have been created for the same data, shown in Figure 38 and Figure 39 (Hodder and Orton 1976, 69-73). This has also been carried out using Microsoft Excel. The sites are sorted into rank order by area, and plotted on a log-log graph of area vs. rank. The 'log-normal' line is also plotted between the largest and the smallest site, to aid classification (Drennan and Peterson 2004). All four graphs are clearly convex, suggesting a non-primate structure, that is to say a structure where there is not a dominance of one or a small number of very large sites, but rather there is a group of larger sites with similar scales of size.

To quantify the degree of variation from a log-normal form, the 'A' coefficient described in Drennan and Peterson (2004) has also been computed. The area under the curve (a_C) is calculated by summing the trapezoidal area defined by successive point pairs. The rectangular area (a_R), defined by the largest and smallest sites, is calculated directly from the log-log data (i.e. the difference between the largest and the smallest values, taking account of sign, times the maximum rank), and the result divided by 2 to give a normalisation factor.

The Drennan and Peterson 'A' coefficient is:

$$A = \frac{a_C}{a_R}.2 - 1$$

A log-normal curve would give a value of 0, convex >0 and primate < 0. The results are shown in Table 2.

Country (modern)	'A' coefficient
Scotland	0.305
Wales	0.289
Ireland	0.255
England	0.488

Table 2 - 'A' coefficients by Country

The value of 0.488 for England reflects the larger proportion of bigger sites than Scotland and Wales particularly. These latter two have strikingly similar forms (Figure 38), with a trend line for the mid to small-range sites, and the larger sizes fall below this, starting approximately at 12 and 7ha respectively, and only 8 or 9 sites larger than this in each case. The forms for Ireland and England (Figure 39) are also similar to one another, with less well fitting trend lines for the mid-range sites. It is suggested in Hodder and Orton (1976, 69-70) that these convex forms are the result of older more complex societies, with many factors influencing the development of sites over a long period of time. They also caution that the sites should be in contemporary use for this kind of analysis, and of course in this study, dating information is very limited and has not been applied.

What could explain these differences? As already noted the scale of the larger sites in Ireland and England in particular are significantly greater than in Scotland and Wales. It is also important to note that rank size analysis is applied in Geography for settlements and populations, and it is far from clear what proportion of hillforts were used as settlements, and for those that were, what the nature of that settlement might be (e.g. Harding (2012), Cunliffe (1991)), so the factors driving the size of hillforts may be rather different than for settlements, at least for some of them. From a topographical perspective, hillforts are generally located in dominant places, frequently making use of natural features such as contour lines on hill tops or promontories. There is thus also a topographical constraint on how large hillforts can become if they are to follow this basic driver, highly dependent of course upon the particular terrain.

A similar argument might apply at the lower end, where there the graph drops down, also noted in Hodder and Orton (1976, 69-73), with the sizes of smaller sites being driven by differences in function and constraints of both topography and perhaps more importantly, the resources needed to construct.

Looking at the maps with site symbols based on area, Ireland (Figure 40) has a particularly striking pattern, with its largest hillfort at 131ha (Spinans Hill 2, Atlas: 0727), but most notably with the arrangement of the larger and largest sites relatively uniformly distributed around the country. This is potentially suggestive of a regional role played by these sites, in a pattern common across much of the island.

Scotland (Figure 41) has a large number of sites (Table 1), but nothing in excess of 16ha, and only 8 bigger than 4ha, spread mainly along the south-west and west coast and islands. Wales (Figure 45) is similar in having a large number of small sites, with a few larger sites in the north-east, and a very small number of large sites, one at the tip of Pembrokeshire (Wooltack Point, Atlas: 1997)), and the other Pen-y-cloddiau (Atlas: 1155) in the Clwydian range. There are also two large sites on the borders, Llanymynech Hill (Atlas: 0071) and Y Breiddin (Atlas: 1276), which will be returned to in more detail later.

Although England (Figure 42, Figure 43 Figure 44) has a smaller number of sites in total (see Table 1) it has proportionally many more mid and large sites than the other regions. These are predominantly along the ridges and escarpments of the hills in south central England, as described by Cunliffe's eponymous 'Hillfort dominated zone' (Cunliffe 1991, Figure 20.6 page 542). The south-west (Figure 43) has a preponderance of promontory forts (also described as 'cliff-castles'), and only two large sites: Dodman Castle (Atlas: 0643) in south Cornwall, and Countisbury Castle (Atlas: 3459) in north Devon. The Cornish site will be discussed later.

Apart from the far north-east it is striking how few hillforts there are in northern England (Figure 44), with only a few along the Pennines, and a notable gap to the south of these hills, largely along the line of the Trent and continuing up to the Mersey. In the far north-east, there is no discernible break in density from south-east Scotland over the modern border, with a taper as it runs down to the line of the Tyne and a break along the course of Hadrian's Wall. This is itself an interesting reflection on the interactions between history and geography (for example see Morris (2010, 26-30)). Furthermore this break boundary matches partially with the clustering observed for randomly distributed sites in Figure 22, discussed earlier.

There is otherwise one large site of 24.5ha at Roulston Scar (Atlas: 1504) on the south-western end of the North York Moors. There are five further sites over 4ha in the Pennines: Ingleborough (Atlas: 2554) in the north-west; Barwick-in-Elmet (Atlas: 1597) between the Wharfe and the Aire;

Markland Grips (Atlas: 2970) in the south-east; Gardom's Edge (Atlas: 3364) above the Derwent and the Chatsworth Estate; and Mam Tor (Atlas: 2968) near the head of Edale. The locations of these sites is consistent with their playing a role in the transfer of goods between catchments and domination of passes through obstacles as suggested for hillforts by Sherratt (1996). However, the nature of other sites and settlements is clearly very different to other hillfort dominated areas of Britain.

All sites over 16ha in the British Isles are shown in Figure 46, which brings out the points raised above, notably the relatively even distribution of the larger sites in Ireland, their total absence in Scotland, a series along the eastern boundary of Wales, and the concentration along the ridges of Wessex and the Cotswolds.

One final observation about the location of hillforts in England and Wales which is particularly clear in Figure 30, is how different this distribution is compared to the classic views of the distribution of wealth, with a line traditionally along the Fosse Way roughly between the Severn and the Humber, dividing the economic zones of the south-east and the north and west, as illustrated for example in Figure 47, from Arcaute, Molinero, Hatna et al. (2016, Figure 3), this will be discussed later. The reasons for this difference could be the changing patterns of trade and consequential wealth through time, as the result of changing technology and other factors in the later Iron Age, as eloquently articulated by Sherratt (1996).

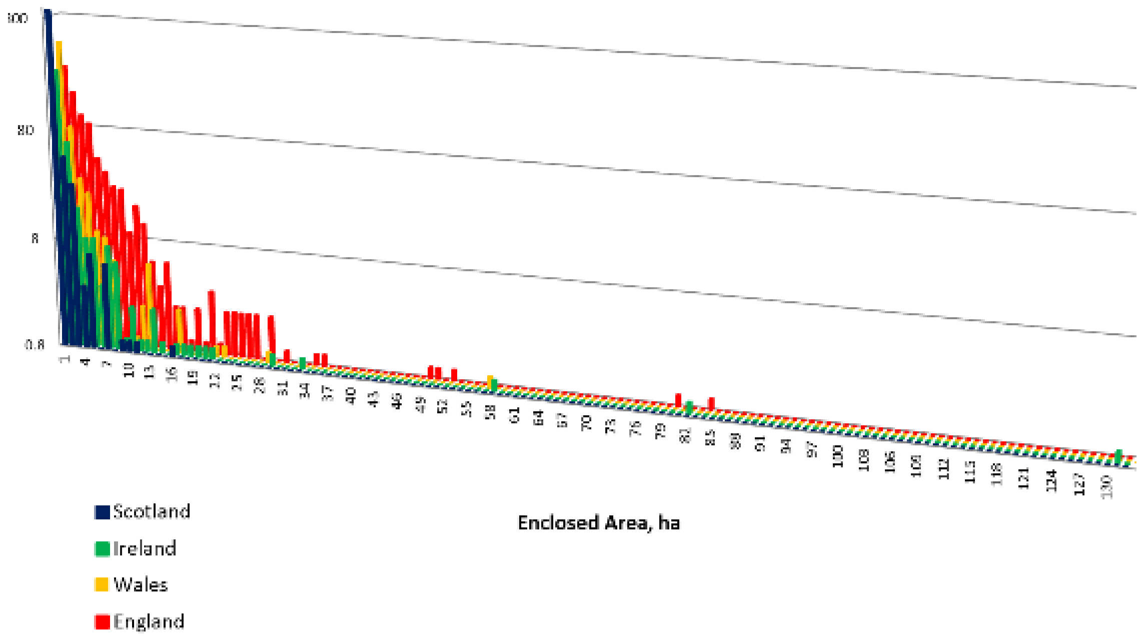


Figure 36 - Enclosed Area Histogram, by Country (log scale)

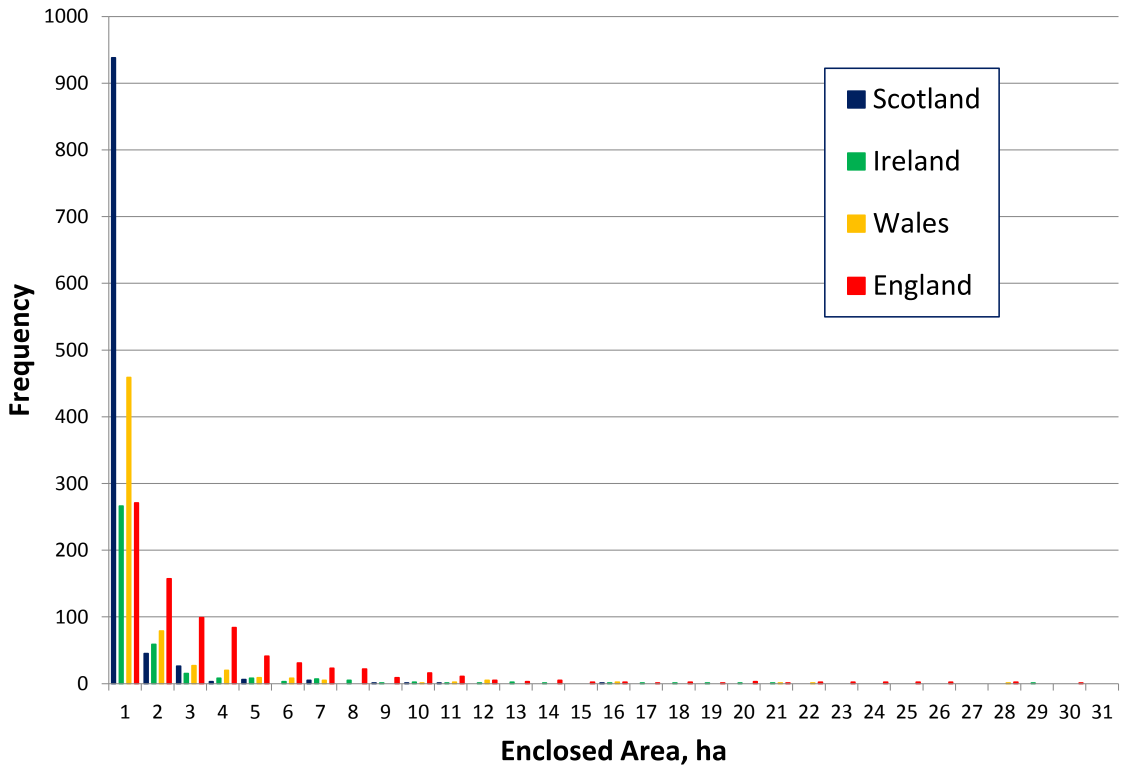


Figure 37 - Enclosed Area Histogram, by Country - to 31 ha

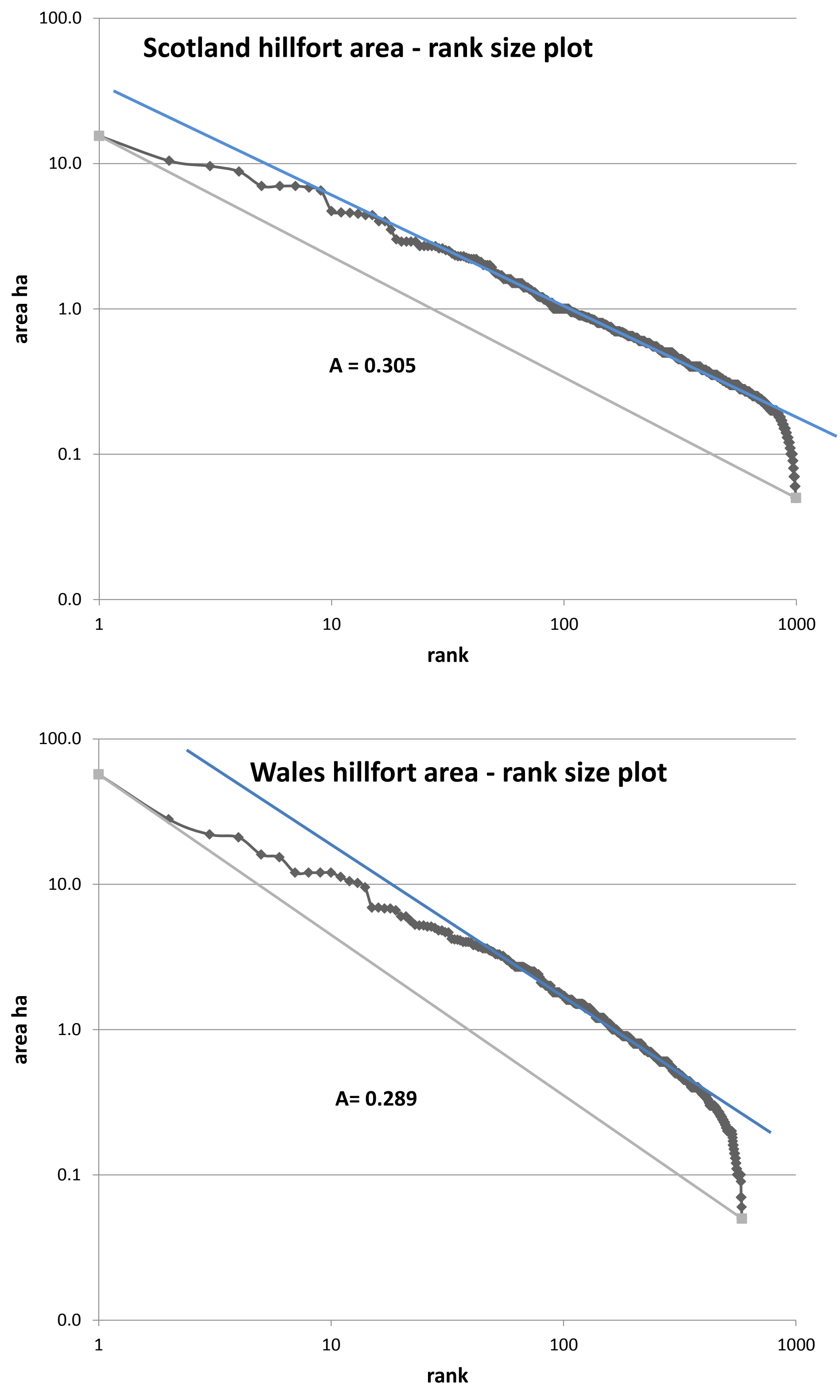
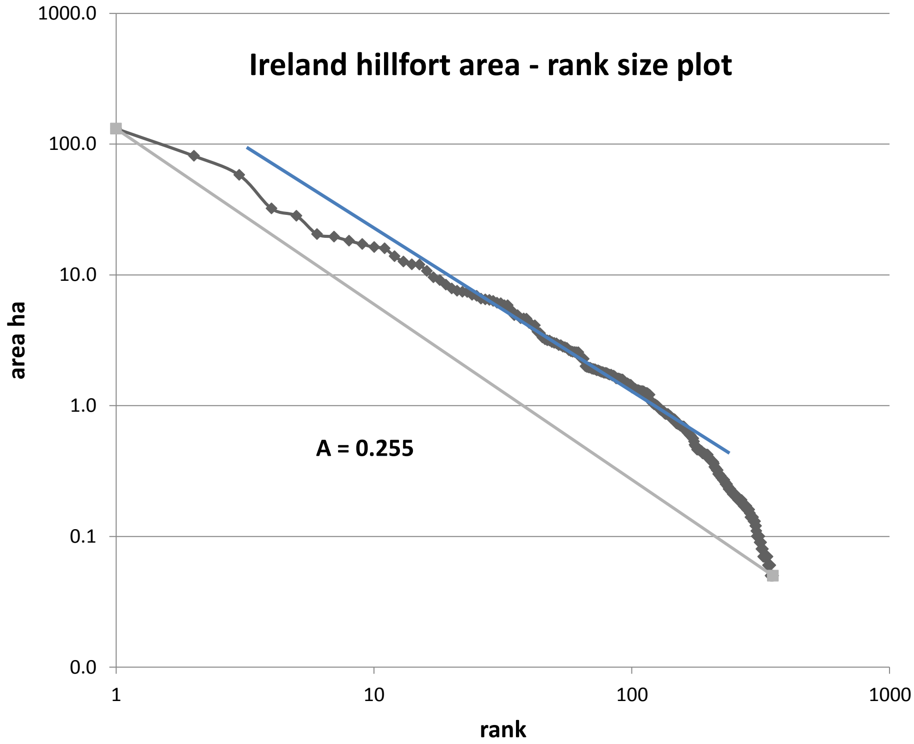


Figure 38 - Rank size plots for Scotland and Wales; blue trend lines for mid-range sites



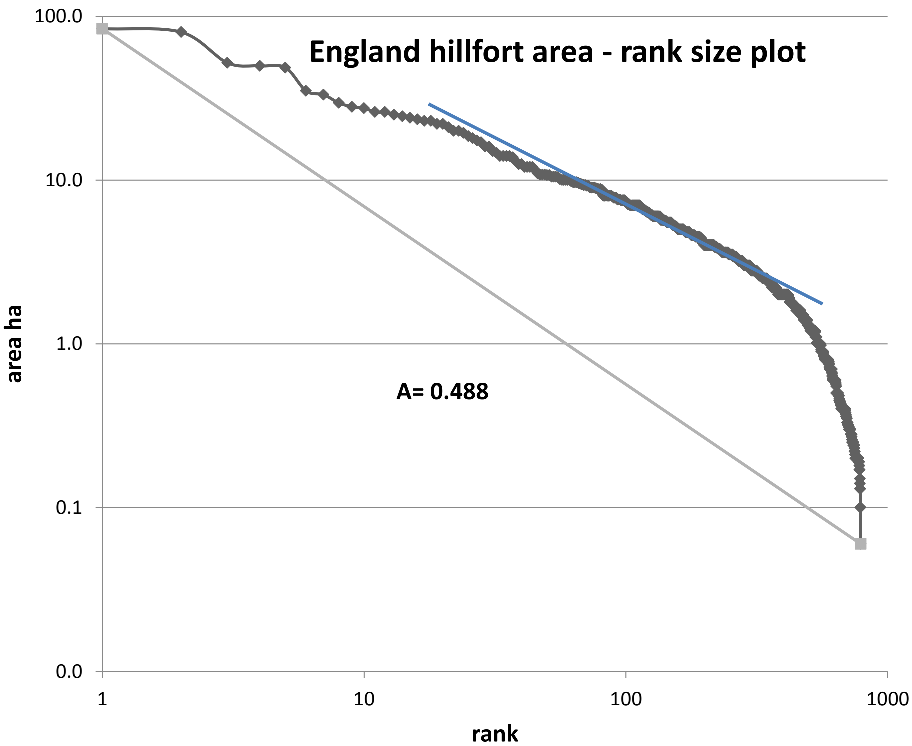


Figure 39 - Hillfort area rank size plots for Ireland and England; blue trend lines for mid-range sites

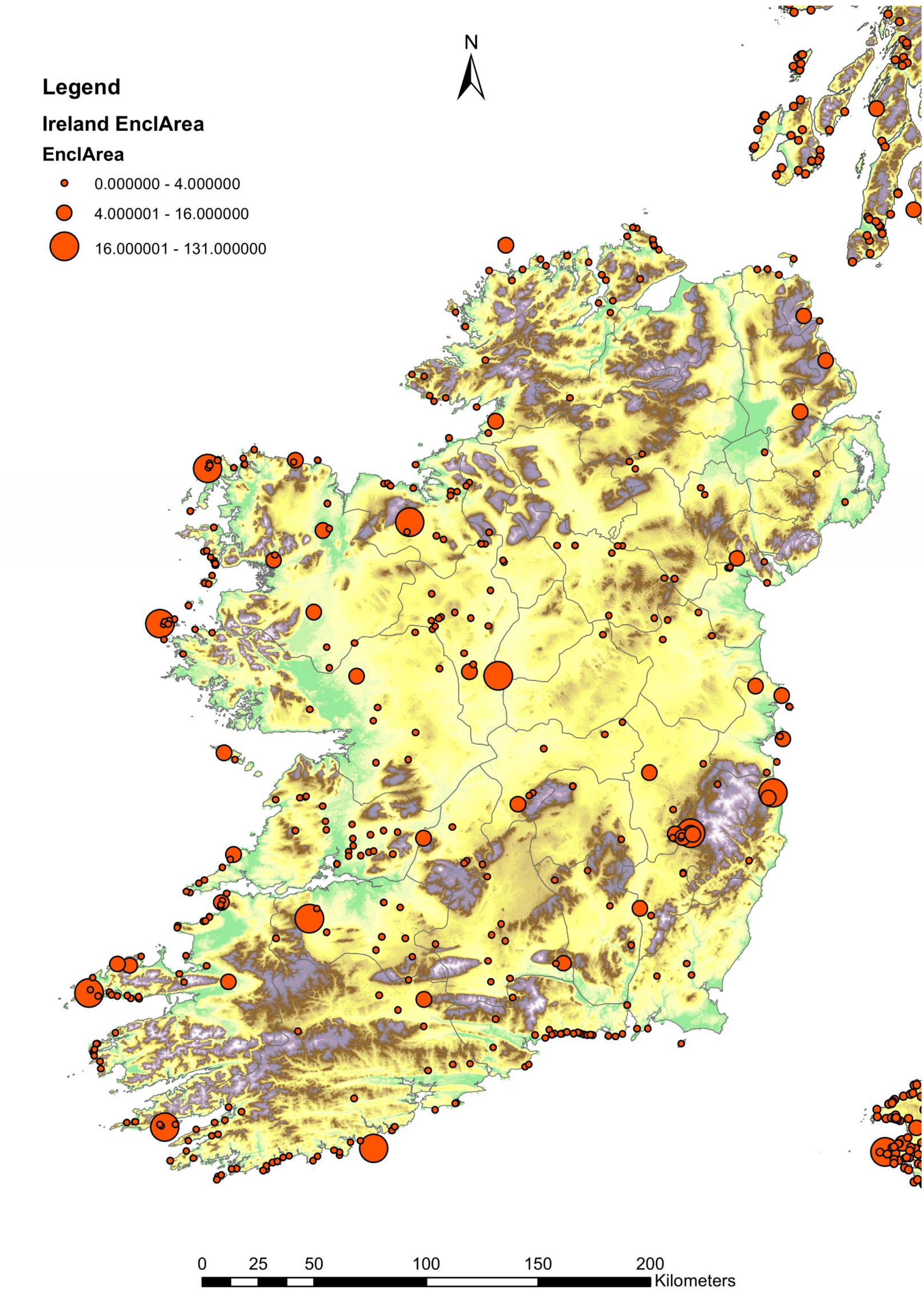


Figure 40 - Hillforts in Ireland, by Enclosed Area

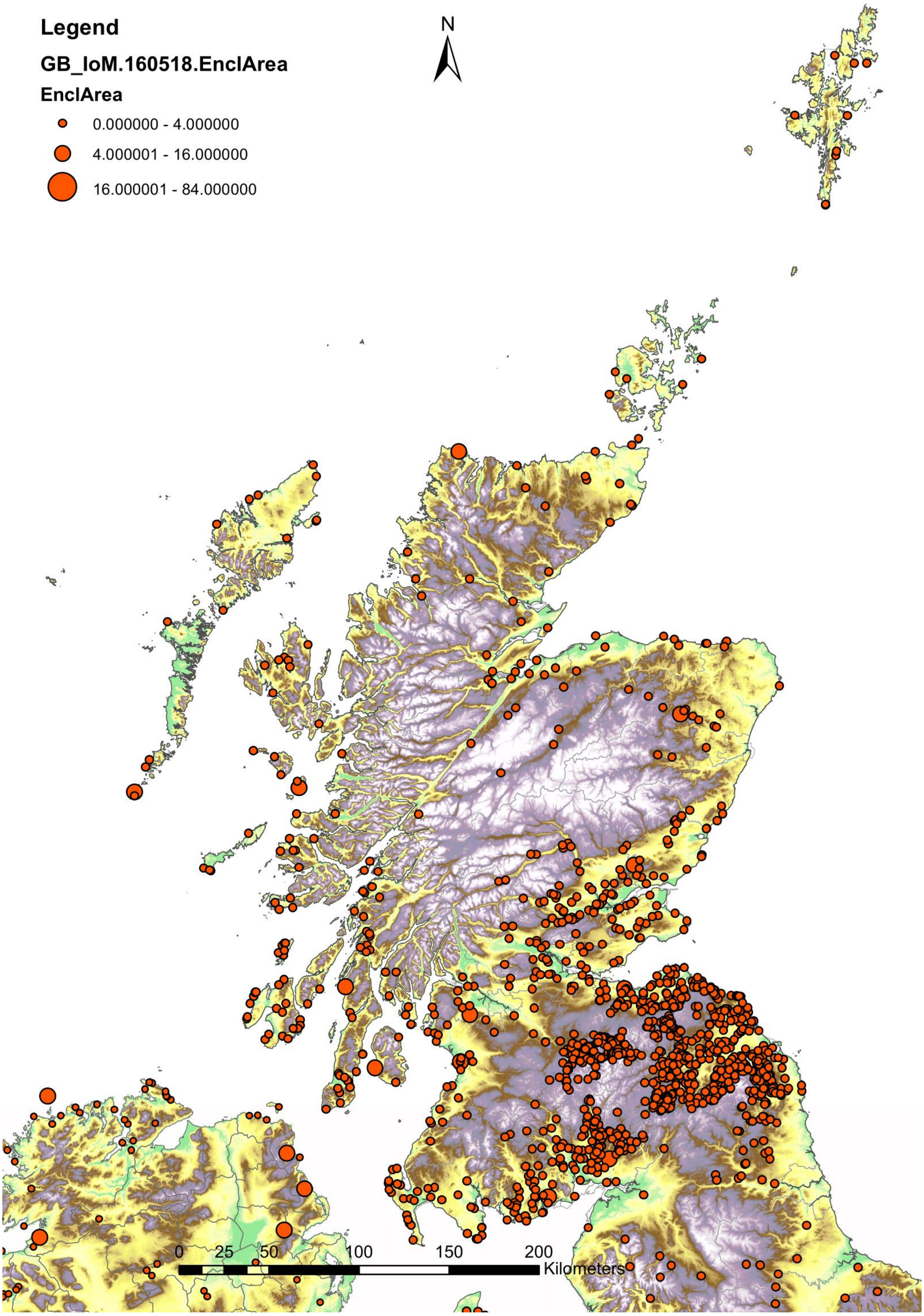


Figure 41 - Hillforts in Scotland, by Enclosed Area

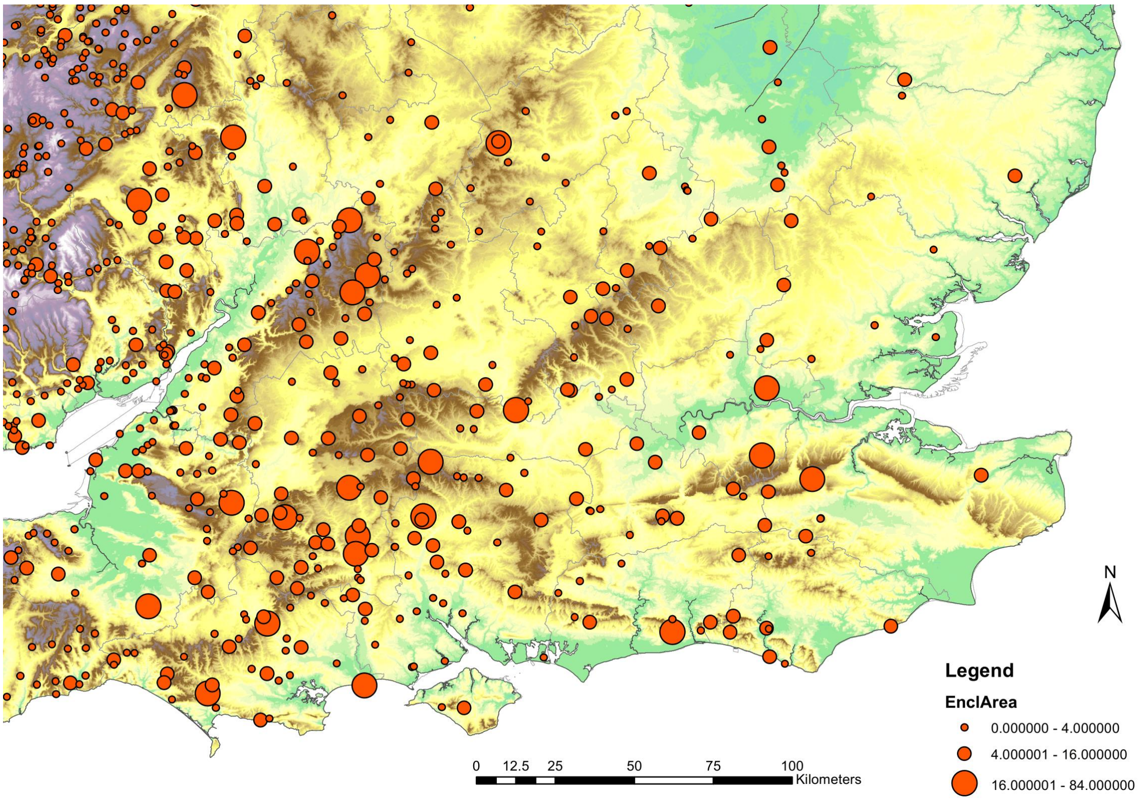


Figure 42 - Hillforts in central & southern England, by Enclosed Area

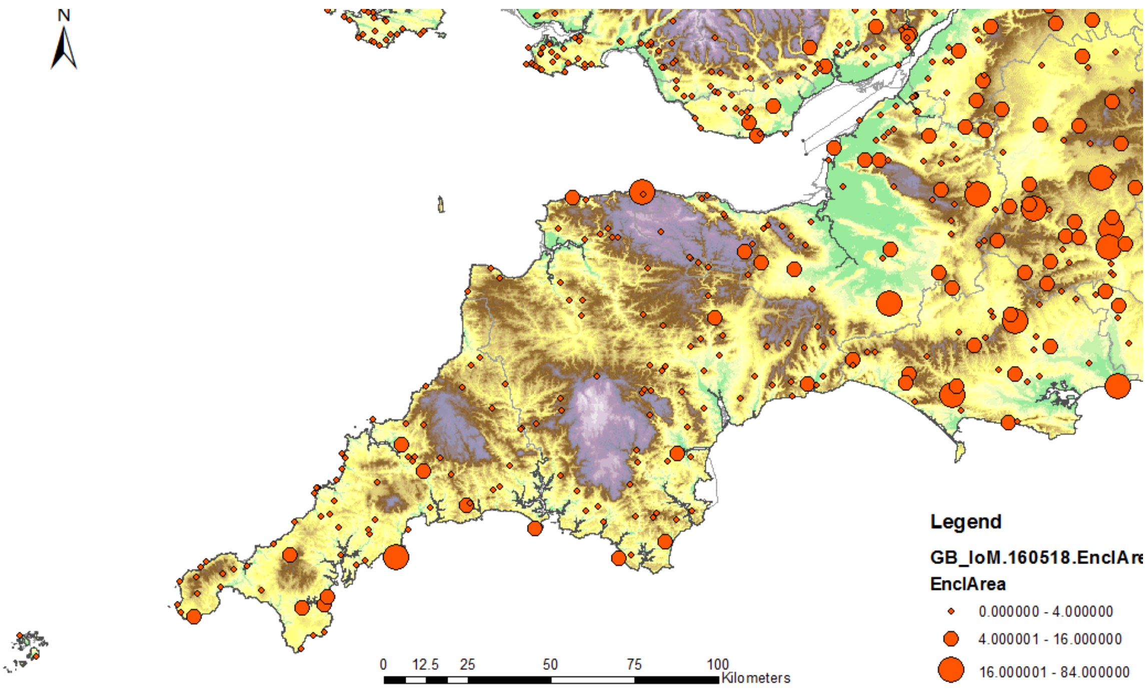


Figure 43 - Hillforts in south-west England, by Area

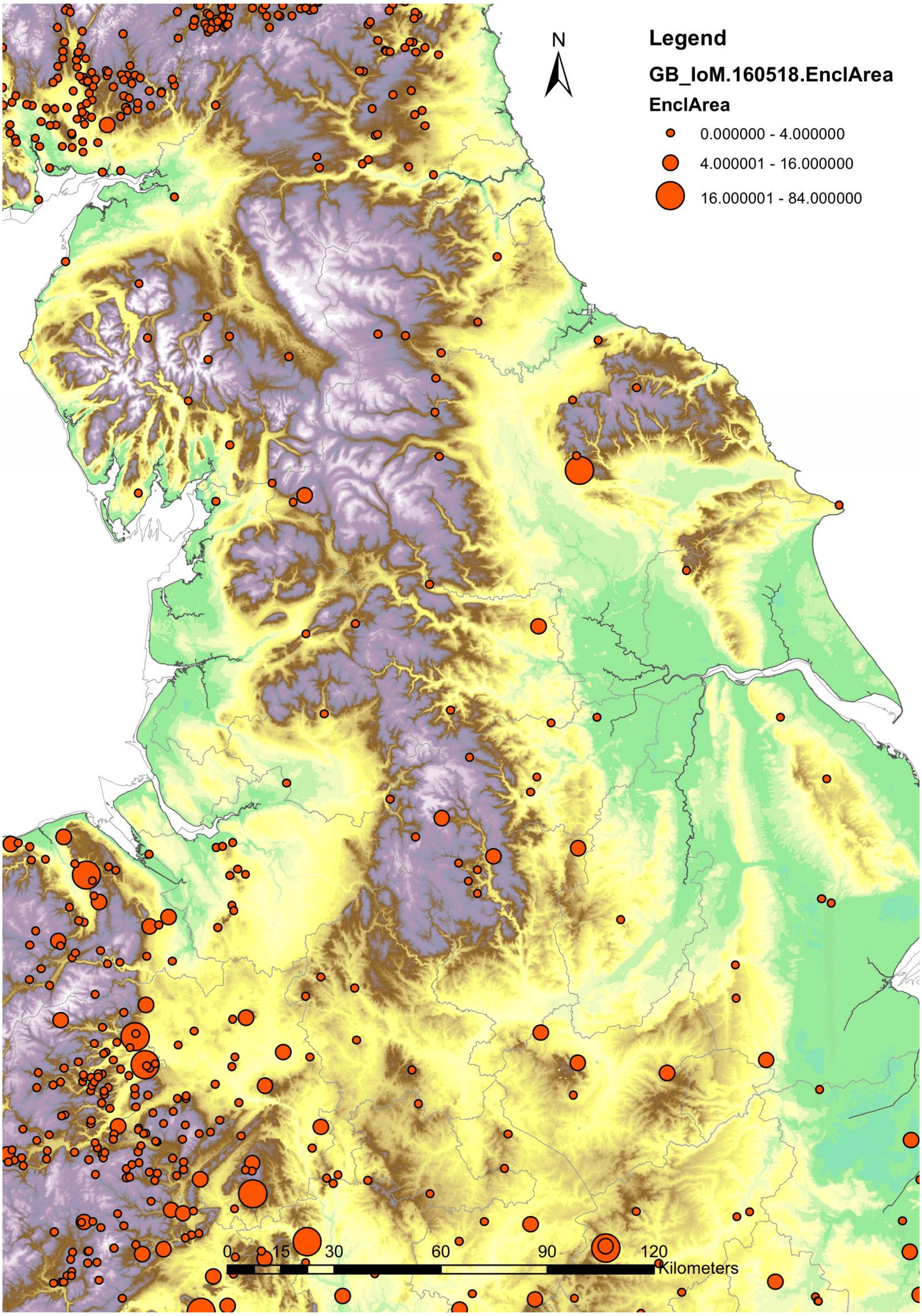


Figure 44 - Hillforts in northern England, by Enclosed Area

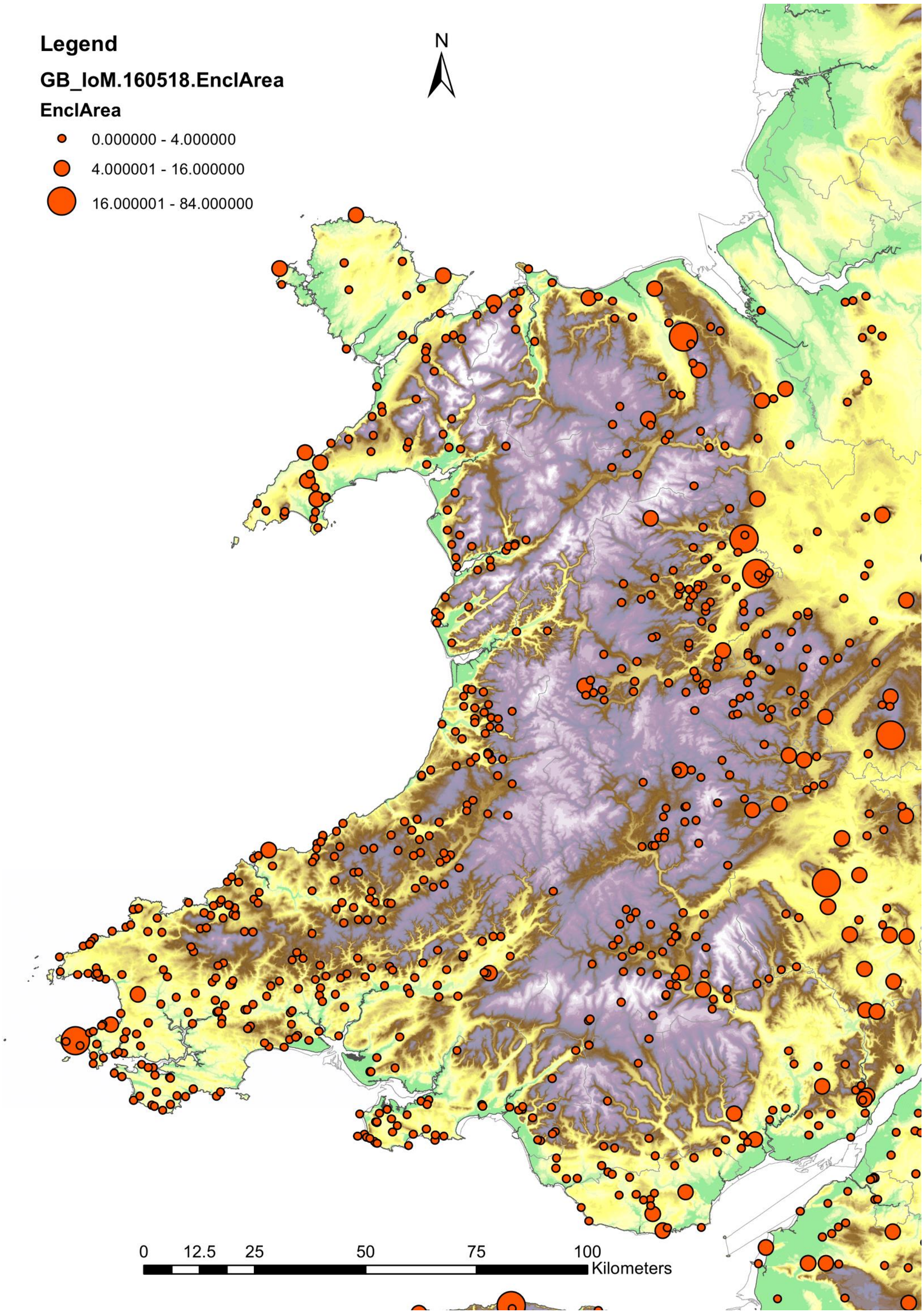


Figure 45 - Hillforts in Wales, by Enclosed Area

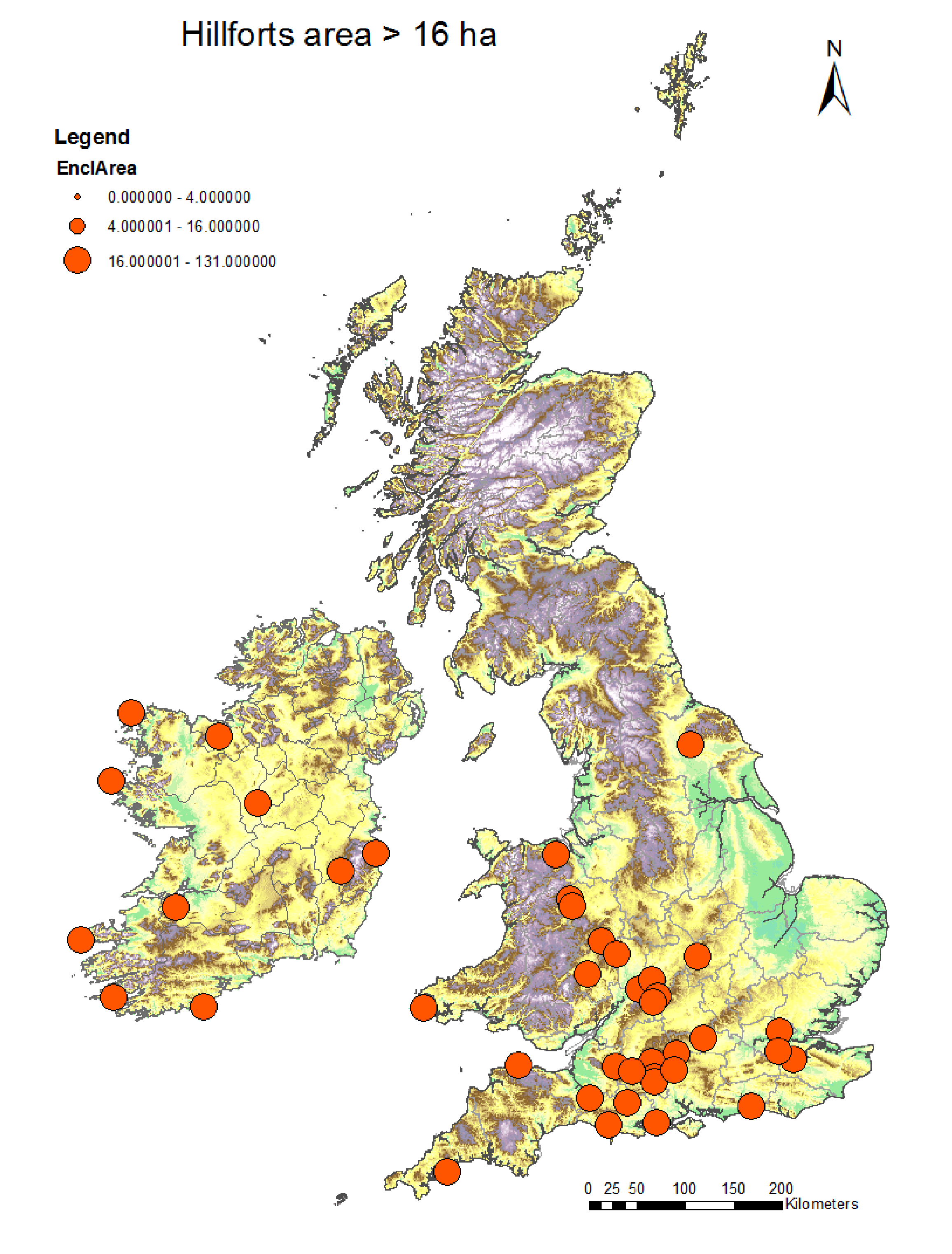


Figure 46 - British Isles: Hillforts > 16ha in Enclosed Area

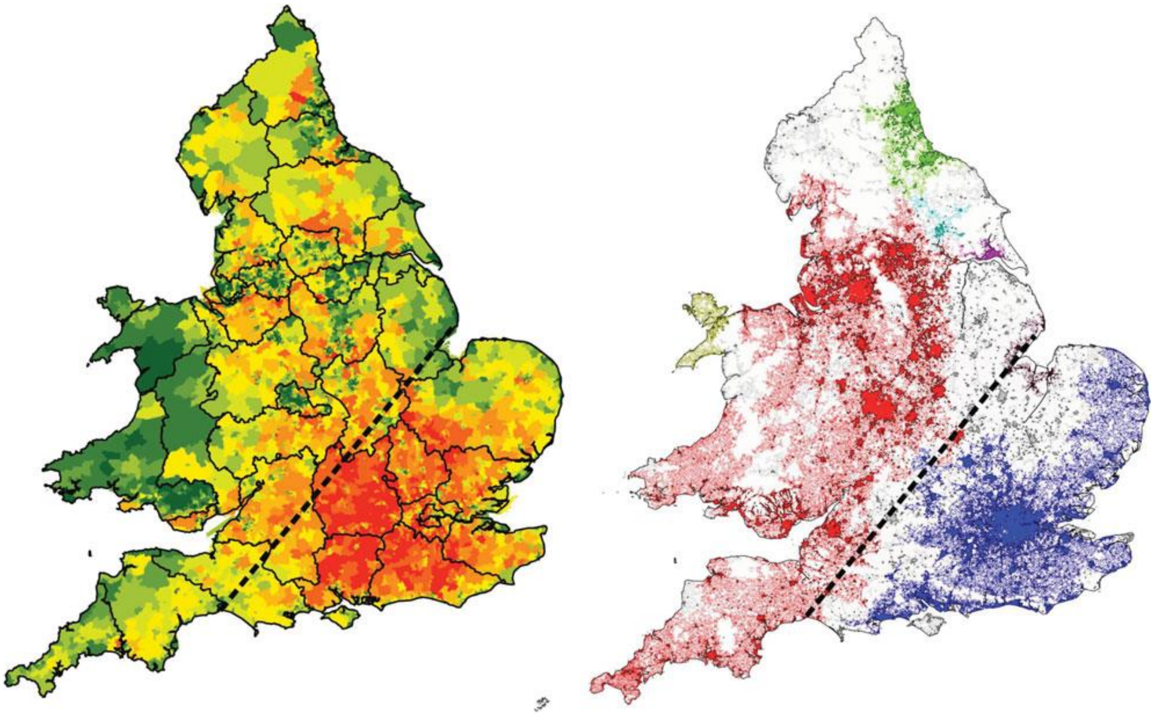


Figure 47 - Income and urbanism in modern England and Wales (Arcaute, Molinero, Hatna et al. 2016, Figure 3)

The discussions so far relate to an absolute comparison of hillforts on the basis of enclosed area, but with the density of sites varying so much within the British Isles it is also worth investigating sites locally according to their ranking compared with those within the surrounding region. Figure 48 shows the result of computations in R establishing the largest sites by size, relative to their nearest 100 neighbours. Note that because the inter-hillfort distance table computed for the percolation analysis was a sparse matrix, with inter-site distances only included below a threshold determined by the start of clusters appearing (40 km, allowing some safety margin), it was necessary to re-compute with a higher threshold (100 km) in order to ensure sufficient neighbours for this analysis to be robust in the less dense areas.

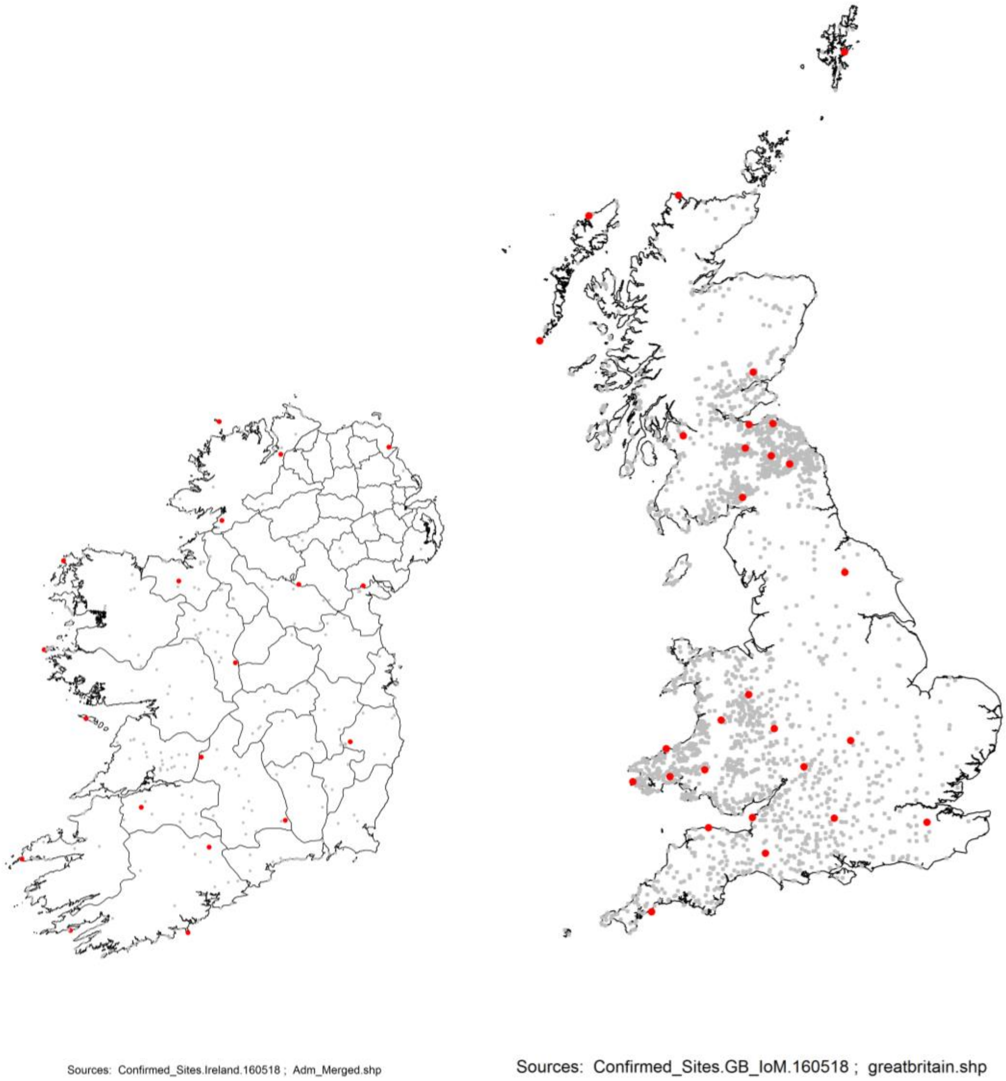


Figure 48 - Largest hillforts in Ireland and Britain relative to nearest 100 neighbours Largest sites red, other sites grey. Maps not to same scale

Comparison of the rank plot for Ireland with the absolute plot (Figure 40) shows much the same pattern but with some of the mid-size range sites emerging particularly in the north. One of the large sites on the east side of the Wicklow Hills, Coolagad (Atlas: 0719), is not shown, being quite close to Spinans Hill (Atlas: 0727), which outranks it in size.

Similar effects are apparent in the British distribution, but it is useful in identifying sites of possible local regional importance. Since the site density varies so considerably across Britain, it would be potentially of value to explore the effect of systematically applying a distance limit for neighbours, to get a more nuanced view of such sites, as a future area of research.

Comparison with other Datasets

Hillfort Kernel Density

A graphic way of showing the density of sites is to generate a kernel density plot (Baddeley and Turner 2005), showing the relative density of sites as a heat plot, overlaid on an outline map.

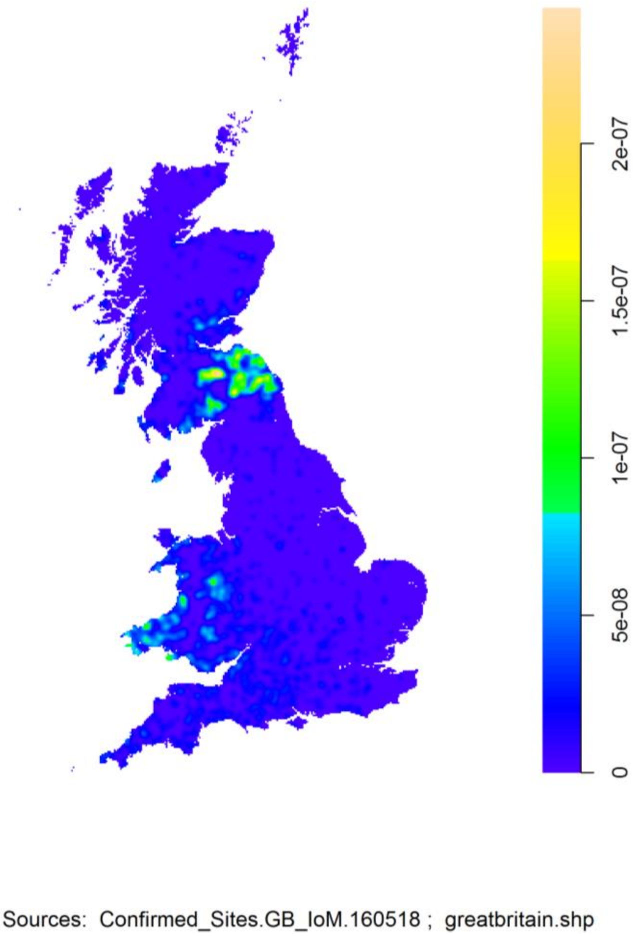


Figure 49 - Hillforts in Britain; Kernel density, sigma value 3974

Figure 49 shows a hillfort kernel density plot for Britain, with a sigma value of 3974 m, i.e. the radius of the probability distribution curve around each site at one standard deviation. The standard deviation value was computed using the Diggle function in the R 'spatstat' package (Baddeley, Turner and Rubak 2016; Diggle 1985) to generate an optimised kernel density smoothing value. Site densities are generally very low, which is why only a few 'hotspots' are apparent, and the probability density legend shows extremely small values. The mean nearest neighbour distance is 3.66 km, and median 2.35 km, as in the histogram in Figure 50, which reflects the impact of the very high concentrations of sites in southern Scotland and south-west Wales. Figure 51 shows a similar plot for Ireland, for a Diggle value of 3158m, illustrating the lower density of sites within the interior. The concentrations along the coasts are also apparent if not so obvious. The mean nearest neighbour distance in Ireland is larger than Britain at 5.19km, and likewise the median at 3.41km.

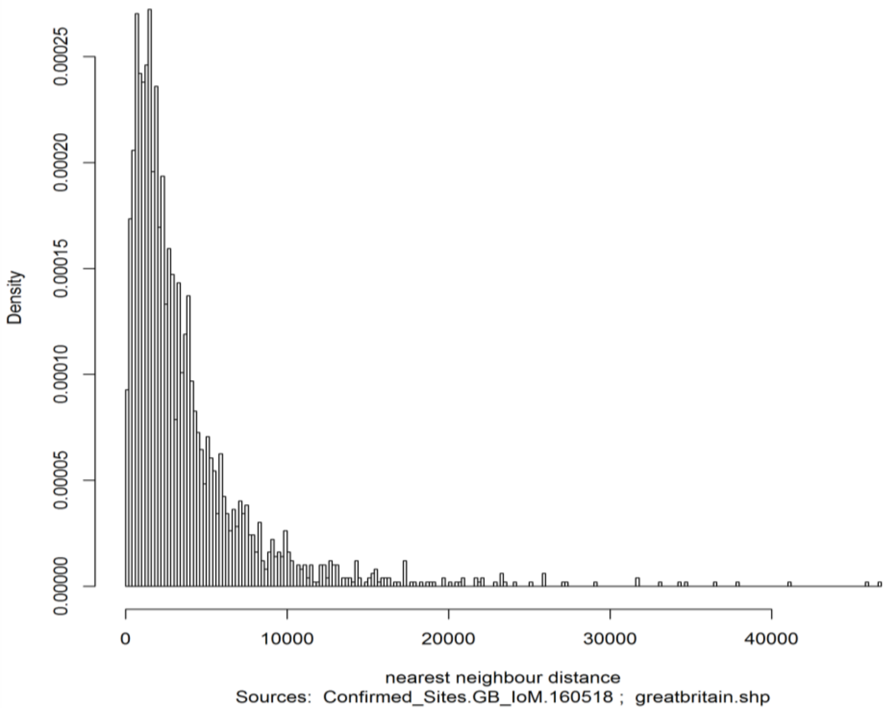


Figure 50 - Hillforts in Britain; nearest neighbour histogram

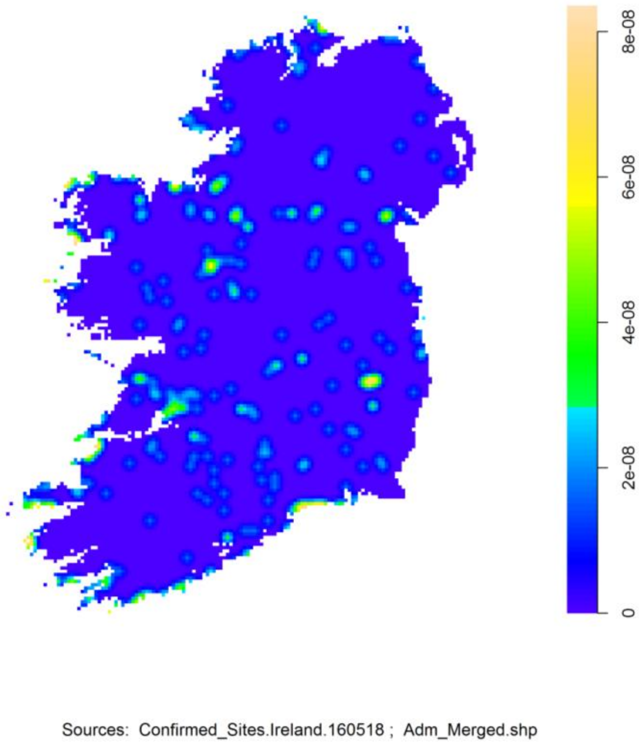


Figure 51 - Hillforts in Ireland; kernel density, sigma value 3158m

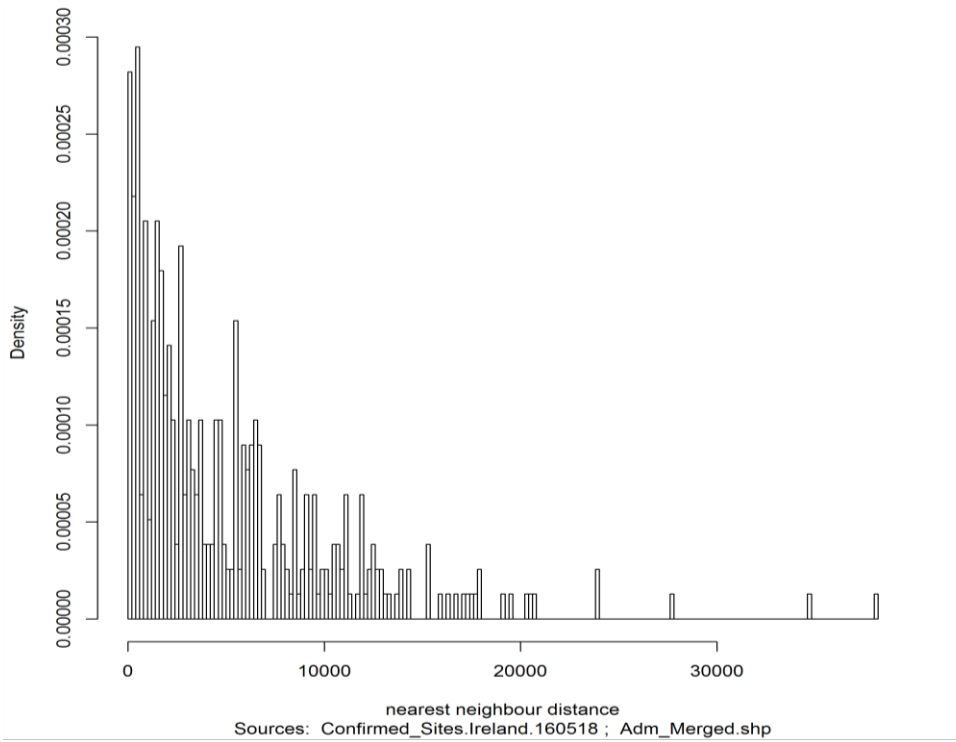


Figure 52 - Hillforts in Ireland; nearest neighbour histogram

Portable Antiquity Scheme – Iron Age Finds

The Portable Antiquities Scheme (PAS) database was used, once rights to access coordinate data were established (*The Portable Antiqutıes Scheme*). A filter on 'Iron Age' was set, and the entire database downloaded, a total of approximately 53,000 points. Of these ca. 16,500 had no coordinates, leaving the balance of ca. 36,000 points that could be plotted. It is presumed that the exact locations of sites with no coordinates were being kept confidential. Nonetheless this still provided a substantial dataset. Figure 54 shows a plot overlaid on the outline and counties of modern England, with hillforts shown in red and finds from the PAS scheme in grey.

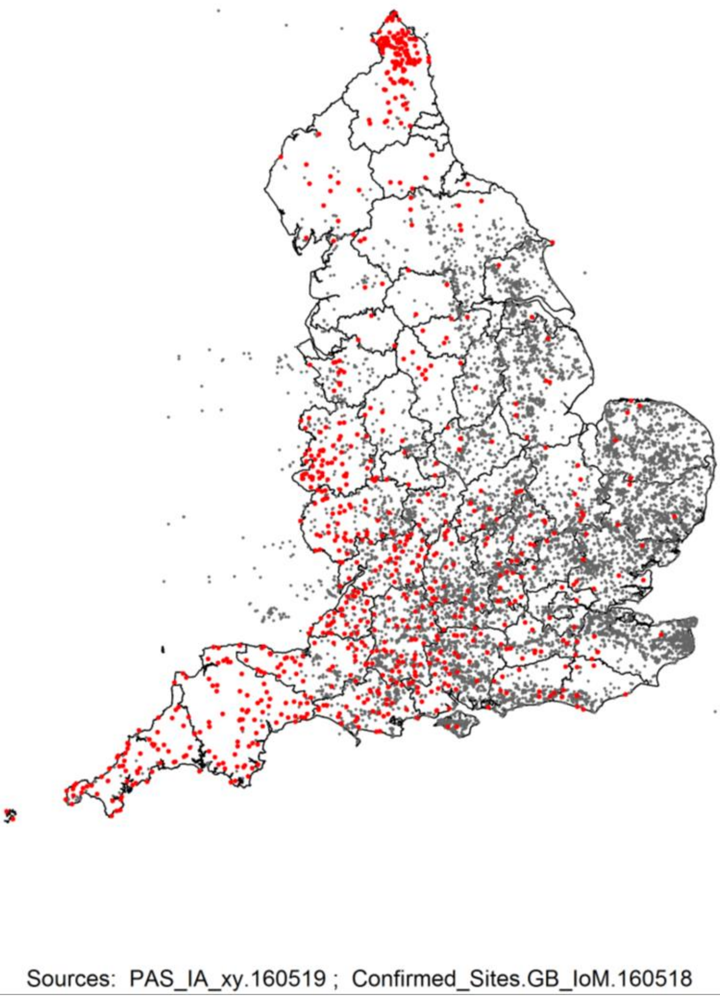


Figure 53 - The Distribution of Hillforts and PAS Iron Age finds in modern England Hillforts red, PAS finds grey

The PAS is a voluntary scheme that applies in both modern England and Wales, although interestingly only a few points in Wales emerge. It is primarily aimed at metal detectorists. There seems to be a reasonable density of finds up to the Welsh border, presumably because there is a community of metal detectorists in the densely populated Midlands who can easily access the borders. It is also possible that some finds are made in Wales and are not recorded, but another factor is the topography of Wales and the lower levels of population. There is however a marked drop off in finds going north towards the Scottish border, partly perhaps because of the lower population in that part of the country, and perhaps also less attraction for metal detectorists near the border, as there are quite different laws in Scotland where all finds must be reported under the laws of Treasure Trove, and there are strict provisions against unauthorised removal of any finds from the country (*Treasure Trove Scotland*).

There is also a paucity of finds in Devon and Cornwall, and at least in part this is due to the widespread landholdings of the Duchy of Cornwall, and their refusal to grant permission for metal detecting on their land (*Duchy of Cornwall, Frequently Asked Questions*). Another generally important factor is that metal detecting is not permitted on protected historic monuments, which excludes many hillforts, and their immediate environs.

Other observations are the few finds in the hinterlands of the Wash, but this is not surprising given that it was largely flooded fen and marshland until the Romans started occupying it (Jones and Mattingly 1990, 11-12); perhaps more surprising is the light density along the south Downs.

The overall observation from this map is that the density of finds is high where there is a marked absence of hillforts, particularly along the east coast from Yorkshire, through Lincolnshire, East Anglia and Kent. Although noting the comments above, this in part reflects the density of metal detectorists, the accessibility of land and ease of detecting (flat farmland is more attractive than rough hills), it nonetheless demonstrates that there was a lot going on in Iron Age England in areas with few if any hillforts. This substantiates the observation made by Hogg noted earlier (1975) that hillforts are not an essential characteristic of Iron Age life in Britain.

This is also shown in the kernel density plot of PAS finds for England, in Figure 54. Figure 55 shows a relative risk plot, computed in R, for PAS finds and Hillforts in England, showing a low degree of coincidence. The plot legend centre (the blue-green transition) is at the point of neutrality, computed as the ratio of the number of hillforts to the number of PAS finds, differences from this being statistically interesting. The low relative probability of hillforts is blue, where Iron Age PAS finds are most concentrated.

Further investigation based on different classes of finds on a more localised basis would be worthy of further study, particularly for later period Iron Age finds such as coins (Andrew Bevan, pers.comm. July 2016).

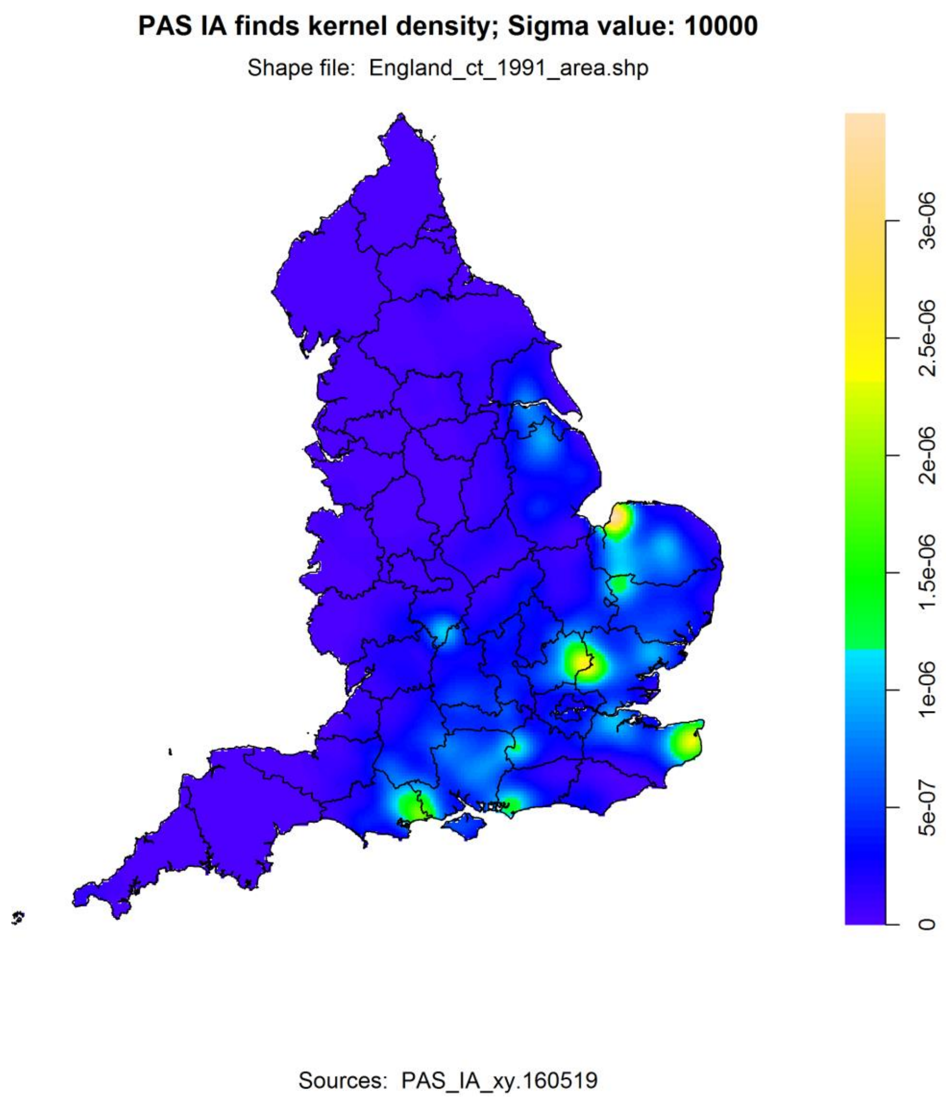


Figure 54 - Modern England; kernel density for PAS Iron Age finds

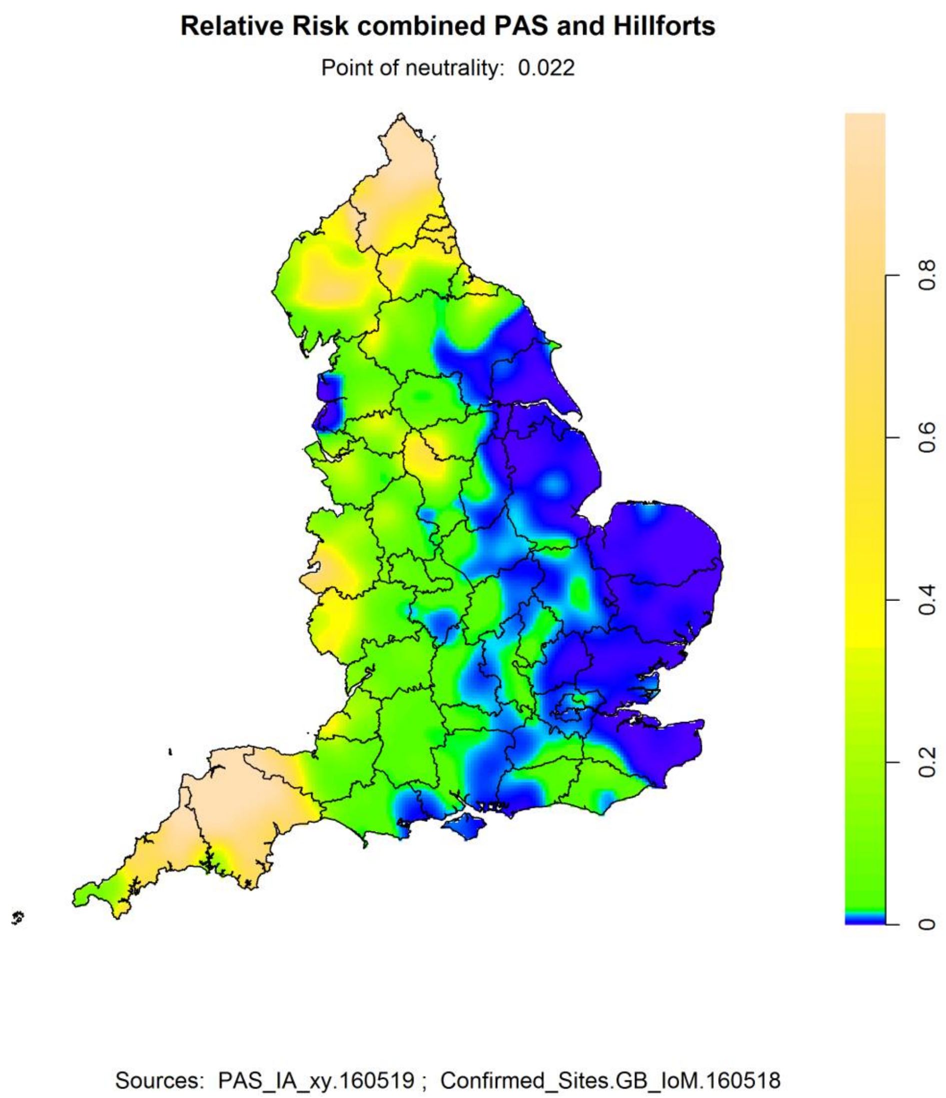


Figure 55 - Modern England; relative risk surface; Hillforts and Iron Age PAS finds (see notes in text)

Stone Circles

Stone circles occur in all parts of the British Isles, although the recorded number is fewer than that of hillforts. A database of stone circles with their coordinates was provided by Damon Ortega (pers. comm., May 2016), derived from Burl (2000), and forms the basis for some of his current research. Although considered to be much older than hillforts, is there some evidence of cultural continuity that can be determined by comparing the two datasets?

Both hillforts and stone circles are generally located in upland areas, and represent a significant amount of effort in terms of sourcing materials, moving them to the site and then undertaking construction. For stone circles the main requirement would have been stone, but by supposition also timber for levers and rollers, as well as other material for tools for excavation. Hillforts also required considerable amounts of material to be moved, and in many early cases this included very large amounts of timber (Brown 2008, 41-49; Ralston 2006, 46-67), although the soil and stones used were almost certainly always obtained in the immediate vicinity (e.g. the formation of embankments by material excavated from an adjacent ditch). Various types of tool of different materials were also presumably required.

By inference both classes of monument were regarded as special places in order to justify this commitment of effort, and there may have been some commonality of function, although hillforts would certainly have had a much broader range of use given their form. Despite stone circles being significantly older than hillforts, they would have been recognisable in the landscape as the work of earlier people. Were similar factors at work in terms of the decisions made as to the siting of these two different types of monument?

Analyses were run for Britain, with 1001 stone circles and in Ireland with 365. It was necessary to do this separately because of the different coordinate systems used in these two regions.

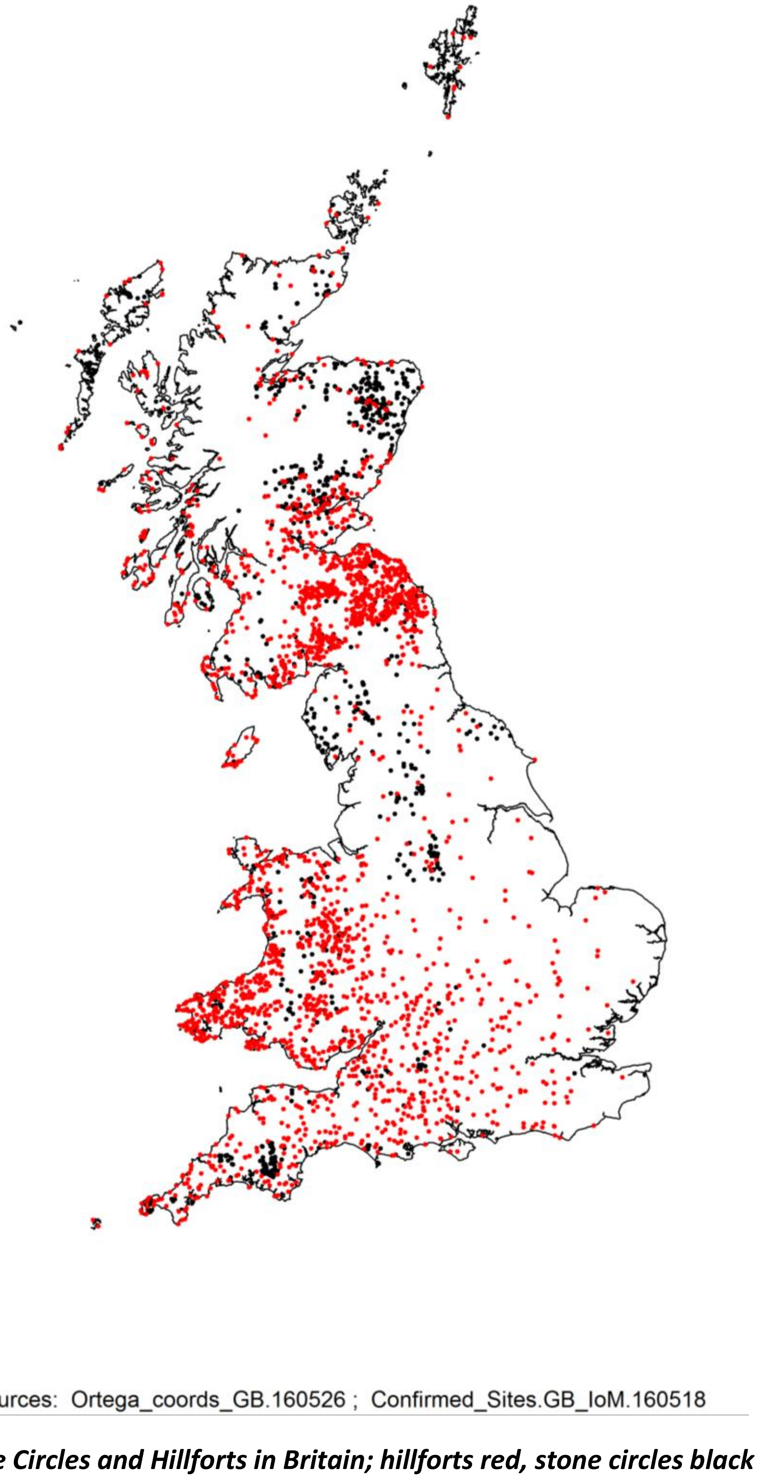
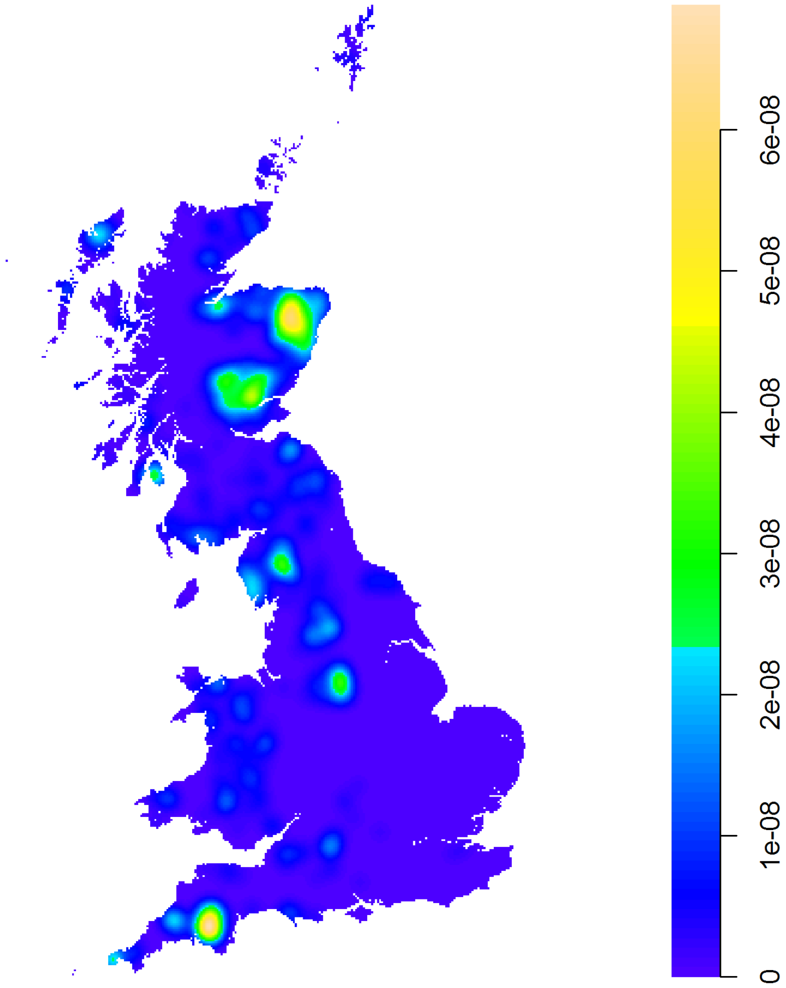


Figure 56 - Stone Circles and Hillforts in Britain; hillforts red, stone circles black

Figure 56 shows a plot of hillforts and stone circles in Britain, showing some modest degree of overlap, but also low densities of hillforts where there are the greatest concentrations of stone circles, such as the Pennines and east central Scotland. Figure 57 shows a kernel density plot of stone circles in Britain, using a sigma value of 10,000m. The computed Diggle value of 1431m gave very little visibility, given the relatively small number of sites.



Sources: Ortega_coords_GB.160526

Figure 57 - Stone Circles in Britain; kernel density, sigma 10,000m

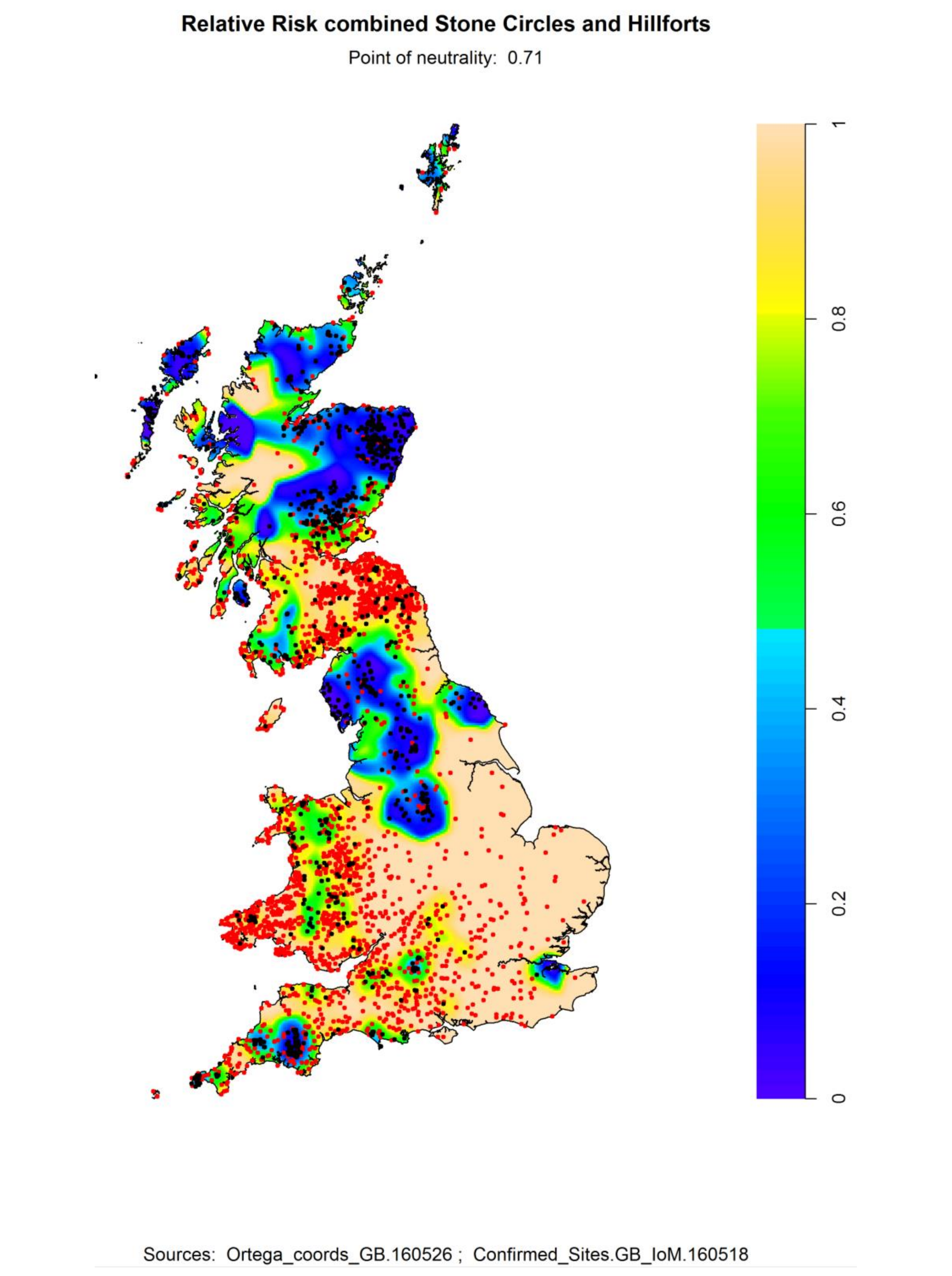


Figure 58 - Stone Circles and Hillforts in Britain relative risk surface
Hillforts red, stone circles black

Figure 58 shows a relative risk surface of Stone Circles and Hillforts, also computed in an R program. The sites of hillforts are shown in red, and the stone circles in black, to aid clarity. The centre of the legend (the yellow end of green) is set at the point of neutrality. This shows higher degrees of coincidence in Wales and the Cotswolds, and south Central Scotland, but low degree of relative probability on Dartmoor and Bodmin Moor, the Pennines and Aberdeenshire, Angus and Fife. Although both stone circles and hillforts predominantly favour elevated sites and hills, there is otherwise little evidence of overlap, and therefore probably little in the way of cultural continuity in their locations. Very similar conclusions can be drawn for Ireland, as shown in Figure 59. Figure 60 shows a relative risk plot, with the point of neutrality shown by the yellow end of green, with a low relative probability of stone circles and hillforts.

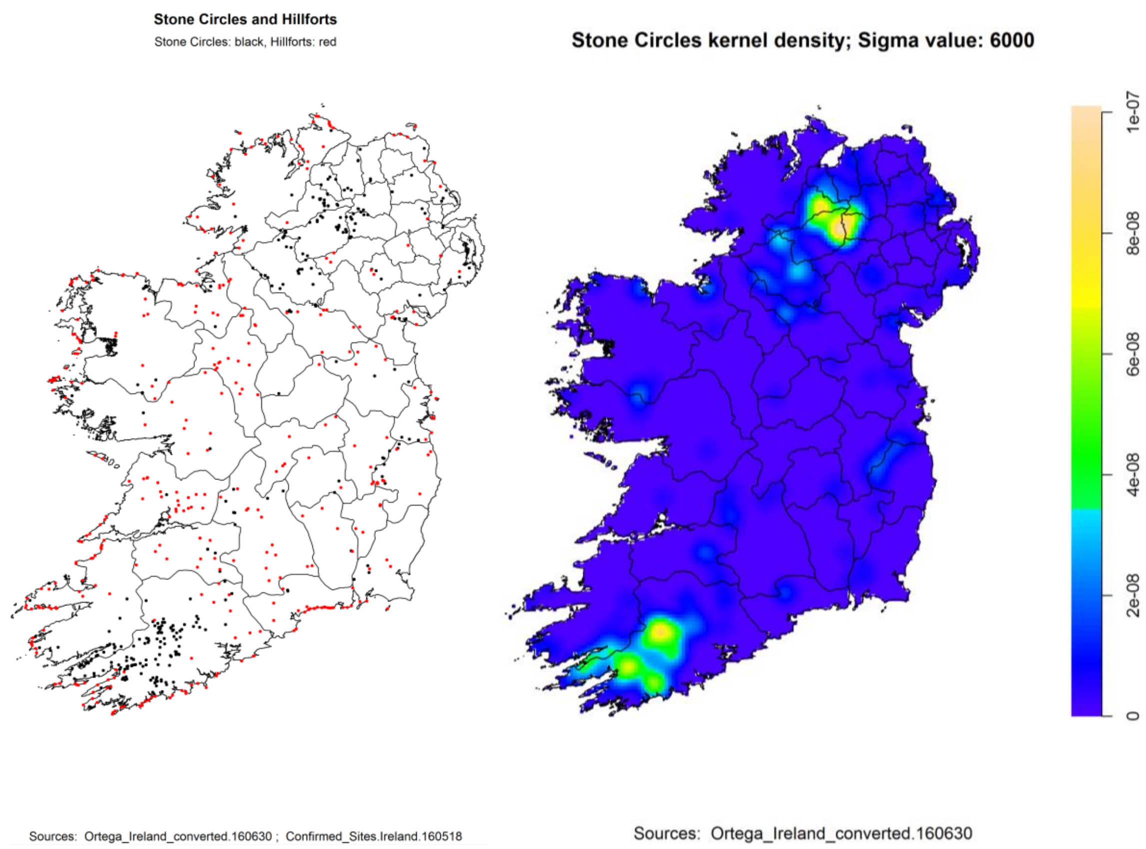


Figure 59 - Ireland: Distribution of Stone Circles and Hillforts; kernel density for Stone Circles Hillforts plotted in red, stone circles in black

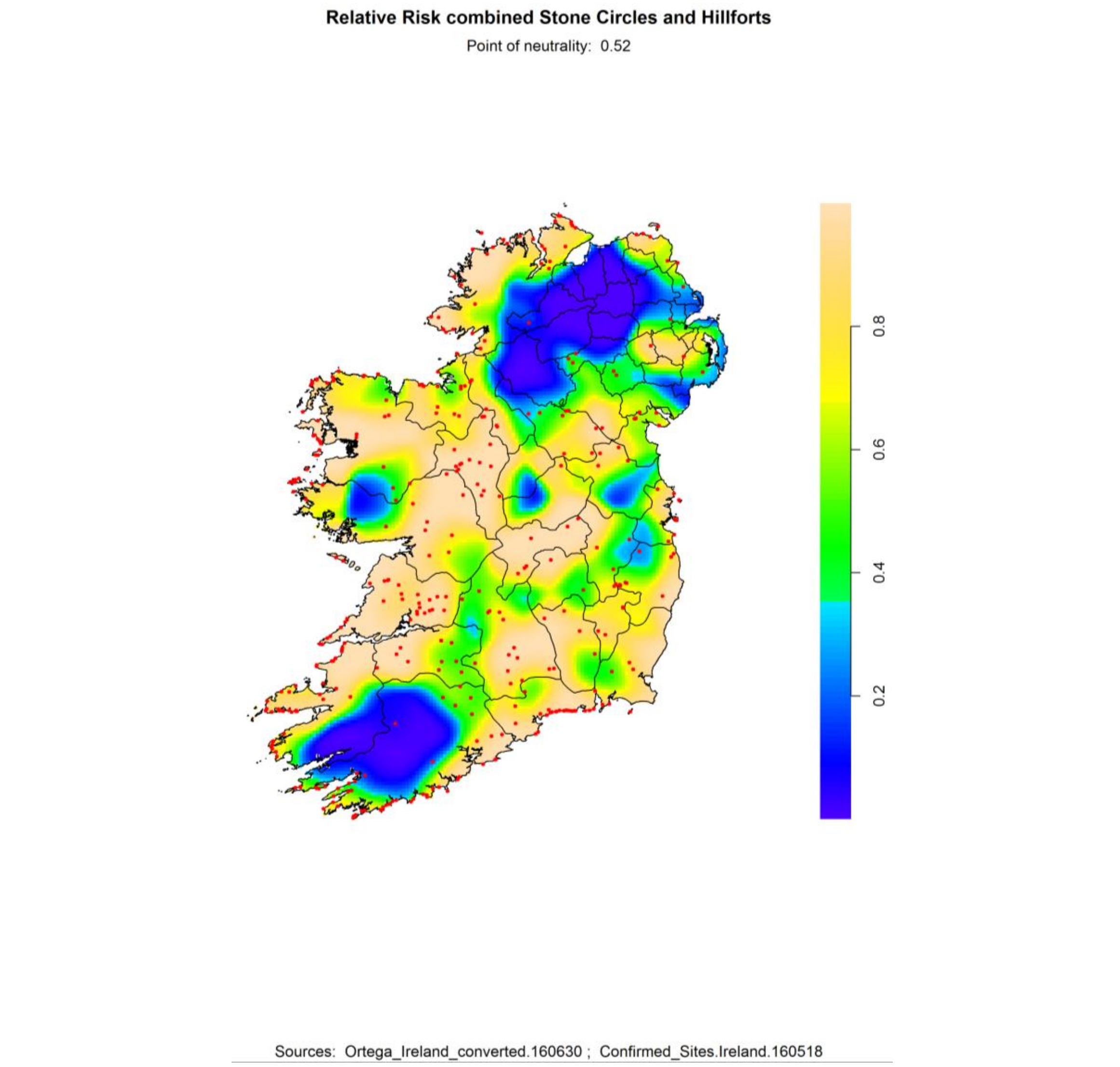


Figure 60 - Ireland: Stone Circles and Hillforts relative risk surface
Colour scale centred on point of neutrality; sigma 10km
Hillforts plotted in red

Irish Ring Forts

Irish Ringforts are an extremely numerous class of monument in Ireland, dating from the early Christian period, between 300AD and 1300AD, but probably mostly constructed between 300AD and 600AD (Stout 1997). In the order of 45,000 sites are known. Mostly these are simple structures with a circular ditch and bank, are generally regarded as single family farmsteads, and reflect an upsurge in population following improvements in agricultural technology. Different classes of site possibly reflect hierarchies with 'kings' occupying the larger more complex sites (Stout 1997).

How might these relate to the distribution of hillforts? Although hillforts are from an earlier age, and have very likely different functions, could the ringforts reflect the locations of the bulk of the population in the earlier Iron Age?

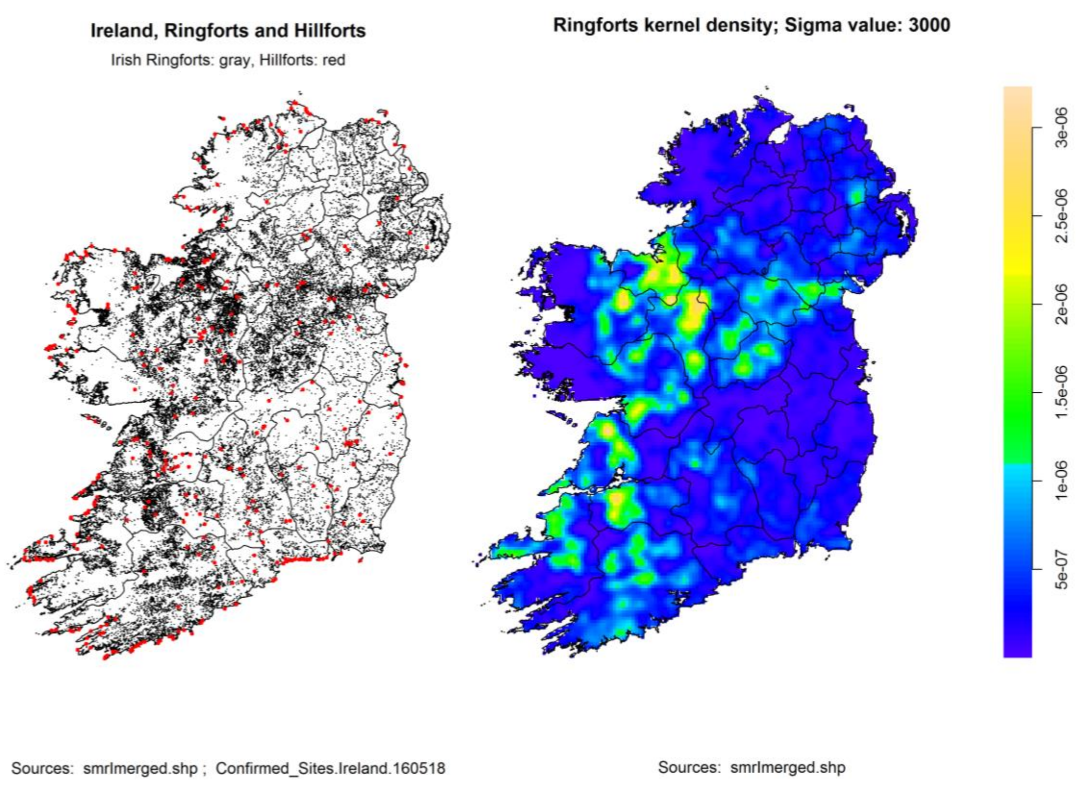


Figure 61 - Distribution of Ringforts and Hillforts in Ireland; Ringfort kernel density Hillforts plotted in red and ringforts black.

Figure 61 shows on the left the distribution of hillforts and ringforts in Ireland, and on the right a kernel density plot of ringforts. Visual inspection suggests rather different patterns, with as noted before many hillforts being located along the coasts.

A relative risk plot of hillforts and ringforts is shown in Figure 62. This shows essentially that ringforts and hillforts are more strongly associated with one another than would be expected from random distributions, with little exception. This is however rather a weak association as can be seen from the legend on the right, with the area of higher probability only marginally above the point of neutrality. This may reflect Stout's view of rapid expansion growth through new technology (page 132), which then opened up new areas of occupation including the hinterland away from the majority of hillforts on the coast in particular and their immediate surroundings.

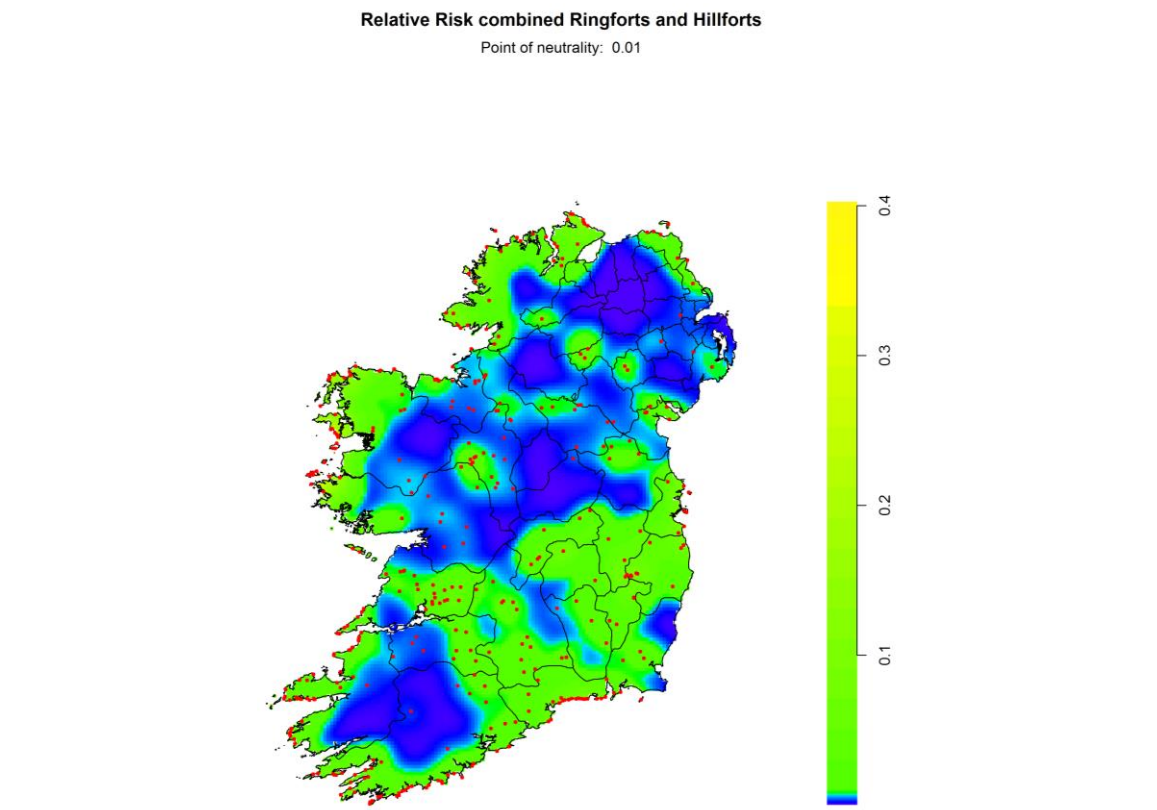
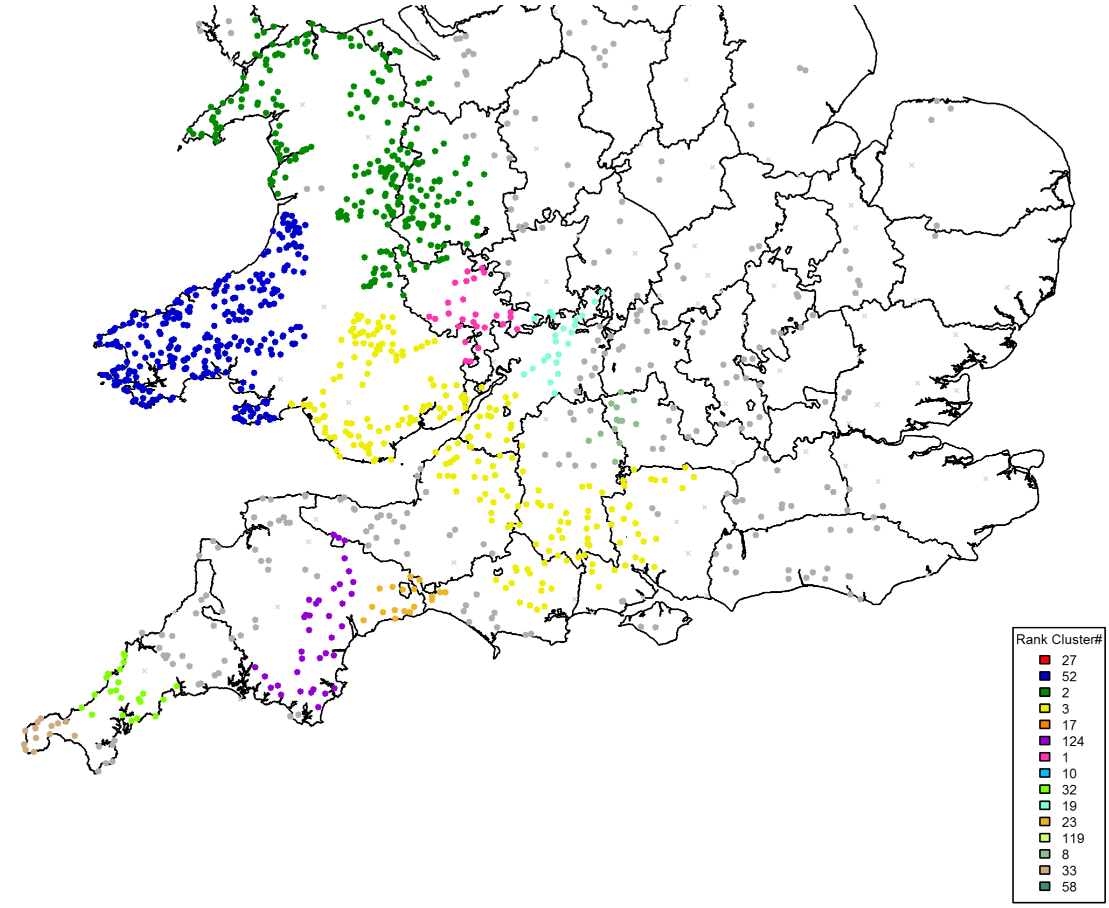


Figure 62 - Ireland, Relative risk plot for Ringforts and Hillforts colour scale centred on point of neutrality; sigma 10km

Domesday Counties

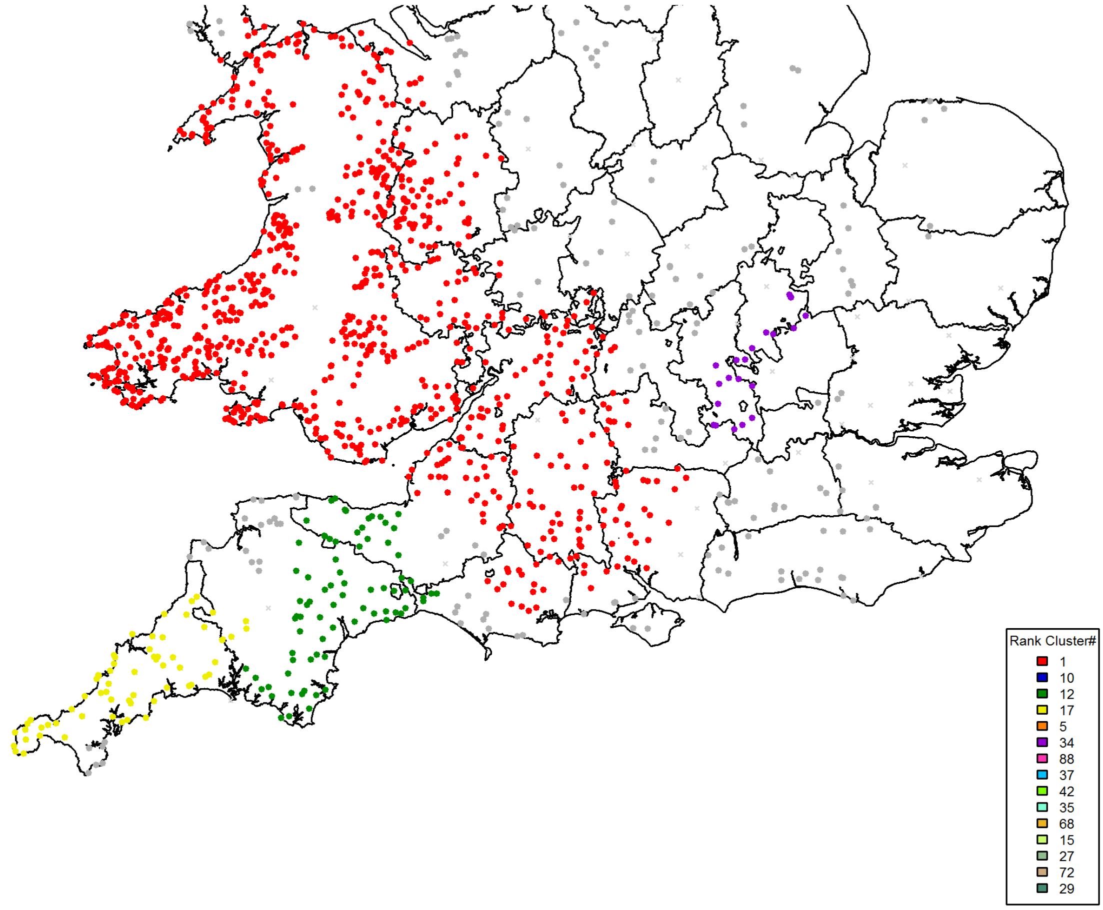
The Domesday counties are a good historical record of England's economy, population and administrative structure in 1086, as noted earlier ('Other Data'). To what extent might these echo patterns of the Iron Age? After all, the county structure of England survives to this day, and was largely unchanged until the latter 20th Century. Plotting hillfort clusters from the Percolation Analyses might give some indication of this.



Source File: Confirmed_Sites.GB_IoM.160518.csv ; percolation distance: 10 km

Figure 63 - Hillfort clusters and Domesday shires; 10km percolation radius

Figure 63 shows a plot of percolation clusters at a radius of 10km, overlaid on an outline of Domesday Shires south of the Dee. Two clusters are particularly notable: the pink one which fits remarkably well into Herefordshire and the second is the turquoise cluster that fits into the north-east part of Gloucestershire. The large lemon yellow cluster stretches from south-east Wales all the way through Wessex (east Somerset, parts of Dorset, Wiltshire and Hampshire).



Source File: Confirmed_Sites.GB_IoM.160518.csv ; percolation distance: 12 km

Figure 64 - Hillfort clusters and Domesday shires; 12km percolation radius

Figure 64 shows the clusters at a percolation radius of 12km. Of interest here is: the yellow cluster in Cornwall, with just a little spill over into Devon; the green cluster that covers a good part of Devon, and part of north-west Somerset; the purple cluster that is in the south of both Buckinghamshire and Bedfordshire (the Chilterns). As noted earlier, percolation analysis does not seek to explain, but it does highlight patterns. In this case it suggests that the some of the geographical factors that influenced the building of hillforts in south-west England and the Home Counties may have influenced the formation of administrative and political units by Domesday. However it also seems more likely that there was a degree of unitary continuity in these territories which directly evolved through the Iron Age to medieval England, being culturally and community based around a common identity. Some of that is self-evident for example in the history of the Cornish language to its decline in the 17th Century (*Maga Kernow, Cornish Language Partnership;* Peters 2005).

Discussion of Specific Regional Clusters

Selected percolation clusters

Four percolation clusters have been selected, as noted earlier: Cornwall; the Cotswolds and the lower Severn Valley; Central Wales and the Marches; the Chilterns. In addition to the percolation criteria described earlier, these have been selected with a focus on more familiar territory for the author and because of their contrasts in terms of topography. Cornwall is a peninsula, central Wales and the Marches a hilly zone away from the coast, the Cotswolds are adjacent to a big river valley and the Chilterns an inland escarpment.

Sites for each selected percolation cluster are plotted on the topography for the local region, with a symbology based on site area, as discussed above. Key rivers have also been plotted, selected from OS Open Rivers. This proved its own challenge, as rivers are not necessarily found under the most obvious name (for example the Severn is a secondary name for much of its upper stretches in Wales), and indeed many have different names for different parts. Experimentation was required, but the most effective was to display all watercourses, and then select the river in question using the information option in ArcGIS, in order to identify the required name(s).

Unless otherwise stated, details of specific sites in this section are taken directly from the Atlas data and from inspection of detailed maps in Streetmap (*Streetmap* 2012).

Cornwall

Cornwall is shown in Figure 65 for a percolation radius of 11km, with sites south-west of the Fowey. From 12km to 14km (the latter in Figure 66) sites east of the Camel and Fowey are included, with four on the northern edge of Dartmoor across the Tamar. This is quite a 'robust' cluster in the sense of its existing with little change over several radii.

There is one large site on the south, Dodman Castle (Atlas: 0603) with commanding views over the sea. Unlike many other promontory forts where only a small part of the headland is used, the ramparts that form it are 600m long, attesting not only its importance, but also the resources required to construct it. It has views across a very wide stretch of sea, from the Salcombe estuary in the east, to the Lizard in the west, encompassing Plymouth Sound and the Fowey, Fal and Helford estuaries. The only other large site is Carn Brea (Atlas: 0609) at the head of the Red River in the long important mining district around modern Camborne, with significant occupation back at least into the Neolithic.

The two medium sites along the line of the Camel and Fowey are Castle Killibury (Atlas: 0615) to the north and Castle Canyke (Atlas: 0613), and are strategically located to command a possible transhipment route between the rivers, the latter also for the tin route off Bodmin Moor down the Fowey. This portage is suggested by Sherratt (1996, Figure 2a, 215) although not explored in any detail in the text. This route would save considerable dangers in rounding Lands' End and the Lizard; the Camel being the only significant estuary on the north coast. This route's possible importance is attested by the location of two of only three known Roman forts in Cornwall, respectively a few miles north and south of Castle Canyke at Nanstallon and Restormel (Quinnell 1986; *Restormel Roman Fort;* Roman fort found in Cornwall 2007). The short lifetime of the Nanstallon fort may reflect the fading of the transhipment route, whilst Restormel's longevity confirms the ongoing importance of tin moved south from Bodmin Moor.

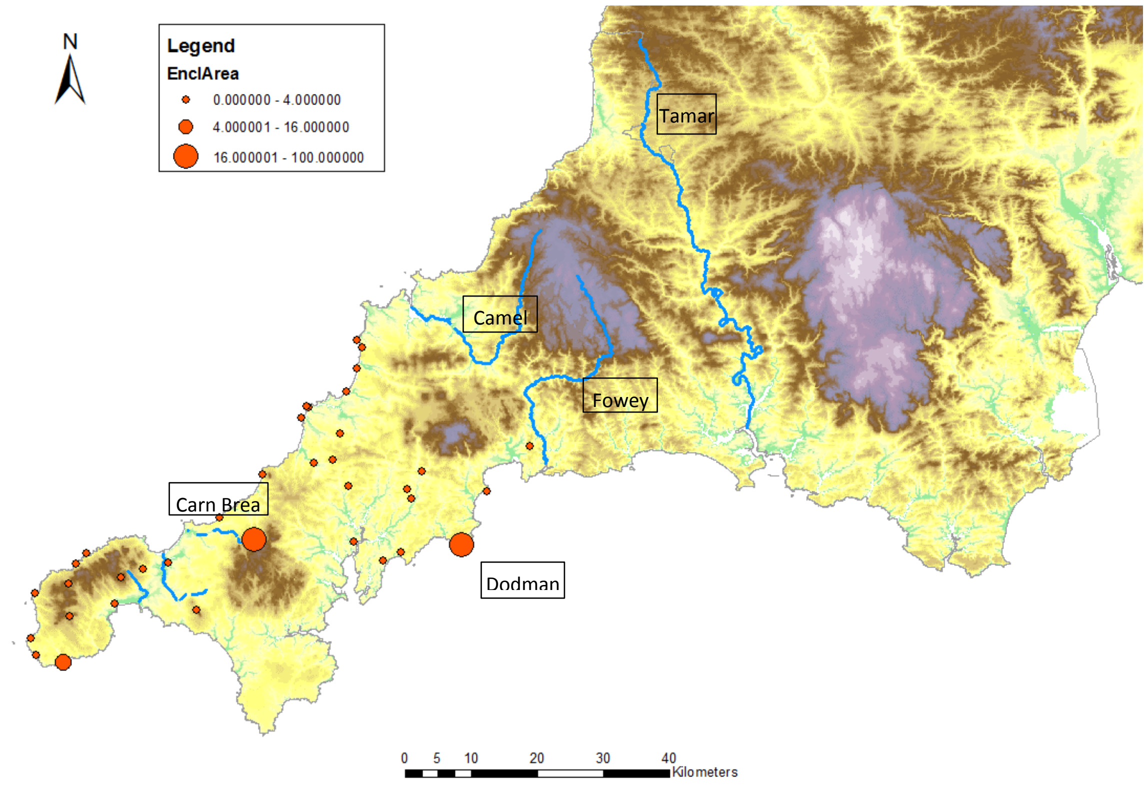


Figure 65 - South-west England cluster; 11km percolation radius, with key rivers

There is also a group of three medium sites around the Helford estuary: Little Dennis (Atlas: 0646) on the south peninsula, Rosemullion (Atlas: 0652) on the north and Gear Farm (Atlas: 0621) up the valley, suggesting control of this river. Similarly located sites around the Fal are listed in Hogg, but if they existed have probably been destroyed by later naval fortifications. Trade along this south-west coastal route is also briefly discussed in Cunliffe (1991, 436-438).

Otherwise hillforts are small, with many located on coastal promontories.

This cluster is consistent with a grouping substantially determined by topography and geography, namely the intrinsic effect of the peninsula, the River Tamar and Dartmoor, and which long maintained its identity politically, culturally and linguistically, to some extent even today (e.g. Peters 2005). Looking at hillforts on the basis of their size suggests a role for the larger sites, consistent with Sherratt's (1996) view of hillforts as being important in controlling trade and portages, probably dominated by metals and other minerals.

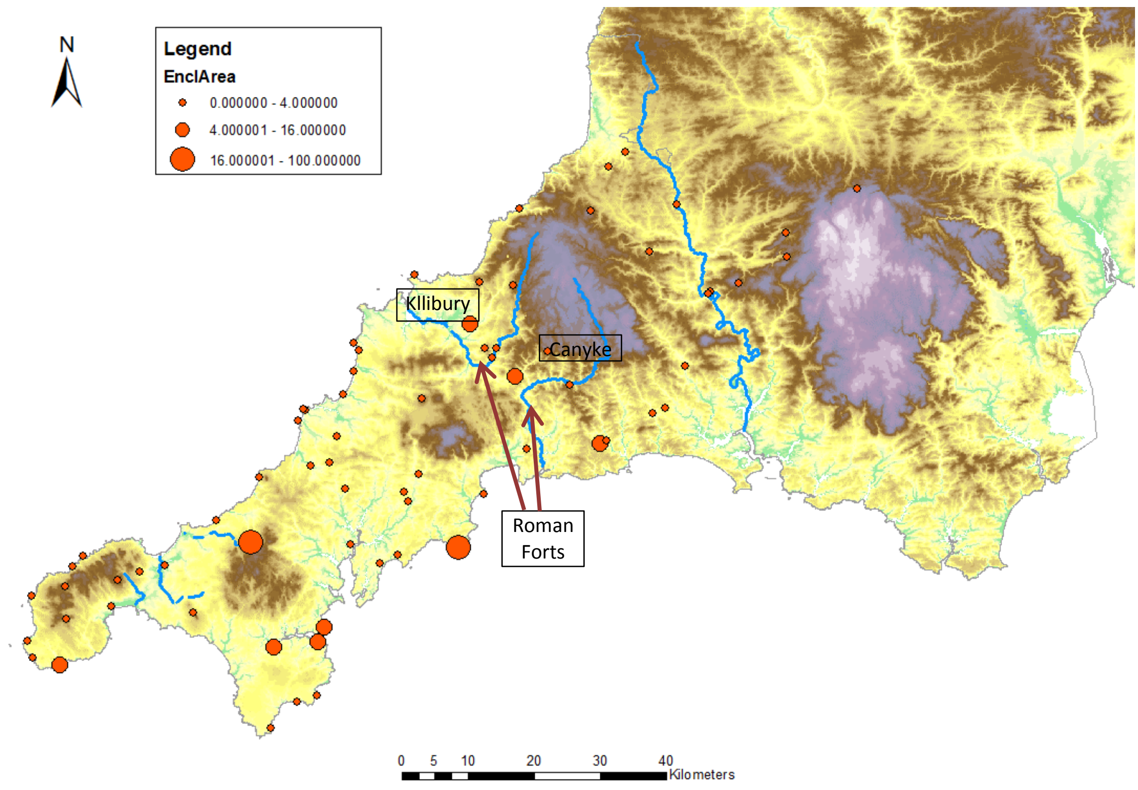


Figure 66 - South-west England cluster; 14km percolation radius, with key rivers

Central Wales and the Marches

The cluster in Central Wales and the Marches is topographically distinct, being situated around a high hilly region, with the upper stretches of the Severn being a key feature. This is a 'robust' cluster, appearing initially at 9km as it breaks out from a larger west, north and central Wales grouping (Figure 68). This incorporates sites that lead around the south to the upper Wye, as well as the large site at Titterstone Clees (Atlas: 0091) over the river Teme to the east, but these fall out at 6km leaving a core stretching out on either side of the Severn, see Figure 67.

There are two dominant sites on the upper reaches of Severn, Llanymynech Hill (Atlas: 0071) and Y Breiddin (Atlas: 1276). Llanymynech Hill is one of the largest sites in Britain at 57ha, and is located where the rivers Vyrnwy, Tanat and Cain reach the Severn Plain. It was very important as a source of copper, zinc and lead through the Iron Age, and there is evidence of metal working dating earlier than this. Y Breiddin at 28ha is located on the southern side of the Severn, shortly below where it is joined by the river Vyrnwy.

These larger sites suggest roles as dominant control points or entrepots for goods moving from the hilly hinterland to plains and low lands beyond, down the river Severn as Brown (2008, 196-204) elaborated in some detail, building on from the ideas of Sherratt (1996).

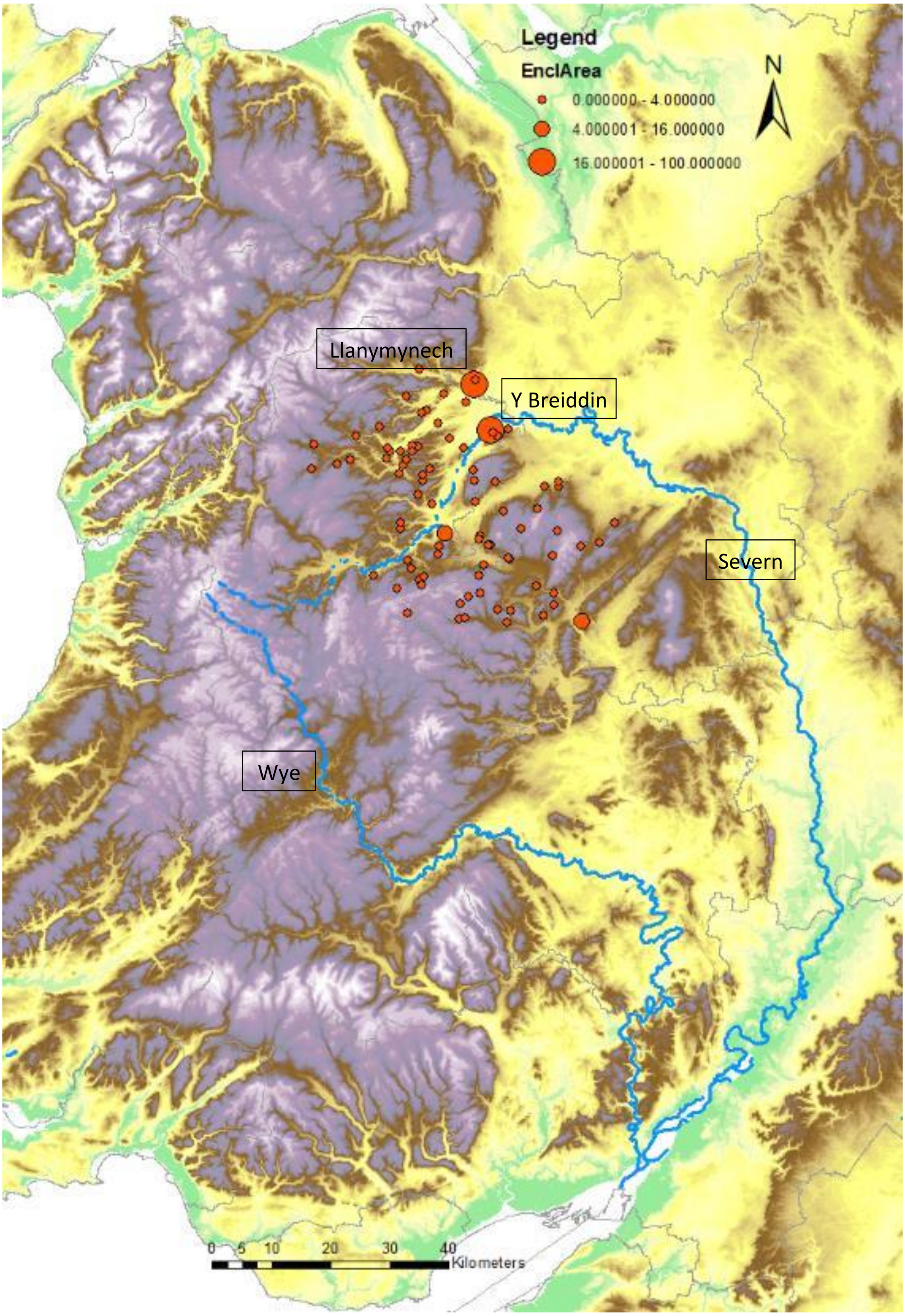


Figure 67 - Wales and the Marches cluster; 6km percolation radius, with key rivers

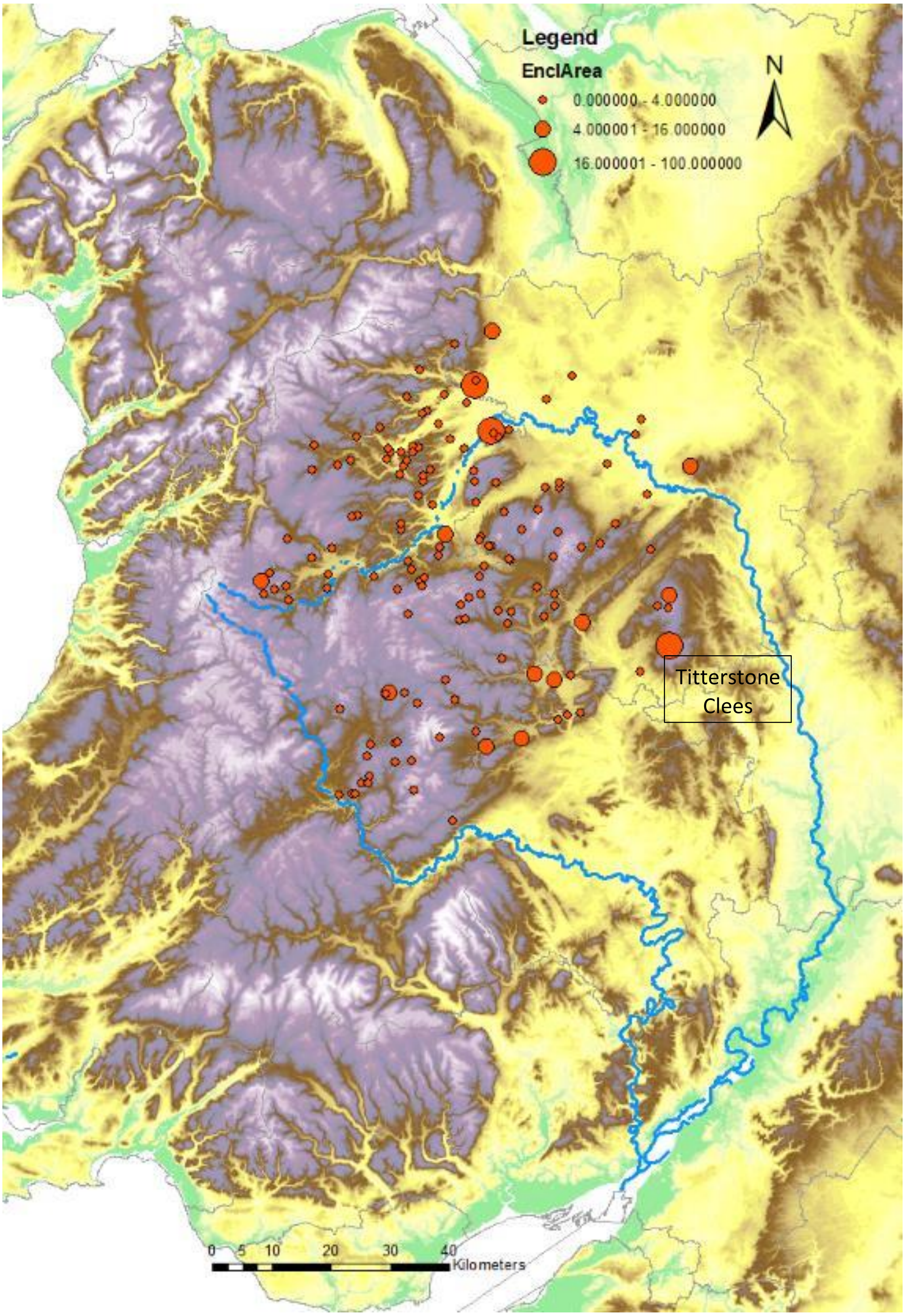


Figure 68 - Wales and the Marches cluster; 9km percolation radius, with key rivers

Severn Valley, Cotswolds

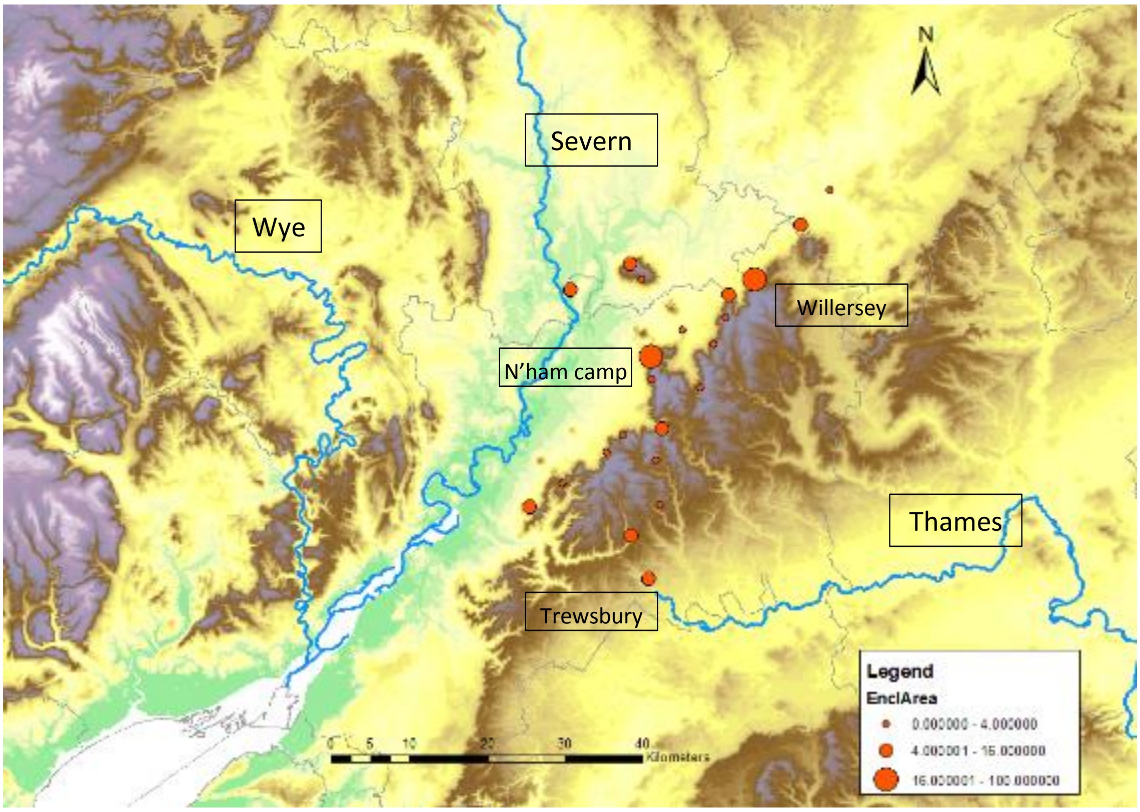


Figure 69 - Severn Valley, Cotswolds cluster; 10km percolation radius; with key rivers

Figure 69 shows the cluster at 10km radius located quite distinctly in the topography of the Cotswolds, with two large and ten other sites on the north-west edge of the escarpment overlooking the Severn Valley, Gloucester and Cheltenham. The largest two are Nottingham Hill Camp (Atlas: 0760) at 48.6ha, north-east of Cheltenham, on a sharply defined promontory, also enclosing a Bronze Age barrow cemetery; to the north-east is Willersey Camp (Atlas: 0756) at 28ha between Broadway and Chipping Camden, with a Neolithic long barrow in its interior.

The upper reaches of the Thames are also shown to the south, with one hillfort, Trewsbury (Atlas: 0767), next to its source, as well as the Fosse Way, a railway and the Thames and Severn Canal. Three other sites are very close to the Severn to the north. The two larger ones are: Towbury Hill Camp (Atlas: 0766) 4.2ha, next to the M50 Severn bridge 4km north of Tewkesbury; Kemerton Camp Bredon Hill (Atlas: 0378) 7.1ha, on a steep sided promontory overlooking sharp bends of the Avon. At a radius of 11km (Figure 70), this cluster expands to include more sites on the south-east of the Cotswolds and another group on the other side of the Severn, north of the Forest of Dean and up the Wye on both sides, although this is through a 'thin' link via Towbury, suggesting that they are probably two distinct groups.

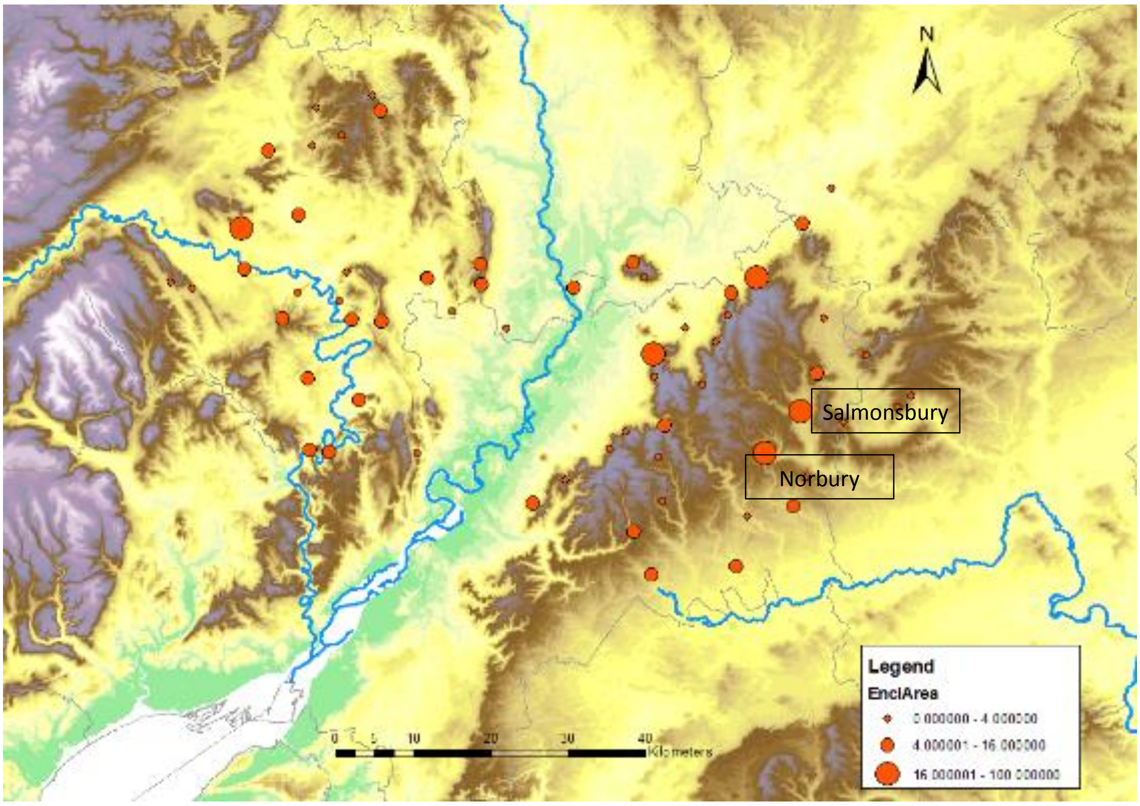


Figure 70 - Severn Valley 11km percolation radius; Enclosed Area

On the south-east of the Cotswolds the two large sites are: Norbury Camp Northleach (Atlas: 0753) 22 ha, right next to the Fosse Way, enclosing a Long Barrow; Salmonsbury (Atlas: 0757) 23ha to its north-east, a marsh fort situated in the fork between the Windrush and the confluence of the Dikler and Eye at Bourton-on-the-Water, and was possibly an enclosed oppida (Cunliffe and Rowley (1976, 330).

The location of the larger sites, reflect not only the longevity of their importance, through incorporation of much older monuments, but also their role in trade, being positioned on key waterways and routes, which have continued in importance, some up to modern times (e.g. Roman road, M50 river crossing, canal and railway).

Chilterns

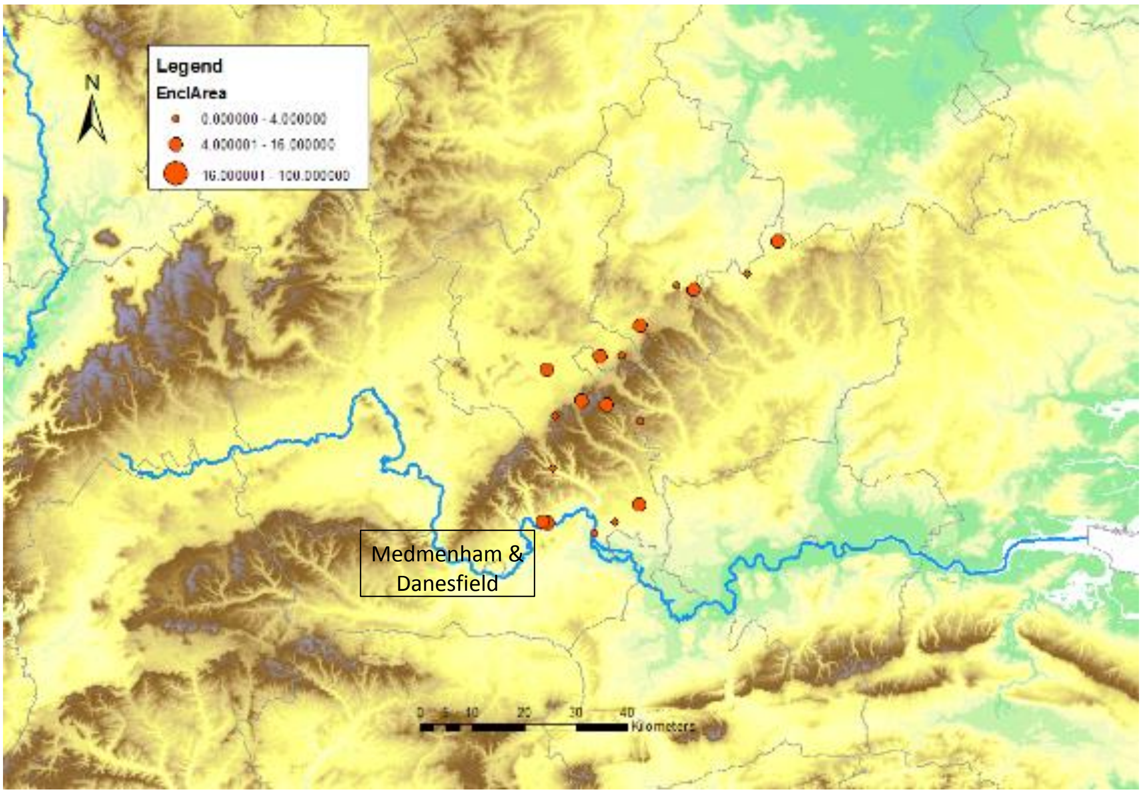


Figure 71 - Chilterns 12km percolation radius; Enclosed Area

Between percolation radii of 12 and 14km another topographically distinct cluster appears in the Chilterns. The sites are mainly along the north-west escarpment, but with the twin sites of Medmenham (Atlas: 3444) 6.6ha and Danesfield (Atlas: 3446) 7.5ha located directly on the Thames. At the larger radius the large Bozedown Camp (Atlas: 0139) 28.5ha sits directly on the Thames at the Goring Gap. Although the position of the sites along the Thames is highly significant, it is perhaps less convincing in this case that they played a function in trade for the sites along the north facing Chiltern escarpment. However a better understanding of the economic activities of the area, and inclusion of other rivers would help. Nonetheless, as with the other clusters in this chapter, it has unquestionably identified a distinct group for further detailed investigation.

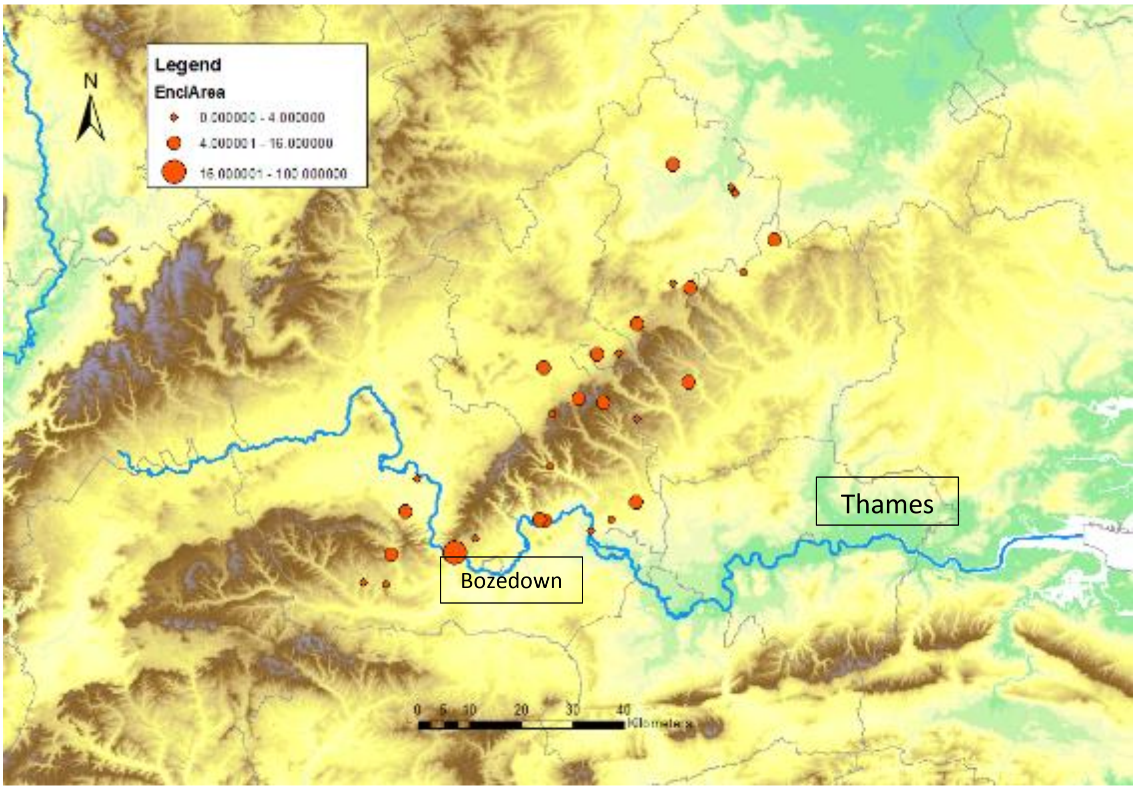


Figure 72 - Chilterns cluster 14km percolation radius; Enclosed Area

Conclusions

This study has investigated the spatial distribution of hillforts in the British Isles, and has excluded other classes of Iron Age site. The development of Percolation Analysis to identify clusters has yielded very interesting results in terms of identifying groups that are in many cases geographically distinct and map well onto topographically defined regions. Four of these in England and Wales were selected for more detailed analysis.

Some caution does however need to be exercised. With the Atlas inclusion requirement of hillforts being dominant, the study does perhaps risk a degree of circularity in terms of hillfort location being topographically significant; further study with datasets including other classes of Iron Age monuments, such as that carried out by Hogg (1972), would potentially address this.

A major weakness of course is that no account has been taken of site dating. The data and the clusters generated reflect the final state of hillfort construction, and do not necessarily represent the situation that may have prevailed in earlier times. However the approach is still valid, and lends itself readily to more refined analyses with such dating information as exists, and of course can be run repeatedly over time as more data becomes available.

Percolation Analysis is not a magic tool to elicit the past, but what is indeed clear from this study is that it generates the starting point for more detailed work, which should then focus on the details of sites and incorporate other sources of data. It hints at possible prehistoric groupings and cultural/ socio-political entities, but as always detailed investigation on a case by case basis is required, as shown so well by Aileen Fox (1958) for example. Potentially this can complement, reinforce or guide studies based on other types of evidence, such as coins, pottery, building styles and so forth, such as those explored in great detail by Cunliffe (1991).

Mapping sites using the enclosed area for symbol scaling has highlighted patterns which are regionally quite distinct. Ireland in particular has a distributed pattern of the largest sites at a global level, which suggest that they possibly played a major or central role in regional territories, and this would be a very interesting topic to explore in detail at a smaller regional scale, comparing with clusters as well as other data sources.

For the specific clusters that were selected for more detailed analysis, it does seem that the larger sites were located to play a role in the control of trade for the cluster, with in many cases their location being on or close to key communication paths, such as rivers, passes or transhipment routes, consistent with the role posited for hillforts by Sherratt (1996) and also by Brown (2008) for Wales. In some cases, these communication paths still play a role in modern times, such as the M50 Severn crossing.

Comparison with other data sets has also been illuminating, particularly that with Domesday shires and PAS finds. The former suggests at least in some areas a long continuity of cultural and/or socio-political entities, some of which still exist in modern England. In the case of the PAS finds, it very clearly demonstrates that Iron Age life in England was most certainly not restricted to the Hillfort Zone, and indeed suggests that perhaps a lot more was going on in other areas, with the uplands being culturally more marginal (as argued by Bradley 2007), a conclusion that would not be surprising today.

The comparisons with Stone Circles and Ringforts are also susceptible to the question of dating. They do however show a degree of distinction between the geographical locations of these classes of site, which likely reflects their different function as well as changes in culture and the habitability of their respective zones over time. The technique can be readily applied to other classes of site, and would be really interesting for other types of Iron Age site, to better highlight the relative roles of hillforts within regional Iron Age socio-political entities.

Another objective of the study was to generate a toolset that could be readily taken up by others, and for application to other datasets. The writing of a program suite that is largely parameter driven, and broken down into functional modules has gone a long way to meeting this goal.

Many opportunities for further investigation have been identified. For Percolation Analysis, this includes exploration and comparison with DBScan, which might assist in identifying the 'linking' sites that bring clusters together as the radius is extended, and also establishing a metric for cluster 'robustness'. For linking sites, such as Towbury Hill (Atlas: 0766) on the Severn, it would enable them to be investigated as to having marginal or key roles. Related to this is the need to find a way of geographically labelling clusters, without relying on a relative rank. This would aid comparisons with other datasets, such as the Domesday sites for example, and other groupings over long periods of time.

Another aspect of clusters to explore is their relation to hillfort hierarchy as implied by enclosed area and rank relative to neighbours. This would be useful for high density areas as in Scotland and Wales. Additionally experimenting with limits on distance for neighbour ranking (rather than a numerical limit) could also be interesting for identifying local relationships and distribution factors. The use of differing thresholds for hillfort size could also be experimented with on a regional basis, as well as revisiting those defined originally by Rivet (1958) and later extended by Hogg (1972; 1975) would also be useful for comparison both globally and locally.

In terms of thresholds and transitions observed in percolation analysis, some comparisons with 'day's walk' distances for different terrains might help explain cluster transition distances in different regions.

There is a lot more scope for analysis of hillforts based on topography, and determining the local and regional factors in their location. This has been started for the four selected clusters, but there is considerably more work that could be done using terrain models, and experimenting with the dominance tools developed by e.g. Llobera (2012).

There is also much more to be done with different classes of site, as well as exploration of the different site attributes, such as ditches and entrances, either individually or in combination, as well as comparisons with other types of data such as pottery styles, and construction styles of sites, such as attempted by Hogg (1972), as well as the significant work by Cunliffe on communities and regions in the Iron Age (e.g. 1991).

In particular the role of hillforts identified by Sherratt (1996) in trade, as entrepots, or in points of control on transhipment routes, portages and rivers, as well as on the coast, and developed by Brown (2002; 2008) for Wales seems to be strongly supported by the investigations of four selected clusters, and a focused examination of their hillforts on the basis of size and relationships to terrain and major rivers. What has been added to those earlier works is the identification of possible groupings, in the form of the clusters. This is one of the more exciting outcomes of this study, and would also be worth wider ranging and deeper exploration, at multiple geographical scales.

Appendix A: Data sources & Data cleaning

Hogg's British Hill-forts

Figure 73 shows a sample scan from Hogg's index of British Hill-forts (1979).

10 km sq	NGR	1" map	1/50,000 map	Name of site	Classn.	Area	Reference
	049 492	108	116	Dinas Melin y Wig	M	4.8	Mer Hist No. 250 P
	055 477	108	116	Mynydd Rhyd-ddu	U	0.33	Mer Hist No. 278
	084 444	108-117	125	Caer Drewyn	U	3.2	Mer Hist No. 255 P (which suggests work is of two periods, the earlier 0.55 ha)
	096 456	108-117	125	Moel Fodig	U	0.28	Mer Hist No. 252 P
SJ 05	080 587	108	116	Pwll y Clai	U	0.40	E. Davies, Denbs 374 P (Date?)
	094 575	108	116	Pen y Gaer, Ffriddlas	X	.-	E. Davies, Denbs 238 . Name only
SJ 07	013 720	108	116	Bedd y Cawr	?	Small	E. Davies, Denbs 67
	063 765	108	116	Moel Hiraddug	M	c.10	Arch Camb 114 (1965) 172 P. E. Davies, Flint 96 P. Flints Hist Soc 19 (1961) 1-20 EP. Also E forthcoming. Multi-period
	095 708	108	116	Moel y Gaer, Bodfari	M	1.8	Arch Camb 114 (1965) 169 P. E. Davies, Flint 37 P

Figure 73 - Sample data scan from Hogg

The complete index was scanned to Adobe pdf files, approximately 20 pages at a time to keep it manageable. After some experimentation with various software packages, and the scanner's own functionality, the scans were processed through the Optical Character Recognition (OCR) feature of Microsoft OneNote 2010, generating Comma Separated Values (csv) files.

This process did not fully respect the source tabulation, and was restructured and cleaned using macros in Microsoft Excel 2010. Macros were also used to extract OS Grid references, and these in turn were used to generate the Latitude/ Longitude and British national XY coordinates for each site using an excellent web based batch coordinate converter (*UK Grid Reference Finder* 2011). A unique numeric identifier was also allocated to each site for simplicity of reference.

The Latitude/ Longitude data was then used to produce a Keyhole Mark-up Language (kml) file, containing the location and name of each site, enabling visual inspection and checking of the data using Google Earth. This enabled gross errors to be resolved, and was also used to identify candidate sites for 'citizen science' surveys in support of that particular element of the Atlas project (*An Atlas of Hillforts in Britain and Ireland, The Survey of Hillforts*). As the project progressed it enabled detailed inspection of sites in Google Earth, and the resolution of less obvious errors in the converted data, as well as a handful in the source index itself.

Hogg's classifications of hillforts are shown in Table 3 below. This classification was not incorporated in the project database, as it was not planned to use it, and partly because the data was not readily convertible into a set of data fields without a lot of manual effort. Significantly more detailed data was generated by the Atlas project for each site in any case.

Classification	***Description***
A	With Annexe, see under Nuclear sites below
C	Coastal
D	Dun, less than 0.1ha. Although a term normally used for Scottish sites, it was found useful elsewhere for sites of this size range
M	Multi-vallate
N	Nuclear forts – probably the result of a superposition of simpler enclosures of different periods
U	Univallate
W (suffix to M)	southwestern type of wide spaced multiple ramparts
X	Rejected by some more recent study, although earlier claimed as a hillfort

Table 3 - Hogg's hillfort classifications

The Atlas of Iron Age Hillforts

Data was extracted from the Atlas database by the Atlas team and provided in the form of a Microsoft .xlsx file.

It took approximately one month to clean up the data, most issues being the result of typographic and transcription errors made during entry, inevitable in such a large project with so many contributors. A similar process was applied as previously for the Hogg database, creating a kml file for inspection in Google Earth. This identified errors in grid coordinates, but also enabled a high level comparison with the Hogg data. The other analyses carried out in this study picked up missing and non-numeric data from numeric fields. Once this issue was identified, a systematic approach was taken to cleaning up numeric fields using Microsoft Excel functions. All data corrections and queries were raised with or fed back to the Atlas project team.

As described in the latter parts of the study, analysis based on the enclosed area of the site proved to be important, and a few sites were found to be missing this data. Values were then either extracted from the Hogg dataset, where available, or were estimated from within Google Earth, using the ruler tool to make an approximate calculation, based on an assumption of either circularity or ovality. This was less critical for smaller sites, but it was very important to ensure that the areas of all large sites were included in the analysis, as is clear in the latter parts of this study. All queries were raised with the Atlas project team, who were invaluably helpful throughout.

The Atlas database comprises 93 data fields, which are listed below. Note that the order of the fields is as exported and does not necessarily reflect the structure of the original database.

- HER
- Site Name
- Alternative Name
- Country Code
- SAM Number
- NMR Mapsheet
- NMR Unique Identifier
- HER PRN
- Atlas Number
- HER Second Identifier
- Summary Description
- Serial Number
- Date Created
- Date Modified
- Create Username
- Modification Username
- Record Completed
- Entrances_No Breaks
- Entrances_Breaks Comments
- Entrances_Original Entrances
- Entrances_Original Entrances Comments
- Entrances_Cheveaux
- Entrances_Cheveaux Comments
- Entrances_Cheveaux Summary
- Date Entry History
- Citizen Science
- Final Sign Off
- Reliability of Data
- Reliability of Interpretation
- Data Comments
- Interpretation Comments
- NGR
- X
- Y
- Latitude
- Logitude
- Current County Unitary Authority
- Historic County
- Current Parish Community Council Townland
- Comments on Condition
- Current Land Category
- Comments on Land Use
- Type of Hillfort Comments
- Topographic Position Comments
- Dominant Topographic Feature
- Altitude

- Site Date
- Comments on Site Dating
- PreHillfort Activity
- PreHillfort Activity Details
- PostHillfort Activity
- PostHillfort Activity Details
- Investigations Summary
- Interior Summary
- Water Source Inside
- Water Source Inside Comments
- Interior Features Surface Evidence Summary
- Interior Features Shown by Excavation Summary
- Interial Features Shown by Geophysics Summary
- Finds Summary
- Finds Summary Comments
- Ramparts Banks Walls Ditches Summary
- Enclosed Area
- Whole Site Footprint Area
- Enclose system visibility multiperiod
- Enclose system visibility multiperiod Comments
- Do Ramparts form a complete circuit
- Do Ramparts form a complete circuit comments
- Rampart Bank Wall Number NE Quadrant
- Rampart Bank Wall Number SE Quadrant
- Rampart Bank Wall Number SW Quadrant
- Rampart Bank Wall Number NW Quadrant
- Rampart Bank Wall Number Quadrant Comments
- Rampart Bank Wall Form Surface Evidence Comments
- Rampart Bank Wall Form Excavated Evidence Comments
- Evidence of Gang Working
- Ramparts Banks Walls
- Ditches
- Comments on Gang Working
- Number of Ditches
- Comments on Ditches
- Annex Present
- Annex Number of Ramparts
- Annex Number of Entrances
- Annex Number of Ditches
- Annex Summary
- Reference
- Deletion
- Boundary
- Boundary Type
- Boundary Type Comments
- Second Country Code
- Second HER
- Second HER PRN
- Second Current County

- Second Historic County
- Second Parish
- Aerial Summary
- Aerial Tick
- Dating Reliability
- Enclosed Area 3
- Enclosed Area 2
- Enclosed Area 4
- Guard Chambers
- Site Date Export
- Monument Condition Export
- Topographic Position Export
- Type of Hillfort Export
- Aspect Export
- Water Source Inside Export
- Interior Features Surface Evidence Export
- Interior Features Shown by Excavation Export
- Interior Features Shown by Geophys Export
- Fields Summary Export
- Rampart Bank Wall Form Surface Evidence Export
- Rampart Bank Wall Form Excavated Evidence Export
- Aerial Tick Export

Key information was extracted from this spreadsheet for the analyses; principally the coordinate data, site name and the unique Atlas index number, as well as a select set of numeric attribute and descriptive data, listed below:

Field name	***Description***	***Type***
AtlasId	Atlas Number	Numeric
Country	Country name	Text
County	County name	Text
Name	Site name	Text
TopoD	Topographic description	Text
StType	Site Type	Text
Aspect	Aspect (orientation)	Text
StDate	Site date information	Text
Ditches	Number of ditches	Numeric
NERamprt	Number of NE quadrant ramparts	Numeric
SERamprt	Number of SE quadrant ramparts	Numeric
SWRamprt	Number of SW quadrant ramparts	Numeric
NWRamprt	Number of NW quadrant ramparts	Numeric
Entrance	Number of original entrances	Numeric
EnclArea	Enclosed area, ha	Numeric
AreaNote	Notes on area (not an original field; notes describing source and computation of data if not given in the Atlas)	Text
WhlArea	Whole site area, ha	Numeric
Area2	Enclosed area 2, ha	Numeric

Area3	Enclosed area 3, ha	Numeric
Area4	Enclosed area 4, ha	Numeric
NGR	National Grid Reference	Text
X	X coordinate (UK or Irish system)	Numeric
Y	Y coordinate (UK or Irish system)	Numeric
StLat	Site Latitude	Numeric
StLong	Site Longitude	Numeric
Alt	Altitude, m	Numeric

The numeric fields were intended for analytical and statistical processing. The text fields were provided for information and display, for example Atlas Id and Name are displayed in the Google Earth kml files. The text was available for viewing without recourse to the original dataset, and was very useful, for example, within ArcGIS for display of data when highlighting a specific site.

Appendix B: kml files, Google Earth

Google Earth proved to be an invaluable tool for providing a basic check on the integrity of the site coordinate data, at a global level (for example, are sites bounded by the coast line?). It also proved extremely useful at a later stage in order to inspect individual sites, through selection and zooming in.

kml files were generated directly within the data Excel spreadsheets, using text functions to insert specific parameters, such as Site Id, Site name, latitude and longitude within the basic code syntax. The resulting code file was then exported as a text file, and a template kml header and footer added.

The file could then be opened directly within Google Earth. It is then possible to select individual or groups of sites to display. Care was taken in designing the kml code to minimise the clutter of names that can occur, and to highlight displayed parameters only when the cursor is in the vicinity of the particular site. Semi-opaque fill was also used, so that overlapping site markers could be identified, and different colours were used for the Hogg and the Atlas data, so that the two could be readily compared.

Code of the file header for the Atlas data is shown below:

```
<?xml version="1.0" encoding="UTF-8"?>
<kml xmlns="http://www.opengis.net/kml/2.2">
<!-- this kml formats placemarks with a partially (ca. 80%) opaque paddle. When hovering over it with the mouse,
then the format changes to 100% opacity, and the placemark name is also displayed -->

<Document>
      <Style id="highlightPlacemark">
          <IconStyle>
            <Icon>
              <href>http://maps.google.com/mapfiles/kml/paddle/wht-blank.png</href>
            </Icon>
          <color>5078FF00</color>
          <!-- green -->
          </IconStyle>
          <LabelStyle>
              <scale>1</scale>
```

```
        </LabelStyle>
    </Style>
      <Style id="normalPlacemark">
        <IconStyle>
          <Icon>
            <href>http://maps.google.com/mapfiles/kml/paddle/wht-blank.png</href>
          </Icon>
        <color>5578FF00</color>
        <!-- green -->
        </IconStyle>
        <LabelStyle>
            <scale>0</scale>
        </LabelStyle>
      </Style>
    <StyleMap id="AtlasStyleMap">
        <Pair>
          <key>normal</key>
          <styleUrl>#normalPlacemark</styleUrl>
        </Pair>
        <Pair>
          <key>highlight</key>
          <styleUrl>#highlightPlacemark</styleUrl>
        </Pair>
      </StyleMap>
```

For each location, a placemark as the following example was generated from the source data. The name field is the Atlas Id and name companded, and the location is defined with the longitude and latitude, in that order:

```
<Placemark><styleUrl>#AtlasStyleMap</styleUrl><name>0002 Bach Camp</name><Point><coordinates>
-2.66479,52.23812,0</coordinates> </Point></Placemark>
```

The Hogg data kml code differed only in the colour selected for the icon.

Appendix C : Percolation Analysis

As a neat visual explanation of percolation analysis, the illustration of the City Clustering Algorithm from Rozenfeld, Rybski, Gabaix et al. (2011, 2208) in Figure 74 and its caption are very helpful.

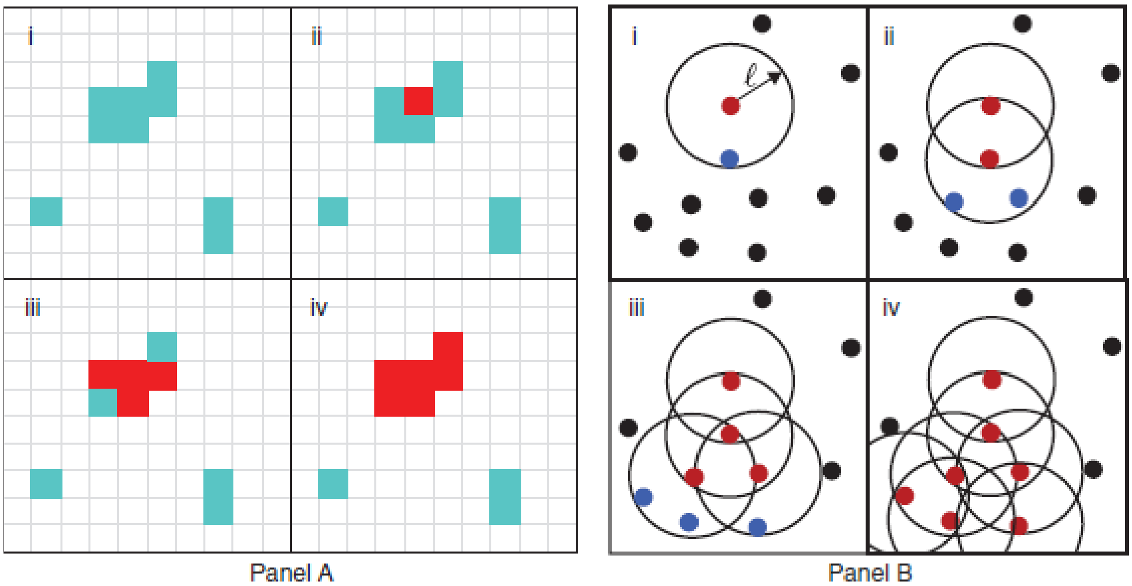


FIGURE 2. THE CCA ALGORITHM FOR GB AND THE US

Notes: Panel A shows the CCA applied to GB (discrete CCA). (i) Cells are colored in blue if they are populated; otherwise they are in white. (ii) We initialize the CCA by selecting a random populated cell (red cell). Then, we merge all populated neighbors of the red cell as shown in (iii). We keep growing the cluster by iteratively merging neighbors of the red cells until all neighboring cells are unpopulated, as shown in (iv). Next, we pick another populated cell and repeat the algorithm until all populated cells are assigned to a cluster. Panel B shows the CCA applied to the US (continuum CCA). The points in this figure denote populated sites or FIPS. For our studies (and in this diagram) we use a density threshold $D_* = 0$. (i) We start a cluster selecting a populated site, red point, among all available populated sites. We draw a circle of radius ℓ and add all populated sites, blue point, that fall within the circle. (ii) We draw a circle from the new member of this cluster and add all populated sites (denoted by the two blue points) within the circle. (iii) Recursively, we keep drawing circles from all new cluster sites. The populated sites inside the circles (three blue points in this case) are merged into the cluster. (iv) The red points are the members of the cluster. Since no black point is at a distance smaller than ℓ from any red point, the cluster does not grow anymore. We start the process again, selecting another initial point that has not been already assigned to any cluster. This process is repeated until all populated sites are assigned to a cluster.

Figure 74 - City Clustering Algorithm

The code of the key percolation cluster programs is included below. The objective was to create a general tool to be readily useable for other databases, with minimal recoding. To this end key process parameters, and the identity of code files to be used were defined within two text files read in at initialisation.

As an example of the ease of processing, when the Irish data became available, it was processed and maps generated within an hour or so.

The cluster generation and analysis programs were created as a series of individual programs that each carried out particular processes. This enabled rapid code development for each element of the overall process, made the processing time of some of the routines manageable, but importantly meant that reprocessing of particular data sets for particular purposes was possible without having to run the whole suite from the beginning. These could undoubtedly be combined with a suitable user interface with more effort, but the suite was efficient and fit for purpose, allowing effort to be focused on the primary task in hand.

The original suite of programs was provided by Elsa Arcaute, written in 'R' for the analysis of Domesday sites, in collaboration with data provided by Stuart Brookes. This suite had a similar structure to below, but has been very extensively modified to meet the objectives of this study, and to provide a generic, parameter driven tool.

The programs for the percolation cluster generation and analysis suite are as follows:

1. **Inter-site distance matrix** – given the dataset of sites and XY coordinates, this computes a partial matrix of inter-site distances. It was obvious after initial experiments that clusters appear only below a certain threshold of radius. In Britain for example around 40km. There was therefore little point in creating a full matrix of inter-site distances, which would become very large and slow to process. The limit was set by a parameter in the configuration file. The inter-site distance is computed using Pythagoras' theorem.
2. **Cluster extraction** – once the inter-site distance matrix has been created, this program identifies all clusters for a specific radius. It steps through a range of radii, the range being set in a configuration file. The data is output as a text csv file. This program is the heart of the analysis; the method used of creating a graph where nodes are connected when the distance between them is the given radius or less, then extracting each cluster, remains as written by Elsa Arcaute.
3. **Mapping clusters** – this program maps clusters for each radius on an outline map overlay. The same data can for example be mapped on a coastal outline, an outline with modern counties, or an outline with historic counties. The outline is defined by an ESRI shape file, and is defined in the configuration file by a text string. The outputs generated were in a multi-page pdf file, but this proved unwieldy, and it was changed to generate png files, which are much easier to import into Microsoft Word and other applications, and to display. Point data is also exported as an ESRI shape file for import directly into ArcGIS. This program also evolved significantly as the study progressed, to obtain clear and good quality outputs for the various different sets of data used.

4. **Percolation cluster size and rank analyses** – this processes the cluster data generated by the cluster extraction to give various different frequency and ranking plots. This program evolved significantly as various different statistical analyses were experimented with, under the guidance of Mark Lake. The final version reflects the latest work by Elsa Arcaute. This program was written by the author.
5. **Random site generation** – not strictly related to percolation, this program generates sites located randomly within a given map outline. The number of sites is determined by a dataset (e.g. the sites defined by Hogg or the Atlas), and the outline by a shape file. Both are set in the generic configuration file. The data generated is in a text csv and an ESRI shape file, and can then be input into the percolation process from the start. This program was written by the author.

To keep the data manageable, a series of folders were set up from the core directory, to hold the source data (i.e. coordinate datasets), working data (i.e. the text output files), maps (i.e. the output from the mapping routines), the map outline shape files, and the data analysis outputs. Two configuration files are used: one to define the source data to be used, and one for the computation parameters for the percolation analysis. Examples are shown below:

source_file.txt

```
source_file,map_name,shape_file_name,DTM_file_name
Confirmed_Sites.GB_IoM.160518.csv,Britain,greatbritain.shp,
UK_hw.asc
```

The source file is the site coordinate data file, the map name is for inclusion within the generated maps. A terrain model is also included, although only a very limited amount of analysis was done with this.

radius_values.txt

```
upper_radius,lower_radius,step_value,limit,upper_threshold,
lower_threshold
40,2,1,100,1,5
```

The upper and lower radii are the limits for the range of percolation cluster generation. The program steps between these two values generating clusters, using the given step value. The limit is the maximum distance for inclusion in the inter-site distance table. The upper and lower thresholds are values included in the analysis program for a computation described an earlier version of the work by Elsa Arcaute, but no longer used. All distances are in km, the XY coordinates in the datasets are to the metre.

Inter-site distance matrix

```
# CREATE NODES LIST SCRIPT
# Original code from Elsa Arcaute, CASA, UCL 28.2.15; extensively modified by Simon Maddison.
# Cite as:  arXiv:1504.08318 [physics.soc-ph]
# This code creates the file nodes_list_d.txt containing the list for all the distances between pairs,
# Distances are computed for each node with each successive node, working through the list.
# This uses an approach that creates file with only distances below a limit, e.g. 100km.
# that is to say a partial matrix of distances between nodes.
# This will then not be a fully interconnected network;
#     need to ensure that this does not adversely affect clustering computation.
# Note original used the bigmemory library but this is not available for windows, so all refs removed.
# Version 3.1: 29/3/16 - add check for null values, display and strip out, also update on duplicate coordinates.
#     This data written out to text files.
# Version 4.1: 4/4/2016 -
#     NOTE: problems with a matrix with #REF entries in .csv -
#     No check for this, and caused problems with matrix manipulation
#     Removed conversions to and from matrix data format for data, from original source. Not necessary
# Version 4.2: 10/4/16 - Minor error in duplicate computation corrected. Old ref to matrix_xy; note also
#     that input data now has additional Lat Long fields added, which are not used here
# Version 4.3: 6/5/16 - Change index name in source file to bring within character limit in ArcGIS -> PlcIndex
# Version 4.4: 8/5/16 - Change of source text file format to include map data
# Version 4.5: 18/5/16 - Change to include output PlcIndex as csv, cleaned up list with only those processed

setwd("C:/Users/Simon/Documents/Iron_Age_Hillforts/Percolation")
# paths
path_source <- paste(getwd(),"/source_data",sep="")
path_results <- paste(getwd(),"/working_data",sep="")

# Read in source file name defined in file source_file.txt
file_name <- paste(path_results,"/","source_file.txt",sep="")
source_files <- read.csv(file_name,header=TRUE,stringsAsFactors=FALSE)
source_file_name <- source_files$source_file
# Edit this file for different input file name

source_file <- paste(path_source,"/",source_file_name,sep="")

# Read in limit value from source file
# Edit the source file radius_values.txt to change this.
file_name <- paste(path_results,"/","radius_values.txt",sep="")
```

```
radius_values <- read.csv(file_name,header=TRUE)
limit <- radius_values$limit

data <- read.csv(source_file)
# This file needs three columns: PlcIndex, easting, Northing in this order.
# Note that the Easting and Northing need to be grid references with any Alphabetic
#  grid square converted to a numeric prefix for each
# List any entries with null values and write to file
data_NA <- data[rowSums(is.na(data)) > 0,]
print(paste("Entries in source file with one or more null values: "))
data_NA
file_name <- paste(path_results,"/","null_entries.txt",sep="")
write.table(data_NA, file_name, row.names=FALSE)
# Remove rows with null values from data
data <- na.omit(data)

# Remove points that are superimposed and keep only the first ID
# - This removes one of two sites that are very close to each other
# Determined on basis of x y coordinates
# Write list to file
duplicate_xy <- data[duplicated(data[,2:3]),]
data_unique <- data[!duplicated(data[,2:3]),]
print(paste("Number of removed superimposed points: ",(nrow(data)-nrow(data_unique))))
duplicate_xy
file_name <- paste(path_results,"/","duplicate_entries.txt",sep="")
write.table(duplicate_xy, file_name, row.names=FALSE)

x_vec <- data_unique$Easting
y_vec <- data_unique$Northing
ID <- data_unique$PlcIndex

# Write file with list of PlcIndex used
PlcIndex_list <- matrix(ID,ncol=1)
file_name <- paste(path_results,"/","PlcIndex.csv",sep="")
write.table(data_unique$PlcIndex, file_name, row.names=FALSE, col.names="PlcIndex")

# Number of points/nodes in file with no duplicates
n <- length(ID)
print(paste('number of points: ',n))
```

```
# Create matrix of internodal distances
# Columns: NodeId1, NodeId2, distance 1-2
col_list <- cbind('ID1','ID2','d12')
ni <- n-1
nj <- n
n_rows <- n*(n-1)/2
nodes_list <- matrix(, nrow = n_rows, ncol = 3)
colnames(nodes_list) <- col_list

row <- 0
i <- 1
for (i in 1:ni)
{
      j1 <- i+1
      for (j in j1:nj)
      {
            # Compute distance between nodes - pythagoras
            #  for full grid references this is in units of 1m
            d <- sqrt((abs(x_vec[i]-x_vec[j]))^2+(abs(y_vec[i]-y_vec[j]))^2)
            # to give distance in km, rounded to 2 decimal places; this also reduces file size
            d <- d/1000
            d <- round(d,2)
            # Include only if less than limit of distances to be included
            if(d < limit)
                  {row <- row+1
                  nodes_list[row,] <- cbind(ID[i],ID[j],as.numeric(d))
            }
      }
}
# Output file name and location writes to a text file
file_name <- paste(path_results,"/","nodes_list_d.txt",sep="")
# Remove the unused rows in the matrix
nodes_list <- nodes_list[-(row+1:n_rows),]
m_nodes <- as.matrix(nodes_list)
# need to write this WITHOUT the row number
write.table(m_nodes, file_name, row.names=FALSE)
print('matrix copied')
```

Cluster extraction

```
# CLUSTERING SCRIPT, RUNS AFTER NODES LIST SCRIPT
# Original code from Elsa Arcaute, CASA, UCL 28.2.15; extensively modified by Simon Maddison.
# Cite as:  arXiv:1504.08318 [physics.soc-ph]
# This code computes the clusters, based on matrix of inter-nodal distances computed in script
#     create_nodes_list_d.vxxx.R
# Note that the list of nodes and internodal distances is not a complete matrix, but only includes
#     distances below a certain threshold. This keeps the data sizes manageable.
# Version 3
# This is version 3, incorporating adjustments and enhancements to code
# Note that the original used library bigmemory but this is not available for windows so all refs removed.
# Note that when nodes are individual, they are no longer in a cluster. (31/1/16)
# Version 3.1: 20/04/16 - Note that this intrinsically cannot generate clusters with only one node.
#     Need to post process means in the analysis script
# Version 3.2: 6/5/16 - Changed Index name to fit within 8 chars -> PlcIndex
# Version 3.3: 18/5/16 - output data into a single file, indexed on PlcIndex and radius instead of separate
#     files. Read in PlcIndex.csv generated by create_nodes_list_d....R

library(igraph)
setwd("C:/Users/Simon/Documents/Iron_Age_Hillforts/Percolation")
# paths
path_source <- paste(getwd(),"/source_data",sep="")
path_results <- paste(getwd(),"/working_data",sep="")
data_file <- paste(path_results,"/","nodes_list_d.txt",sep="")

# Read in list of site Indices; file created by create_nodes_list_d.R
file_name <- paste(path_results,"/","PlcIndex.csv",sep="")
# This is used as the basis for the output file
mem_clust_by_r <- read.csv(file_name,header=TRUE)

# Read in distance thresholds - this ensures same values used in all scripts
# Edit the file radius_values.txt to change these.
file_name <- paste(path_results,"/","radius_values.txt",sep="")
radius_values <- read.csv(file_name,header=TRUE)
upper_radius <- radius_values$upper_radius
lower_radius <- radius_values$lower_radius
step_value <- radius_values$step_value

print(paste("Radii used for cluster analysis, km: upper ",upper_radius,
```

```
      "; lower ",lower_radius,"; step ", step_value))

# The data table of nodes and internode distances is a Text file, with headers
matrix_IDs_distance <- read.table(data_file,header=TRUE)
# Columns are: node Id 1, node Id 2, distance between them. Note that this is generated
#  with a limit to the maximum distance to reduce overall matrix size, and hence
#  creates a partial matrix

# Define range of percolation radius and step value to progressively reduce it
# Radius in km
radius_values <- seq(upper_radius,lower_radius,by=-step_value)

# Repeat for each radius value

for (i in radius_values)
{
      print(i)
      # Create sub-matrix such that all internode distances d12<=radius
      matrix_radius <- matrix_IDs_distance[matrix_IDs_distance[,3]<=i,]
      # Create graph (note that characters and numerics are treated differently)
      matrix2 <- matrix_radius[,1:2]
      matrix2[,1] <- as.character(matrix2[,1])
      matrix2[,2] <- as.character(matrix2[,2])
      # This creates a graph from the sub-matrix matrix2
      # Directed means that the edges will only be 'counted' once
      # In order to interpret matrix2 as a matrix added 'as.matrix'
      g <- graph.edgelist(as.matrix(matrix2), directed=TRUE)

      #take subcomponents - description of how this works
      #http://stackoverflow.com/questions/20725411/finding-strong-and-weak-clusters-and-their-membership-in-r

      # Identifies clusters in the graph; creates list of nodes and associated cluster id
      # weak refers to the mechanism used to generate the clusters and relates to computational efficiency
      # Note that this does not include clusters of 1 node,
      #  so the counts are not really meaningful at the lower limit
      member_of_cluster_id <- clusters(g, mode="weak")$membership
      # V processes vertices of graph
      m <- cbind(V(g)$name,member_of_cluster_id)
      new_col_name <- paste("ClstRad",i,sep="")
```

```
      colnames(m) <- c("PlcIndex",new_col_name)
      mem_clust_by_r <-  merge(mem_clust_by_r, m, by="PlcIndex", all.x=TRUE)

file_name <- paste(path_results,"/","member_cluster_by_radius.csv",sep="")
# Write file without row names
write.csv(mem_clust_by_r,file_name,quote=FALSE,row.names=FALSE)
```

Mapping Clusters

```
# CLUSTER MAPPING SCRIPT
# Original code from Elsa Arcaute, CASA, UCL 28.2.15; extensively modified by Simon Maddison.
# Cite as:  arXiv:1504.08318 [physics.soc-ph]
# Version 4: 24/3/16. Changes to plots made to:
#     Include identity of clusters (i.e.cluster index) in legend - done
#     Plot the source of the original data (e.g. Atlas, Hogg etc.) (taken from source file name)- done
#     Parameterise the shape file to be used for plotting ...
#     Take out graded grey colours for clusters > 15
# Version 4.1: 29/3/16 changed outputs to be multi-page pdf
#     NOTE: Ireland not plotting points TO BE RESOLVED
# Version 4.2: 6/4/16 - error in plot of all site corrected; removed 'xy'; all refs now to xy_data
# Version 4.3: 10/4/16 - output shape file with all data points, with same projection as input map
#     NOTE: esri shape file driver truncates names to 10 characters but this should not be a
#     problem at the moment. Prob of e.g. PlacesIndex, ClusterNumber etc.
#     shape file name is name of source file without .csv etc. extension
# Version 4.4: 29/4/16 - paging upgraded to export to A4 and use the full page
# Version 4.5: 6/5/16 - changed Index name to fit within 8 chars -> PlcIndex
# Version 4.6: 8/5/16 - working on Ireland plot to fix. Problem seemed to be an old Ireland map with
#     an obsolete coordinate system
#     Add in external input to country and shape file for source,
#     file is added to source_file.txt - map name and shape file name, headers also added
#     source_file, map_name, shape_file
#     REQUIRES CHANGE TO ALL OTHER PROGRAMS (done)
# Version 4.7: 18/5/16 -  Cluster data now comes in a single file (member_cluster_by_radius.csv)
#     output cluster ranks as new file site_cluster_rank_by_radius, lists by PlcIndex the cluster rank
#     indexed by radius
# Version 4.8: 30/7/16 - Plot as png file output. Earlier issues with resolution resolved.
#     Easier to include in word document than pdf. However each plot is a separate file,
#     with suitable name increment
```

```
# Version 4.9: 03/08/2016 - minor change to put legend on right for Domesday counties

library(maptools)
library(rgdal)
library(stringr)

setwd("/Users/Simon/Documents/Iron_Age_Hillforts/Percolation")
# paths for results and output maps
path_source <- paste(getwd(),"/source_data",sep="")
path_results <- paste(getwd(),"/working_data",sep="")
path_maps <- paste(getwd(),"/maps",sep="")
path_shape <- paste(getwd(),"/shape_files",sep="")

# File is the list of nodes and x y grid references
# Read in source file name defined in file source_file.txt
# Edit this file for different input file name
file_name <- paste(path_results,"/","source_file.txt",sep="")
source_files <- read.csv(file_name,header=TRUE,stringsAsFactors=FALSE)
source_file_name <- source_files$source_file
source_file <- paste(path_source,"/",source_file_name,sep="")
map_name <- source_files$map_name
shape_file_name <- source_files$shape_file

# Load shape file
file_name <- paste(path_shape,"/",shape_file_name,sep="")
# Truncate shape file name to remove extension
layer_name <- substr(shape_file_name,1,(nchar(shape_file_name)-4))
print(map_name)
map_outline <- readOGR(dsn=file_name,layer=layer_name)

# Read in nodes and grid coordinates
xy_data <- read.csv(source_file, as.is = FALSE)
# fields of interest:
# PlcIndex Easting Northing

# Convert data to SpatialPointsDataFrame, using project4string from selected map file
# get projection string for current map
projection_string <- proj4string(map_outline)
crs_projection <- CRS(projection_string)
```

```
# turn data into SpatialPointsDataFrame
coordinates(xy_data) <- c("Easting", "Northing")
proj4string(xy_data) <- crs_projection

# Read in data file with cluster id for each site and radius
file_name <- paste(path_results,"/","member_cluster_by_radius.csv",sep="")
ranked_mem_clust_by_r <- read.csv(file_name, header=TRUE)

# Read in distance thresholds - this ensures same values used as in clustering script
file_name <- paste(path_results,"/","radius_values.txt",sep="")
radius_values <- read.csv(file_name,header=TRUE)
upper_radius <- radius_values$upper_radius
lower_radius <- radius_values$lower_radius
step_value <- radius_values$step_value

radius_values <- seq(upper_radius,lower_radius,by=-step_value)

# Define colours for clusters, top 15 and then grey for the rest
top15_colours <- colors()[c(553,29,258,654,91,115,456,122,48,8,149,86,102,40,12)]
the_rest_colour <- "#AEAEAE"

# Plot map outline with all sites, uniform colour
file_map_png <- paste(path_maps,"/","percolation_plots_",map_name, "_all.png",sep="")
png(file=file_map_png, units="cm", width=21, height=29.7, res=300)

plot(map_outline, col="white",border=TRUE)

# get from original file xy_data coords and create table
# Easting | Northing | ID | cluster | colour
if(nrow(xy_data) < 1000)
      {point_dia <- 0.8
      } else {
      point_dia <- 0.4}
points(xy_data$Easting, xy_data$Northing, col='red', pch=20, cex=point_dia)
number_of_sites <- paste("Number of sites: ",nrow(xy_data))
mtext(number_of_sites)
plot_title <- str_to_title(map_name,locale="")
title(paste("Iron Age Hillforts in ", plot_title), sub=paste("Sources: ",source_file_name,"; ",shape_file_name))
dev.off()
```

```
# Generates maps for each of percolation radii in range of radius values
for(i in radius_values)
{
      file_map_png <- paste(path_maps,"/","percolation_plots_",map_name, "_rad_", i, ".png",sep="")
      png(file=file_map_png, units="cm", width=21, height=29.7, res=300)
      # Uses data for percolation radius computed, steps through each column for radius value i
      ClstRad_col <- paste("ClstRad",i,sep="")
      # extract data for this radius from cluster member by radius
      member_cluster <- ranked_mem_clust_by_r[c("PlcIndex", ClstRad_col)]
      colnames(member_cluster) <- c("node","cluster")

      # Ranks clusters by size, i.e. number of nodes for each cluster
      #  creates the number of instances for each cluster id ...
      frequency_clusters <- data.frame(table(member_cluster$cluster))
      names(frequency_clusters) <- c('cluster','number_of_points')
      ranked_clusters <- frequency_clusters[order(frequency_clusters$number_of_points, decreasing=TRUE),]

      # Number of nodes total for this radius
      total_nodes <- sum(ranked_clusters$number_of_points)
      # The number of clusters for this radius
      total_clusters <- nrow(ranked_clusters)
      print(paste("For radius: ",i," there are: ",total_clusters," clusters and ",total_nodes," nodes"))

      # Pick top 15 (if there are as many as this)
      number_colours <- 15
      if(total_clusters < 15)
      {
            number_colours <- total_clusters
      }
      # Giving colours to the top 15 or the top ones when less than 15
      ranked_clusters$col = ''
      ranked_clusters$col[1:number_colours] <- top15_colours[1:number_colours]
      if(total_clusters>15)
      {
            # add colours of grey for the rest
            ranked_clusters$col[16:total_clusters] <- the_rest_colour
      }
```

```
	#get sub-data from original xy_data
	xy_at_d <- xy_data[xy_data$PlcIndex %in% member_cluster$node,]

	#first add colour column to member_cluster then to xy_at_d
	member_cluster$col <- ranked_clusters$col[match(member_cluster$cluster, ranked_clusters$cluster)]
	xy_at_d$col <- member_cluster$col[match(xy_at_d$PlcIndex, member_cluster$node)]

	plot(map_outline, col="white", border=TRUE)
	# get from original file xy_data coords and create table
	# Easting | Northing | ID | cluster | colour
	# this following to check all points included
	# plot all points in grey, as background
	points(xy_data$Easting, xy_data$Northing, col='grey85', pch=4, cex=.3)
	# plot
	points(xy_at_d$Easting, xy_at_d$Northing, col=xy_at_d$col, pch=20, cex=point_dia)
	if(plot_title=="Ireland" | plot_title=="Domesday")
	{legend_loc <- "bottomright"
	} else {
	legend_loc <- "bottomleft"}
	legend(legend_loc, inset= .01, title="Rank Cluster#",legend=ranked_clusters$cluster[1:number_colours], cex=.5,
	      fill=ranked_clusters$col[1:number_colours])

	title(main=paste("Iron Age Hillforts in ", plot_title), sub=paste("Source File: ",source_file_name,
	      "; percolation distance:",i,"km"))

	# Create data for output in csv file
	# create ranking index for cluster numbers
	ranked_clusters$rank <- 1:nrow(ranked_clusters)
	ranked_clusters$number_of_points <- NULL
	ranked_clusters$col <- NULL
	new_col_name <- paste("RankRad", i, sep="")
	colnames(ranked_clusters) <- c("cluster",new_col_name)
	# add cluster ranking as column into data frame with clusters and PlcIndex
	ranked_mem_clust_by_r <- merge(ranked_mem_clust_by_r, ranked_clusters, by.x=ClstRad_col, by.y="cluster",
	       all.x=TRUE)
	# Remove original cluster column for this radius
	ranked_mem_clust_by_r[,ClstRad_col] <- NULL
	dev.off()
}
```

```
#dev.off()

# convert NA's in ranked_mem_clust_by_r to zero
ranked_mem_clust_by_r[is.na(ranked_mem_clust_by_r)] <- 0
# Write file with PlcIndex and cluster rank for each radius
# Output file name and location write to a text file
file_name <- paste(path_results,"/","site_cluster_rank_by_radius.csv",sep="")
# need to write this WITHOUT the row number
write.csv(ranked_mem_clust_by_r[order(ranked_mem_clust_by_r$PlcIndex),], file_name, row.names=FALSE)

# Write shape file with all data points
# First convert data to SpatialPointsDataFrame, using project4string from selected map file
# get projection string for current map
projection_string <- proj4string(map_outline)
crs_projection <- CRS(projection_string)

# name of output shapefile
# Truncate source file name to remove extension
source_file_name <- substr(source_file_name,1,(nchar(source_file_name)-4))
output_shape_file <- paste(path_maps,"/",source_file_name,".shp",sep="")
writeOGR(xy_data,output_shape_file,layer=source_file_name,driver="ESRI Shapefile",overwrite_layer=TRUE)
```

Percolation cluster size and rank analyses

```
# CLUSTER FREQENCY PLOTTING SCRIPT
# Simon Maddison, 10th May 2015
# Version 2.12: 30/7/16

# This script extracts cluster frequency data and plots it for a range of percolation radii
# Other graphical/ statistical analyses added as study progressed
# Computation and plot of cluster rank size statistics, as per Arcaute latest paper
# The data is in the working directory in text files;
#   each file is a list of Place indices, with the identity of each cluster which it is a member of
setwd("C:/Users/Simon/Documents/Iron_Age_Hillforts/Percolation")
# paths
path_source <- paste(getwd(),"/source_data",sep="")
path_working <- paste(getwd(),"/working_data",sep="")
path_results <- paste(getwd(),"/analysis_results",sep="")
```

```
library(calibrate)

# Read in source file name defined in file source_file.txt
file_name <- paste(path_working,"/","source_file.txt",sep="")
source_files <- read.csv(file_name,header=TRUE,stringsAsFactors=FALSE)
source_file_name <- source_files$source_file
file_name <- paste(path_source,"/",source_file_name,sep="")

# Read source file to establish number of nodes (entries in the file)
# subtract 1 for header line
total_nodes <- length(readLines(file_name)) -1

# Reads in file of data giving cluster ids indexed by radius, for each node
file_name <- paste(path_working,"/","member_cluster_by_radius.csv",sep="")
mem_clust_by_r <- read.csv(file_name, header = TRUE)

# Read in distance thresholds - this ensures same values used in all scripts
# Edit the source file radius_values.txt to change these.
file_name <- paste(path_working,"/","radius_values.txt",sep="")
radius_values <- read.csv(file_name,header=TRUE)
upper_radius <- radius_values$upper_radius
lower_radius <- radius_values$lower_radius
step_value <- radius_values$step_value
upper_threshold <- radius_values$upper_threshold
lower_threshold <- radius_values$lower_threshold

radius_values <- seq(lower_radius,upper_radius,by=step_value)

# Create matrix of number of clusters and number of nodes (max, mean, median, S_mean), for each radius
# NOTE: S_mean statistic superseded in Arcaute's later paper, different plot generated as well
# Columns:  Radius, number of clusters, max cluster size, mean cluster size, median cluster size, S mean
col_list <-
cbind('radius','num_clust','max_clust_size','max_normalized','mean_clust_size','median_clust_size','S_mean')
n_rows <- upper_radius - lower_radius +1
analysis_by_radius <- matrix(, nrow = n_rows, ncol = 7)
colnames(analysis_by_radius) <- col_list
row <- 1

for(i in radius_values)
```

```
{
	# Reads in data for each percolation radius that has been computed
	# This lists each node and the id of the cluster to which it is a member, for this radius
	ClstRad_col <- paste("ClstRad",i,sep="")
	member_cluster <- mem_clust_by_r[c("PlcIndex", ClstRad_col)]
	names(member_cluster) <- c('node','cluster')
	member_cluster <- na.omit(member_cluster)

	# number of nodes in clusters, i.e. excluding single nodes
	total_clustered_nodes <- nrow(member_cluster)
	# Ranks clusters by size, i.e. number of nodes for each cluster
	frequency_clusters <- data.frame(table(member_cluster$cluster))
	names(frequency_clusters) <- c('cluster','number_of_nodes')
	ranked_clusters <- frequency_clusters[order(frequency_clusters$number_of_nodes, decreasing=TRUE),]
	# Convert cluster column from factor to numeric, else problems later
	ranked_clusters$cluster <- as.numeric(as.character(ranked_clusters$cluster))
	number_of_clusters <- nrow(ranked_clusters)

	# add in clusters of size 1; compute number of these clusters
	add_row_total <- total_nodes - total_clustered_nodes
	if(add_row_total>0)
	{
		supp_array <- data.frame(seq((number_of_clusters+1),(number_of_clusters + add_row_total)))
		one_col <- data.frame(rep(1,times=add_row_total))
		supp_array <- cbind(supp_array,one_col)
		names(supp_array) <- c('cluster','number_of_nodes')
		# ranked_clusters now includes clusters of size 1
		# makes consistent measure of means medians for different percolation radii
		ranked_clusters <- rbind(ranked_clusters,supp_array)
	}
	# revised number of clusters
	number_of_clusters <- nrow(ranked_clusters)
	# The number of nodes in the first ranked cluster is maximum
	max_nodes <- ranked_clusters$number_of_nodes[1]
	max_normalized <- max_nodes/total_nodes
	mean_nodes <- mean(ranked_clusters$number_of_nodes)
	median_nodes <- median(ranked_clusters$number_of_nodes)

	# 'S mean' for clusters, excluding largest, and also excluding clusters of smaller size
```

```
	# These upper and lower thresholds defined in input radius_values text file
	# lower_threshold - read in from text file. Clusters with this number of nodes or less excluded
	# upper_threshold - read in from text file. Number of largest clusters to be excluded
	constrained_ranked_clusters <- ranked_clusters[ranked_clusters$number_of_nodes > lower_threshold,]
	constrained_number_of_clusters <- nrow(constrained_ranked_clusters)
	if((upper_threshold+1) <= constrained_number_of_clusters)
	{
		S_mean <-
		mean(constrained_ranked_clusters$number_of_nodes[(upper_threshold+1):constrained_number_of_clusters])
	} else {
		S_mean <- 0
	}

	analysis_by_radius[row,] <-
		cbind(i,number_of_clusters,max_nodes,max_normalized,mean_nodes,median_nodes,S_mean)
	row <- row + 1

	#plot(ranked_clusters$number_of_nodes,
	#	main=paste("Cluster rank; radius: ",i,"km"),
	#	sub=paste("Source File: ",source_file_name),
	#	xlab="cluster",ylab="number of nodes in cluster")
	#title(paste("Source File: ",source_file_name), outer=TRUE, line=3, cex=0.8)

}

#dev.off()
analysis_by_radius <- as.data.frame(analysis_by_radius)
# Save the analysis data as a csv file
file_name <- paste(path_results,"/","analysis_by_radius.csv",sep="")
write.table(analysis_by_radius,file=file_name,col.names=TRUE,sep=",",row.names=FALSE,quote=FALSE)

output_file <- paste(path_results,"/cluster_plots","_%02d.png",sep="")
png(file=output_file, units="cm", width=21, height=21, res=300)
#par(mfrow=c(1,1), oma=c(1,3,4,3))

plot(analysis_by_radius$num_clust,analysis_by_radius$max_clust_size,
	xlim=NULL, ylim=NULL,
	main="max, mean, median size of cluster vs #clusters",
	sub=paste("Source File: ",source_file_name),
```

```
      xlab="# clusters, showing radius value", ylab="cluster size")
textxy(analysis_by_radius$num_clust,analysis_by_radius$max_clust_size,analysis_by_radius$radius, col="red",
      cex=.8)
points(analysis_by_radius$num_clust,analysis_by_radius$mean_clust_size)
textxy(analysis_by_radius$num_clust,analysis_by_radius$mean_clust_size,analysis_by_radius$radius, col="blue",
      cex=.8)
points(analysis_by_radius$num_clust,analysis_by_radius$median_clust_size)
textxy(analysis_by_radius$num_clust,analysis_by_radius$median_clust_size,analysis_by_radius$radius, col="green",
      cex=.8)

# Plot mean and median on linear scales
plot(analysis_by_radius$num_clust,analysis_by_radius$mean_clust_size,
      xlim=NULL, ylim=NULL,
      main="mean (blue), median (green) size of cluster vs #clusters",
      sub=paste("Source File: ",source_file_name),
      xlab="# clusters, showing radius value", ylab="cluster size")
textxy(analysis_by_radius$num_clust,analysis_by_radius$mean_clust_size,analysis_by_radius$radius, col="blue",
      cex=.8)
points(analysis_by_radius$num_clust,analysis_by_radius$median_clust_size)
textxy(analysis_by_radius$num_clust,analysis_by_radius$median_clust_size,analysis_by_radius$radius, col="green",
      cex=.8)

# Plot mean on log log scale
plot(analysis_by_radius$num_clust,analysis_by_radius$mean_clust_size,
      xlim=NULL, ylim=NULL, log="xy",
      main="mean (blue) size of cluster vs #clusters, log-log",
      sub=paste("Source File: ",source_file_name),
      xlab="# clusters, showing radius value", ylab="cluster size")
textxy(analysis_by_radius$num_clust,analysis_by_radius$mean_clust_size,analysis_by_radius$radius, col="blue",
      cex=.8)

# Plot radius vs S_mean
plot(analysis_by_radius$radius,analysis_by_radius$S_mean,
      main="S_mean cluster size vs radius",
      sub=paste("Source File: ",source_file_name),
      xlab="radius km", ylab="S_mean cluster size")
lines(analysis_by_radius$radius,analysis_by_radius$S_mean, type="b")
mtext(paste("excluded: clusters ",upper_threshold," largest & with members <=",lower_threshold))
textxy(analysis_by_radius$radius,analysis_by_radius$S_mean,analysis_by_radius$radius, col="red", cex=.8)
```

```
# Plot radius vs max_clust_size
plot(analysis_by_radius$radius,analysis_by_radius$max_clust_size,
      main="Max cluster size vs radius",
      sub=paste("Source File: ",source_file_name),
      xlab="radius km", ylab="max cluster size")
lines(analysis_by_radius$radius,analysis_by_radius$max_clust_size, type="b")
textxy(analysis_by_radius$radius,analysis_by_radius$max_clust_size,analysis_by_radius$radius, col="red", cex=.8)

# Plot radius vs normalized max_clust_size
plot(analysis_by_radius$radius,analysis_by_radius$max_normalized,
      main="Max cluster size (normalized) vs radius",
      sub=paste("Source File: ",source_file_name),
      xlab="radius km", ylab="max cluster size (normalized)")
lines(analysis_by_radius$radius,analysis_by_radius$max_normalized, type="b")
textxy(analysis_by_radius$radius,analysis_by_radius$max_normalized,analysis_by_radius$radius, col="red", cex=.8)

dev.off()
```

Random site generation

```
# SYNTHETIC - RANDOM - SITES GENERATION
# Code generates sites synthetically within bounds of a defined area and plots out
# Sites data output for analysis and plotting with existing packages, csv and shape files
# Version 1: 26/04/2016
# Version 1.1: 6/5/16 - Changed name of Index in coordinates file to be within 8 chars
# Version 1.2: 8/5/16 - Changed source text file format to incorporate map data, read from file
# Version 1.3: 19/5/16 - Amended to print on full A4 page format

library(maptools)
library(rgdal)
library(sp)
library(rgeos)
library(stringr)

setwd("/Users/Simon/Documents/Iron_Age_Hillforts/Percolation")
# paths for results and outputs
path_source <- paste(getwd(),"/source_data",sep="")
```

```
path_results <- paste(getwd(),"/working_data",sep="")
path_maps <- paste(getwd(),"/maps",sep="")
path_shape <- paste(getwd(),"/shape_files",sep="")

# File is the list of nodes and OS grid references
# Read in source file name defined in file source_file.txt
# Edit this file for different input file name
file_name <- paste(path_results,"/","source_file.txt",sep="")
source_files <- read.csv(file_name,header=TRUE,stringsAsFactors=FALSE)
source_file_name <- source_files$source_file
source_file <- paste(path_source,"/",source_file_name,sep="")
map_name <- source_files$map_name
shape_file_name <- source_files$shape_file

# Load shape file
file_name <- paste(path_shape,"/",shape_file_name,sep="")
# Truncate shape file name to remove extension
layer_name <- substr(shape_file_name,1,(nchar(shape_file_name)-4))
print(map_name)
map_outline <- readOGR(dsn=file_name,layer=layer_name)

# Find number of sites in source file
# Read source file to establish number of nodes (entries in the file)
# subtract 1 for header line
total_nodes <- length(readLines(source_file)) - 1
# Generate random nodes and grid coordinates, within defined outline
#    this is a spatial data frame, with coordinate system from map outline
xy_data <- spsample(map_outline,total_nodes,type="random")

# plot all sites
file_map_pdf <- paste(path_maps,"/","random_sites_plot_",map_name,".pdf",sep="")
#pdf(file=file_map_pdf)
pdf(file=file_map_pdf, paper="a4", width=21/2.54, height=29.7/2.54)

plot(map_outline, col="white",border=TRUE)

# Plot random sites
if(total_nodes < 1000)
{point_dia <- 0.8
```

```
} else {
point_dia <- 0.4}
points(xy_data$x, xy_data$y, col='red', pch=20, cex=point_dia)
number_of_sites <- paste("Number of sites: ",total_nodes)
mtext(number_of_sites)
plot_title <- str_to_title(map_name,locale="")
title(paste("Random sites in ",plot_title), sub=paste("Number as in ",source_file_name))

dev.off()

# Write as csv file for other analyses
xy_data <- as.data.frame(xy_data)
row_names <- rownames(xy_data)
rownames(row_names) <- NULL
xy_data <- cbind(row_names,xy_data)
colnames(xy_data) <- c("PlcIndex","Easting","Northing")
curr_date <- format(Sys.Date(),"%y%m%d")
file_name <- paste(path_source,"/random_sites_",map_name,".",curr_date,".csv",sep="")
write.table(xy_data,file=file_name,col.names=TRUE,sep=",",row.names=FALSE,quote=FALSE)

# Write shape file with all data points
# First convert data to SpatialPointsDataFrame, using project4string from selected map file
# get projection string for current map
projection_string <- proj4string(map_outline)
crs_projection <- CRS(projection_string)
# turn data into SpatialPointsDataFrame
coordinates(xy_data) <- c("Easting", "Northing")
proj4string(xy_data) <- crs_projection

# name of output shapefile
file_name <- paste("Random_sites_",map_name)
output_shape_file <- paste(path_maps,"/",file_name,".shp",sep="")
writeOGR(xy_data,output_shape_file,layer=source_file_name,driver="ESRI Shapefile",overwrite_layer=TRUE)
```

Appendix D: Attribute Based Analysis, and Comparison with other Datasets

As described earlier in this study, various other analyses have been done, based on different datasets, and also on the basis of specific site attributes. Only Enclosed Area was used for this, generating site rankings for example, at different radii. This program is not designed for ready adaption to different datasets, and evolved quite quickly as different datasets were processed and specific outputs were required for this dissertation, so source file names are hard coded, for example. Experimentation was carried out with the various density plots, to get results that displayed well, and this was also hard coded rather than being parameter driven.

Kernel density and other data comparative analyses and plots

```
# KERNEL DENSITY AND RELATED PLOTS
# Simon Maddison, MSc GIS and Spatial Analysis in Archaeology, 2016
# This code generates kernel density and other plots based on various data sets for Iron Age Hillforts
# Outlines derived from shape files of coastlines
# Data from Atlas and other sources
# Version 1.0: 10/06/16 -  initial version
# Version 1.1: 18/06/2016 - amended method of relative risk analysis
# Version 1.2: 18/06/2016 -
#     - step through different rank thresholds
#     - run for neighbour count = 100, 75, 50; combine thresholds to single pdf
#     - also run for Ireland; tweaks on plots
# Version 2.0: 22/06/2016
#     - add in analysis of PAS
#     - add in calculation of 'point of neutrality' for relative risk plots
#     - add in colour plot for hillfort sites based on size rank with neighbours
#     - add in analysis of neolithic stone circles (Burl) from Damon Ortega
#     - repeat for Ireland with stone circles
# Version 2.1: 06/07/2016
#     - ranking not correct, by inspection of plots
#     - checks and changes to ranking computation to fix
#     - amended to prevent clipping of sites by map_outline where they fall on coast
# Version 2.2: 31/07/2016
#     - amended to output plots to png files - easier to incorporate in Word doc
#     - bug in relative rank plots by area, used last defined value of n_value
# Version 2.3: 03/08/2016
#     - minor changes to which sigma values used for Stone Circles and PAS
```

```
#     - change PAS to use GB shape file and overlay with english counties
#     - change colour and size of points for PAS and Stone Circles
#     - add nearest neighbour histogram without x limit
#     - additional comments
# Version 2.4: 11/08/2016
#     - added in computations for Irish Ring Forts, similar to that for Stone Circles
#     - core extraction code for database provided by Alex and Marta; database
#           is provided as a shape file
#     - for density plots, added in colour ramp that centres on point of neutrality
#     - for PAS analysis restrict sites to England and Wales
# Version 2.5: 04/09/2016
#     - minor edits to change nearest neighbour plot output names
#     - filter applied to PAS analyses to include only English hillfort sites
# Version 2.5: 04/09/2016
#     - minor edits to change nearest neighbour plot output names
#     - filter applied to PAS analyses to include only English hillfort sites
# =======================================================================
# Setup and configuration; read in data

library(spatstat)
library(maptools)
library(rgdal)
library(stringr)
library(data.table)

setwd("/Users/Simon/Documents/Iron_Age_Hillforts/Percolation")
# paths for results and output maps
path_source <- paste(getwd(),"/source_data",sep="")
path_results <- paste(getwd(),"/working_data",sep="")
path_plots <- paste(getwd(),"/kernel_d_plots",sep="")
path_shape <- paste(getwd(),"/shape_files",sep="")

# File is the list of nodes and x y grid references
# Read in source file name defined in file source_file.txt
# Edit this file for different input file name
file_name <- paste(path_results,"/","source_file.txt",sep="")
source_files <- read.csv(file_name,header=TRUE,stringsAsFactors=FALSE)
source_file_name <- source_files$source_file
source_file <- paste(path_source,"/",source_file_name,sep="")
```

```
map_name <- source_files$map_name
shape_file_name <- source_files$shape_file
# Truncate source file name for inclusion in output file names
source_file_name <- substr(source_file_name, 1, (nchar(source_file_name)-4))

# Load shape file
file_name <- paste(path_shape,"/",shape_file_name,sep="")
# Truncate shape file name to remove extension
layer_name <- substr(shape_file_name,1,(nchar(shape_file_name)-4))
print(map_name)
map_country <- map_name
map_outline <- readOGR(dsn=file_name,layer=layer_name)

# Read in nodes and grid coordinates
xy_data <- read.csv(source_file, as.is = FALSE)
# as.is is redundant as far as I can see as just confirms default behaviour of read.csv
# fields of interest:
# PlcIndex Easting Northing
# Number of sites:
num_sites <- nrow(xy_data)
# Indices are not continuous, highest index value:
max_site_index <- max(xy_data$PlcIndex)

# Convert data to SpatialPointsDataFrame, using project4string from selected map file
# get projection string for current map
projection_string <- proj4string(map_outline)
crs_projection <- CRS(projection_string)
# turn data into SpatialPointsDataFrame
coordinates(xy_data) <- c("Easting", "Northing")
proj4string(xy_data) <- crs_projection

# create map bounds of the map outline
map_bounds <- boundingbox(as.owin(map_outline))
plot(map_bounds)
plot(map_outline,add=TRUE)

# create outline window object
e_map <- as.owin(map_outline)
```

```
# sites is the same coordinate data, constrained by the map bounds
sites <- as.ppp(coordinates(xy_data), as.owin(map_bounds))
# note: using the map outline would likely remove some sites located close to the outline edge
# this would result in a distorted density plot
# therefore compute the density plots first and then clip with the coastline in map

print("preliminary setup complete.  Next step is density distributions")
readline(prompt="Press [enter] to continue")

dev.off()
# ---------------------------------------------------------------------------------------------#
# Generate site kernel density plots

# determine 'optimum' standard deviation (sigma) using Diggle
diggle_sigma <- bw.diggle(sites)
# Create kernel density plots for sites
# select sigma values to compute by trial and error; slow to do them all
sites_2500 <- density(sites, sigma=2500, edge=TRUE, eps=2500)
sites_6000 <- density(sites, sigma=6000, edge=TRUE, eps=2500)
#sites_8000 <- density(sites, sigma=8000, edge=TRUE, eps=2500)
#sites_10000 <- density(sites, sigma=10000, edge=TRUE, eps=2500)
#sites_12500 <- density(sites, sigma=12500, edge=TRUE, eps=2500)
# Create plot with diggle generated sigma, and eps for finer resolution than default
sites_diggle <- density(sites, sigma=diggle_sigma, edge=TRUE, eps=2500)

# Clip plots to actual map_outline
#e_map <- as.owin(map_outline)
crop_s_2500 <- sites_2500[e_map, drop=FALSE]
crop_s_6000 <- sites_6000[e_map, drop=FALSE]
#crop_s_8000 <- sites_8000[e_map, drop=FALSE]
#crop_s_10000 <- sites_10000[e_map, drop=FALSE]
#crop_s_12500 <- sites_12500[e_map, drop=FALSE]
crop_s_diggle <- sites_diggle[e_map, drop=FALSE]

# changed plots to png ... easier to incorporate in documents
# Plot Kernel Density Plots for Hillfort Sites
file_map_png <- paste(path_plots,"/",source_file_name,"_kD_%02d.png",sep="")
png(file=file_map_png, units="cm", width=21, height=21, res=300)
```

```
plot(crop_s_2500, main="", col=topo.colors(100), box=FALSE)
title("Hillfort kernel density", sub=paste("Sources: ",source_file_name,"; ",shape_file_name))
mtext("Sigma value: 2500")
plot(crop_s_6000, main="", col=topo.colors(100), box=FALSE)
title("Hillfort kernel density", sub=paste("Sources: ",source_file_name,"; ",shape_file_name))
mtext("Sigma value: 6000")
#plot(crop_s_8000, main="", col=topo.colors(100), box=FALSE)
#title("Hillfort kernel density", sub=paste("Sources: ",source_file_name,"; ",shape_file_name))
#mtext("Sigma value: 8000")
#plot(crop_s_10000, main ="", col=topo.colors(100), box=FALSE)
#title("Hillfort kernel density", sub=paste("Sources: ",source_file_name,"; ",shape_file_name))
#mtext("Sigma value: 10000")

plot(crop_s_diggle, main="", col=topo.colors(100), box=FALSE)
title("Hillfort kernel density", sub=paste("Sources: ",source_file_name,"; ",shape_file_name))
mtext(paste("Diggle sigma value: ",round(diggle_sigma, digits=0)))
#plot(crop_s_12500, main="", col=topo.colors(100), box=FALSE)
#title("Hillfort kernel density", sub=paste("Sources: ",source_file_name,"; ",shape_file_name))
#mtext("Sigma value: 12500")

dev.off()
# ----------------------------------------------------------------------------------------------#
# Nearest neighbour analysis and histograms
sites_dist <- nndist(sites)

file_map_png <- paste(path_plots,"/",source_file_name,"_nearest_neigbour_%02d.png",sep="")
png(file=file_map_png, units="cm", width=21, height=21, res=300)
hist(sites_dist, xlab="nearest neighbour distance",
      main="Hillfort: nearest neighbour histogram")
title("", sub=paste("Sources: ",source_file_name,"; ",shape_file_name))
hist(sites_dist, breaks=240, xlim=c(0, 20000), prob=TRUE, xlab="nearest neighbour distance",
      main="Hillfort: nearest neighbour histogram")
title("", sub=paste("Sources: ",source_file_name,"; ",shape_file_name))
hist(sites_dist, breaks=240, prob=TRUE, xlab="nearest neighbour distance",
      main="Hillfort: nearest neighbour histogram")
title("", sub=paste("Sources: ",source_file_name,"; ",shape_file_name))
dev.off()

# Record histogram data and Clark Evans test result
```

```
file_map_txt <- paste(path_plots,"/ClarkEvans_",source_file_name,".txt",sep="")
capture.output(file=file_map_txt, paste("Mean: ",round(mean(sites_dist),digits=-1)/1000))
capture.output(file=file_map_txt, append=TRUE, paste("Median: ",round(median(sites_dist),digits=-1)/1000))
# Clark-Evans test
capture.output(file=file_map_txt, append=TRUE, clarkevans.test(sites, corrections="none"))

print("next step is to generate relative ranking dependent on number of neighbours")
#readline(prompt="Press [enter] to continue")
# ---------------------------------------------------------------------------------------------#
# Create map plots based on area rank for Hillforts

# nodes_list_d lists sites pairwise with the distance between them up to computed threshold
file_name <- paste(path_results,"/nodes_list_d.txt", sep="")
node_distance_table <- read.table(file_name,header=TRUE)

# read in site area from key_params file; ignore other data
# Indexed by AtlasId; area in hectares
file_name <- paste(path_source,"/Confirmed_sites.Key_params.160610.csv", sep="")
# Read in site areas, according to AtlasId
site_area <- fread(file_name, header=TRUE, select=c("AtlasId","EnclArea"))

map_name <- paste(source_file_name, "_Hillforts_rank_area")
# Change to png file output, built into loop

# number of neighbours is n_value
for (n_value in c(100,75,50,30,20,15,10))
{
      # Establish site size rank relative to nearest 'n_value' neighbours
      # create object to rank site areas relative to neighbours
      site_area_n_rank <- xy_data
      # add columns for rank and for area - for visual inspection
      site_area_n_rank$area_rank <- 0
      site_area_n_rank$EnclArea <- 0

      # run for all sites in xy_data
      for (i in 1:num_sites)
      {
            # identity of the site = PlcIndex
            site_id <- xy_data$PlcIndex[i]
```

```
# Find the area of the site in question
area_i <- site_area$EnclArea[site_area$AtlasId == site_id]
# write to file, for ease of checking only
site_area_n_rank$EnclArea[i] <- area_i

# neighbours drawn from both ID1 and ID2, as it is a sparse matrix
node_i_neighbours <- subset(node_distance_table, (ID1==site_id) | (ID2==site_id))

# tidy up table, to ensure area lookup works
# number of neighbours
num_neighbours <- nrow(node_i_neighbours)
# prevents problems with some exceptions, go to next item in main loop
# rank defaults to zero if so
if (num_neighbours == 0) {next}

for (j in 1:num_neighbours)
{
      if(node_i_neighbours$ID1[j]==site_id)
      {
      # exchange ID1 and ID2, so site id is always in ID2 column
            node_i_neighbours$ID1[j] <- node_i_neighbours$ID2[j]
            node_i_neighbours$ID2[j] <- site_id
      }
}
# rank on distance, nearest first
ranked_i_neighbours <- node_i_neighbours[order(node_i_neighbours$d12),]
# Check if more or less neighbours than the preset limit

if (num_neighbours < n_value)  {
      n_limit <- num_neighbours
} else {
      # preset value for number of neighbours to rank
      n_limit <- n_value
}

# need to check if no neighbours and set rank value, else does not work
if (n_limit != 0){
      # for neighbours to n_limit (unless not that many)
      neighbour_areas <- ranked_i_neighbours[1:n_limit,]
```

```
            # add in the areas for these sites in a new column
            neighbour_areas$EnclArea <- site_area$EnclArea[site_area$AtlasId %in% neighbour_areas$ID1]
            # add in the area of the current site to the array, in order for the ranking to work
            i_row <- c(site_id,site_id,0,area_i)
            neighbour_areas <- rbind(i_row,neighbour_areas)
            # rank the neighbour sites in order of area, ascending
            # top rank will come out as n_limit+1 as ranking is order in the list, normalise
            rank_i <- round(match(area_i, sort(neighbour_areas$EnclArea, decreasing=FALSE))/(n_limit+1),
                  digits=2)
      } else {
            # if no neighbours then rank 0 - (need to consider alternatives ...??)
            rank_i <- 0
      }

      #print(paste("Node ",site_id,", area rank: ",rank_i,", neighbours: ",n_limit, sep=""))
      # Write rank value into ppp object
      site_area_n_rank$area_rank[i] <- rank_i
}

# Save the rank file
file_name <- paste(path_results,"/","site_area_rank_by_",n_value,"_neighbours.csv",sep="")
write.csv(site_area_n_rank,file_name,row.names=FALSE)

# Print sites with colours based on rank ...
rank_threshold_range <- c(0.99)
# only do this for top ranking now ..

for (rank_threshold in rank_threshold_range)
{
      file_map_png <- paste(path_plots,"/",map_name,"_neighbours_", n_value, "_", rank_threshold,
            ".png",sep="")
      png(file=file_map_png, units="cm", width=21, height=29.7, res=300)
      site_area_n_rank$col <- ifelse(site_area_n_rank$area_rank>rank_threshold,"red","grey")
      plot(map_outline)
      # plot low rank sites
      points(site_area_n_rank, pch=19, cex=ifelse(site_area_n_rank$area_rank>rank_threshold,0.0,0.5),
            col=site_area_n_rank$col)
      # plot high rank sites (to avoid being overplotted)
      points(site_area_n_rank, pch=19, cex=ifelse(site_area_n_rank$area_rank>rank_threshold,0.9,0.0),
```

```
                  col=site_area_n_rank$col)
            title(paste("Hillforts ranked on area rel to: ",n_value," neighbours"),
                  sub=paste("Sources: ",source_file_name,"; ",shape_file_name))
            mtext(paste("Rank threshold: ", rank_threshold))
            dev.off()
      }
}


print("next step is to generate relative risk surface for different rank thresholds")
#readline(prompt="Press [enter] to continue")
# -------------------------------------------------------------------------------------------#
# Relative risk surfaces for different rank thresholds
# edit value to be most appropriate for this data set
n_value <- 100
# for Ireland 100
# Note that this requires a sufficient number of neighbours
# as the distance table is sparse, this may need extending the upper range to be effective
# Node distance table recomputed with an upper limit of 100km to aid this

# read in corresponding data file, written out above

file_name <- paste(path_results,"/","site_area_rank_by_",n_value,"_neighbours.csv",sep="")
site_area_n_rank <- read.csv(file_name)
# convert to ppp object
coordinates(site_area_n_rank) <- c("Easting", "Northing")
proj4string(site_area_n_rank) <- crs_projection

# Create cut relative risk surfaces for different area rank thresholds
# Plot for different rank thresholds and probability thresholds
# only do this for top threshold now
rank_threshold_range <- c(0.99)

for (rank_threshold in rank_threshold_range)

{
      print(paste("rank threshold: ", rank_threshold))
      # Create marked ppp data object for small site ranks
      small_sites <- site_area_n_rank[site_area_n_rank$area_rank <= rank_threshold,]
```

```
# constrain by map_bounds, else some sites will be lost on edges
msmall_sites <- as.ppp(coordinates(small_sites), as.owin(map_bounds))
marks(msmall_sites) <- as.factor("small")
# Create marked ppp data object for large site ranks
big_sites <- site_area_n_rank[site_area_n_rank$area_rank > rank_threshold,]
mbig_sites <- as.ppp(coordinates(big_sites), as.owin(map_bounds))
marks(mbig_sites) <- as.factor("big")

# combine into single marked object
all_sites <- superimpose(mbig_sites, msmall_sites)

# Point of neutrality, i.e. ratio of Big sites to all sites, numbers of points
neutral_point <- signif(mbig_sites$n/all_sites$n, digits=2)

file_map_png <- paste(path_plots,"/",map_name,"_rrD_n", n_value, "_%02d.png",sep="")
png(file=file_map_png, units="cm", width=21, height=29.7, res=300)

plot(map_outline)
points(all_sites, main="big sites red, small sites grey", pch=19, cex=ifelse(all_sites$marks=="big",0.9,0.4),
      col=(ifelse(all_sites$marks=="big","red","grey")))
title(paste("Hillforts, area rank, threshold: ",rank_threshold,", neighbours: ",n_value))

# Create relative risk surface
rrd_all_sites <- relrisk(all_sites, sigma= diggle_sigma, edge=TRUE, eps=2500)
crop_rrd_all_s <- rrd_all_sites[e_map, drop=FALSE]

# use density plots for the subset of sites after using owin(map_bounds)
# created earlier

plot(crop_s_diggle, main="", col=topo.colors(100), box=FALSE)
title("Hillfort kernel density used for cropping rrD plots")
mtext(paste("Diggle sigma value: ",round(diggle_sigma, digits=0)))
#plot(crop_s_12500, box=FALSE)
crop_sites <- crop_s_diggle
#crop_sites <- crop_s_12500

# create versions with different cut off probability thresholds from all sites
rrd_all_sites4 <- crop_rrd_all_s
rrd_all_sites4[as.matrix(crop_sites)<(4e-08)] <- NA
```

```
rrd_all_sites3 <- crop_rrd_all_s
rrd_all_sites3[as.matrix(crop_sites)<(3e-08)] <- NA
rrd_all_sites2 <- crop_rrd_all_s
rrd_all_sites2[as.matrix(crop_sites)<(2e-08)] <- NA
rrd_all_sites1 <- crop_rrd_all_s
rrd_all_sites1[as.matrix(crop_sites)<(1e-08)] <- NA
rrd_all_sites075 <- crop_rrd_all_s
rrd_all_sites1[as.matrix(crop_sites)<(0.75e-08)] <- NA

# plot relative risk surface, with "small" and "big" sites, i.e. relative to 'n' neighbours
plot(crop_rrd_all_s, box=FALSE, main="")
points(big_sites, pch=19, cex=0.9, col="red")
title(paste("Hillfort relative risk big vs. small sites, rank threshold: ",rank_threshold,", neighbours: ",
      n_value))
mtext(paste("big sites, red; point of neutrality: ", neutral_point))

# create density plot for the subset of sites after using owin(map_outline)
#plot(crop_s_diggle, box=FALSE, main="")
#title(paste("Hillfort kernel density plot, neighbours: ",n_value))
#mtext(paste("Diggle sigma value: ",round(diggle_sigma, digits=0)))

plot(rrd_all_sites4, main="", box=FALSE)
plot(map_outline, add=TRUE, border="grey")
points(big_sites, pch=19, cex=0.9, col="red")
title(paste("Hillfort relative risk big vs. small, rank threshold: ",rank_threshold,", neighbours: ",
      n_value))
mtext(paste("probability threshold: 4e-08; point of neutrality: ",
      neutral_point))

plot(rrd_all_sites3, main="", box=FALSE)
plot(map_outline, add=TRUE, border="grey")
points(big_sites, pch=19, cex=0.9, col="red")
title(paste("Hillfort relative risk big vs.small, rank threshold: ",rank_threshold,", neighbours: ",
      n_value))
mtext(paste("probability threshold: 3e-08; point of neutrality: ",
      neutral_point))

plot(rrd_all_sites2, main="", box=FALSE)
plot(map_outline, add=TRUE, border="grey")
```

```
	points(big_sites, pch=19, cex=0.9, col="red")
	title(paste("Hillfort relative risk big vs. small, rank threshold: ",rank_threshold,", neighbours: ",
		n_value))
	mtext(paste("probability threshold: 2e-08; point of neutrality: ",
		neutral_point))

	plot(rrd_all_sites1, main="", box=FALSE)
	plot(map_outline, add=TRUE, border="grey")
	points(big_sites, pch=19, cex=0.9, col="red")
	title(paste("Hillfort relative risk big vs. small, rank threshold: ",rank_threshold,", neighbours: ",
		n_value))
	mtext(paste("probability threshold: 1e-08; point of neutrality: ",
		neutral_point))

	plot(rrd_all_sites075, main="", box=FALSE)
	plot(map_outline, add=TRUE, border="grey")
	points(big_sites, pch=19, cex=0.9, col="red")
	title(paste("Hillfort relative risk, rank threshold: ",rank_threshold,", neighbours: ",
		n_value))
	mtext(paste("probability threshold: 0.75e-08; point of neutrality: ",
		neutral_point))
}
dev.off()

print("next step is to generate relative risk surface for Stone Circles")
#readline(prompt="Press [enter] to continue")

# ---------------------------------------------------------------------------------------------#
# Stone Circle data analysis; for both GB and Ireland, run separately for each, different coord systems
# Source files:
# Ortega_coords_GB.160526.csv
# Ortega_coords_Ireland.160526.csv

SC_source_file_name <- "Ortega_coords_GB.160526.csv"
#SC_source_file_name <- "Ortega_Ireland_converted.160630.csv"
source_file <- paste(path_source,"/",SC_source_file_name,sep="")
SC_source_file_name <- substr(SC_source_file_name,1,(nchar(SC_source_file_name)-4))

# stone circle coordinates
```

```
SC_xy_data <- read.csv(source_file, as.is = FALSE)
coordinates(SC_xy_data) <- c("Easting", "Northing")
proj4string(SC_xy_data) <- crs_projection
# SC_sites is the same data, within the bounds of the map outline
SC_sites <- as.ppp(coordinates(SC_xy_data), as.owin(map_bounds))
# note: this avoids excluding sites near boundary
n_sites <- SC_sites$n

# determine 'optimum' standard deviation (sigma) using Diggle
diggle_sigma <- bw.diggle(SC_sites)
# Create kernel density plots for sites

SC_sites_diggle <- density(SC_sites, sigma=diggle_sigma, edge=TRUE, eps=2500)
SC_sites_4000 <- density(SC_sites, sigma=4000, edge=TRUE, eps=2500)
SC_sites_6000 <- density(SC_sites, sigma=6000, edge=TRUE, eps=2500)
#SC_sites_10000 <- density(SC_sites, sigma=10000, edge=TRUE, eps=2500)
#SC_sites_12500 <- density(SC_sites, sigma=12500, edge=TRUE, eps=2500)
# crop to map outline
crop_SC_s_diggle <- SC_sites_diggle[e_map, drop=FALSE]
crop_SC_s_4000 <- SC_sites_4000[e_map, drop=FALSE]
crop_SC_s_6000 <- SC_sites_6000[e_map, drop=FALSE]
crop_SC_s_10000 <- SC_sites_10000[e_map, drop=FALSE]
#crop_SC_s_12500 <- SC_sites_12500[e_map, drop=FALSE]

map_name <- SC_source_file_name

file_map_png <- paste(path_plots,"/",map_name,"_kD_%02d.png",sep="")
png(file=file_map_png, units="cm", width=21, height=21, res=300)

plot(map_outline)
points(SC_sites,pch=19,cex=0.3,col="black")
title(paste("Stone Circles; ", n_sites),
      sub=paste("Source: ",SC_source_file_name))

plot(crop_SC_s_diggle, main="", col=topo.colors(100), box=FALSE)
title(paste("Stone Circles kernel density; Sigma value: ",round(diggle_sigma, digits=0)),
      sub=paste("Sources: ",SC_source_file_name))
plot(map_outline, main="", add=TRUE)
```

```
plot(crop_SC_s_4000, main="", col=topo.colors(100), box=FALSE)
title("Stone Circles kernel density; Sigma value: 4000",
      sub=paste("Sources: ",SC_source_file_name))
plot(map_outline, main="", add=TRUE)

plot(crop_SC_s_6000, main="", col=topo.colors(100), box=FALSE)
title("Stone Circles kernel density; Sigma value: 6000",
      sub=paste("Sources: ",SC_source_file_name))
plot(map_outline, main="", add=TRUE)

#plot(crop_SC_s_10000, main="", col=topo.colors(100), box=FALSE)
#title("Stone Circles kernel density; Sigma value: 10000",
#     sub=paste("Sources: ",SC_source_file_name))
#plot(map_outline, main="", add=TRUE)

SC_dist <- nndist(sites)

hist(SC_dist, xlab="nearest neighbour distance",
      main="Stone Circles: nearest neighbour histogram")
title("", sub=paste("Sources: ",SC_source_file_name,"; ",shape_file_name))
hist(SC_dist, breaks=240, xlim=c(0, 10000), prob=TRUE, xlab="nearest neighbour distance",
      main="Stone Circles: nearest neighbour histogram")
title("", sub=paste("Sources: ",SC_source_file_name,"; ",shape_file_name))
hist(SC_dist, breaks=240, prob=TRUE, xlab="nearest neighbour distance",
      main="Stone Circles: nearest neighbour histogram")
title("", sub=paste("Sources: ",SC_source_file_name,"; ",shape_file_name))
dev.off()

# Create marked ppp object for Stone Circles
mSC_sites <- (SC_sites)
marks(mSC_sites) <- as.factor("SC")
# Create marked ppp data object for Hillforts
msites <- sites
marks(msites) <- as.factor("Hillforts")

# combine into single marked object
comb_sites <- superimpose(mSC_sites, msites)

# Calculate point of neutrality
```

```
neutral_point <- signif(sites$n/comb_sites$n, digits=2)
# Create density surface for combined sites
comb_sites_10000 <- density(comb_sites, sigma=10000, edge=TRUE, eps=2500)
crop_comb_s_10000 <- comb_sites_10000[e_map, drop=FALSE]

# Create relative risk surface
rrd_comb_sites <- relrisk(comb_sites, sigma= 10000, edge=TRUE, eps=2500)
crop_rrd_comb_s <- rrd_comb_sites[e_map, drop=FALSE]

map_name <- paste(SC_source_file_name, "_density_plots")

file_map_png <- paste(path_plots,"/",map_name,"_%02d.png",sep="")
png(file=file_map_png, units="cm", width=21, height=29.7, res=300)

#create breaks for colour palette, centred on the point of neutrality
lbrk <- seq(0, neutral_point, by=(neutral_point/50))
ubrk <- seq(neutral_point, 1, by=((1-neutral_point)/49))
brk <- c(lbrk, ubrk)

plot(map_outline, main="")
points(comb_sites, pch=19, cex=0.5,col=(ifelse(comb_sites$marks=="Hillforts","red","black")))
title("Stone Circles and Hillforts",
      sub=paste("Sources: ",SC_source_file_name,"; ",source_file_name))
mtext("Stone Circles: black, Hillforts: red")

plot(crop_rrd_comb_s, main="", col=topo.colors(100), breaks=brk, box=FALSE)
title("Relative Risk combined Stone Circles and Hillforts",
      sub=paste("Sources: ",SC_source_file_name,"; ",source_file_name))
mtext(paste("Point of neutrality: ", neutral_point))
plot(map_outline, add=TRUE)
points(sites,pch=19,cex=0.5,col="red")
points(SC_sites,pch=19,cex=0.5,col="black")

rrd_comb_sites1 <- crop_rrd_comb_s
rrd_comb_sites1[as.matrix(crop_comb_s_10000)<1e-08] <- NA
plot(rrd_comb_sites1, main="", col=topo.colors(100), breaks=brk, box=FALSE)
title("Relative Risk combined Stone Circles and Hillforts: 1e-08",
      sub=paste("Sources: ",SC_source_file_name,"; ",source_file_name))
mtext(paste("Point of neutrality: ", neutral_point))
```

```
plot(map_outline, add=TRUE)
points(sites,pch=19,cex=0.5,col="red")

rrd_comb_sites15 <- crop_rrd_comb_s
rrd_comb_sites15[as.matrix(crop_comb_s_10000)<(1.5e-08)] <- NA
plot(rrd_comb_sites15, main="", col=topo.colors(100), breaks=brk, box=FALSE)
title("Relative Risk combined Stone Circles and Hillforts: 1.5e-08",
      sub=paste("Sources: ",SC_source_file_name,"; ",source_file_name))
mtext(paste("Point of neutrality: ", neutral_point))
plot(map_outline, add=TRUE)
points(sites,pch=19,cex=0.5,col="red")

rrd_comb_sites2 <- crop_rrd_comb_s
rrd_comb_sites2[as.matrix(crop_comb_s_10000)<(2e-08)] <- NA
plot(rrd_comb_sites2, main="", col=topo.colors(100), breaks=brk, box=FALSE)
title("Relative Risk combined Stone Circles and Hillforts: 2e-08",
      sub=paste("Sources: ",SC_source_file_name,"; ",source_file_name))
mtext(paste("Point of neutrality: ", neutral_point))
plot(map_outline, add=TRUE)
points(sites,pch=19,cex=0.5,col="red")

rrd_comb_sites3 <- crop_rrd_comb_s
rrd_comb_sites3[as.matrix(crop_comb_s_10000)<(3.0e-08)] <- NA
plot(rrd_comb_sites3, main="", col=topo.colors(100), breaks=brk, box=FALSE)
title("Relative Risk combined Stone Circles and Hillforts: 3e-08",
      sub=paste("Sources: ",SC_source_file_name,"; ",source_file_name))
mtext(paste("Point of neutrality: ", neutral_point))
plot(map_outline, add=TRUE)
points(sites,pch=19,cex=0.5,col="red")

dev.off()
readline(prompt="Press [enter] to continue")

# ---------------------------------------------------------------------------------------------#
# Irish Ring Fort data analysis; for Ireland only
# Data provided in the form of a shape file for all Irish SMR by Marta and Alex
# Extraction code provided by them also

# This is the SMR record for all sites of various classifications
```

```
RF_source_file_name <- "smrImerged.shp"
RF_source_file <- paste(path_shape, "/", RF_source_file_name, sep="")

# This is the same map outline used for Ireland in other analyses
#shape_file_name <- paste(path_shape, "/Adm_Merged.shp", sep="")
#map_name <- "Ireland"

# Read in shape file with data
mergedSMRI <- readOGR(RF_source_file, layer="smrImerged")

# not necessary with my way of using map bounds
#ireland <- readOGR(map_file, layer="Adm_Merged")
#myrows <- row.names(mergedSMRI[is.na(over(mergedSMRI, ireland))[,1],])
#SMRInew <- mergedSMRI[!row.names(mergedSMRI)%in% myrows,]
SMRInew <- mergedSMRI
nrow(SMRInew)

# Check for pairs of points that overlap entirely
zd <- zerodist(SMRInew)
# Have a look at how this reports results
head(zd)
# remove duplicate locations
SMRInewBU <- SMRInew
SMRInew <- SMRInew[-zd[,2], ]
# counts of before and after removal of duplicates
nrow(SMRInewBU)
nrow(SMRInew)

# pull out Ringfort sites
RF_sites <- subset(SMRInew, (SMRInew$SiteType =="Ringfort"))
# NOTE for this study other types of site not of interest

# number of ringforts
n_RF_sites <- nrow(RF_sites)
# convert to ppp object
RF_sites <- as.ppp(coordinates(RF_sites), as.owin(map_bounds))

plot(map_outline)
points(RF_sites, pch=19, cex=0.01)
```

```
#diggle_sigma <- bw.diggle(RF_sites)
# Create kernel density plots for sites

#RF_sites_diggle <- density(RF_sites, sigma=diggle_sigma, edge=TRUE, eps=2500)
RF_sites_2000 <- density(RF_sites, sigma=2000, edge=TRUE, eps=2500)
RF_sites_3000 <- density(RF_sites, sigma=3000, edge=TRUE, eps=2500)
RF_sites_4000 <- density(RF_sites, sigma=4000, edge=TRUE, eps=2500)
#RF_sites_6000 <- density(RF_sites, sigma=6000, edge=TRUE, eps=2500)
#RF_sites_10000 <- density(RF_sites, sigma=10000, edge=TRUE, eps=2500)
#RF_sites_12500 <- density(RF_sites, sigma=12500, edge=TRUE, eps=2500)
# crop to map outline
#crop_RF_s_diggle <- RF_sites_diggle[e_map, drop=FALSE]
crop_RF_s_2000 <- RF_sites_2000[e_map, drop=FALSE]
crop_RF_s_3000 <- RF_sites_3000[e_map, drop=FALSE]
crop_RF_s_4000 <- RF_sites_4000[e_map, drop=FALSE]
#crop_RF_s_6000 <- RF_sites_6000[e_map, drop=FALSE]
#crop_RF_s_10000 <- RF_sites_10000[e_map, drop=FALSE]
#crop_RF_s_12500 <- RF_sites_12500[e_map, drop=FALSE]
map_name <- RF_source_file_name
file_map_png <- paste(path_plots,"/",map_name,"_kD_%02d.png",sep="")
png(file=file_map_png, units="cm", width=21, height=21, res=300)

plot(map_outline)
points(RF_sites,pch=19,cex=0.01,col="black")
title(paste("Irish Ringforts; ", n_RF_sites),
      sub=paste("Source: ",RF_source_file_name))

plot(map_outline)
points(RF_sites,pch=19,cex=0.01,col="black")
points(sites,pch=19,cex=0.5, col="red")
title("Ireland, Ringforts and Hillforts",
      sub=paste("Sources: ",RF_source_file_name, "; ", source_file_name))
mtext("Irish Ringforts: gray, Hillforts: red")

#plot(crop_RF_s_diggle, main="", col=topo.colors(100), box=FALSE)
#title(paste("Ringforts kernel density; Sigma value: ",round(diggle_sigma, digits=0)),
#     sub=paste("Sources: ",RF_source_file_name))
#plot(map_outline, main="", add=TRUE)
```

```
plot(crop_RF_s_2000, main="", col=topo.colors(100), box=FALSE)
title("Ringforts kernel density; Sigma value: 2000",
      sub=paste("Sources: ",RF_source_file_name))
plot(map_outline, main="", add=TRUE)

plot(crop_RF_s_3000, main="", col=topo.colors(100), box=FALSE)
title("Ringforts kernel density; Sigma value: 3000",
      sub=paste("Sources: ",RF_source_file_name))
plot(map_outline, main="", add=TRUE)

plot(crop_RF_s_4000, main="", col=topo.colors(100), box=FALSE)
title("Ringforts kernel density; Sigma value: 4000",
      sub=paste("Sources: ",RF_source_file_name))
plot(map_outline, main="", add=TRUE)

#plot(crop_RF_s_6000, main="", col=topo.colors(100), box=FALSE)
#title("Ringforts kernel density; Sigma value: 6000",
#     sub=paste("Sources: ",RF_source_file_name))
#plot(map_outline, main="", add=TRUE)

RF_dist <- nndist(RF_sites)

hist(RF_dist, xlab="nearest neighbour distance",
      main="Ringforts: nearest neighbour histogram")
title("", sub=paste("Sources: ",RF_source_file_name,"; ",shape_file_name))
hist(RF_dist, breaks=240, xlim=c(0, 10000), prob=TRUE, xlab="nearest neighbour distance",
      main="Ringforts: nearest neighbour histogram")
title("", sub=paste("Sources: ",RF_source_file_name,"; ",shape_file_name))
hist(RF_dist, breaks=240, prob=TRUE, xlab="nearest neighbour distance",
      main="Ringforts: nearest neighbour histogram")
title("", sub=paste("Sources: ",RF_source_file_name,"; ",shape_file_name))
dev.off()

# Create marked ppp object for Ringforts
mRF_sites <- RF_sites
marks(mRF_sites) <- as.factor("Ringfort")
# Create marked ppp data object for Hillforts
msites <- sites
```

```
marks(msites) <- as.factor("Hillforts")

# combine into single marked object
comb_sites <- superimpose(mRF_sites, msites)

# Calculate point of neutrality
neutral_point <- signif(sites$n/comb_sites$n, digits=2)
# Create density surface for combined sites
comb_sites_10000 <- density(comb_sites, sigma=10000, edge=TRUE, eps=2500)
crop_comb_s_10000 <- comb_sites_10000[e_map, drop=FALSE]

# Create relative risk surface
rrd_comb_sites <- relrisk(comb_sites, sigma= 10000, edge=TRUE, eps=2500)
crop_rrd_comb_s <- rrd_comb_sites[e_map, drop=FALSE]

map_name <- paste(RF_source_file_name, "_density_plots")

file_map_png <- paste(path_plots,"/",map_name,"_%02d.png",sep="")
png(file=file_map_png, units="cm", width=21, height=29.7, res=300)

#create breaks for colour palette, centred on the point of neutrality
lbrk <- seq(0, neutral_point, by=(neutral_point/50))
ubrk <- seq(neutral_point, 1, by=((1-neutral_point)/49))
brk <- c(lbrk, ubrk)

plot(map_outline, main="")
points(comb_sites, pch=19,
cex=(ifelse(comb_sites$marks=="Hillforts",0.5,0.01)),col=(ifelse(comb_sites$marks=="Hillforts","red","black")))
title("Ringforts and Hillforts",
      sub=paste("Sources: ",RF_source_file_name,"; ",source_file_name))
mtext("Ringforts: black, Hillforts: red")

plot(crop_rrd_comb_s, main="", col=topo.colors(100), breaks=brk, box=FALSE)
plot(map_outline, main="", add=TRUE)
points(sites,pch=19,cex=0.5, col="red")
title("Relative Risk combined Ringforts and Hillforts",
      sub=paste("Sources: ",RF_source_file_name,"; ",source_file_name))
mtext(paste("Point of neutrality: ", neutral_point))
```

```
rrd_comb_sites1 <- crop_rrd_comb_s
rrd_comb_sites1[as.matrix(crop_comb_s_10000)<1e-07] <- NA
plot(rrd_comb_sites1, main="", col=topo.colors(100), breaks=brk, box=FALSE)
title("Relative Risk combined Ringforts and Hillforts: 1e-07",
      sub=paste("Sources: ",SC_source_file_name,"; ",source_file_name))
mtext(paste("Point of neutrality: ", neutral_point))
plot(map_outline, main="", add=TRUE)
points(sites,pch=19,cex=0.5, col="red")

rrd_comb_sites15 <- crop_rrd_comb_s
rrd_comb_sites15[as.matrix(crop_comb_s_10000)<(0.5e-07)] <- NA
plot(rrd_comb_sites15, main="", col=topo.colors(100), breaks=brk, box=FALSE)
title("Relative Risk combined Ringforts and Hillforts: 0.5e-07",
      sub=paste("Sources: ",RF_source_file_name,"; ",source_file_name))
mtext(paste("Point of neutrality: ", neutral_point))
plot(map_outline, main="", add=TRUE)
points(sites,pch=19,cex=0.5, col="red")

rrd_comb_sites2 <- crop_rrd_comb_s
rrd_comb_sites2[as.matrix(crop_comb_s_10000)<(2e-07)] <- NA
plot(rrd_comb_sites2, main="", col=topo.colors(100), breaks=brk, box=FALSE)
title("Relative Risk combined Ringforts and Hillforts: 2e-07",
      sub=paste("Sources: ",RF_source_file_name,"; ",source_file_name))
mtext(paste("Point of neutrality: ", neutral_point))
plot(map_outline, main="", add=TRUE)
points(sites,pch=19,cex=0.5, col="red")

rrd_comb_sites3 <- crop_rrd_comb_s
rrd_comb_sites3[as.matrix(crop_comb_s_10000)<(3.0e-07)] <- NA
plot(rrd_comb_sites3, main="", col=topo.colors(100), breaks=brk, box=FALSE)
title("Relative Risk combined Ringforts and Hillforts: 3e-07",
      sub=paste("Sources: ",RF_source_file_name,"; ",source_file_name))
mtext(paste("Point of neutrality: ", neutral_point))
plot(map_outline, main="", add=TRUE)
points(sites,pch=19,cex=0.5, col="red")

dev.off()

readline(prompt="Press [enter] to continue")
```

```
# ------------------------------------------------------------------------------------------#
# PAS data analysis - Iron Age filter applied: only for England (although a few sites in Wales)
# Data file: PAS_IA_xy.160519.csv
PAS_source_file_name <- "PAS_IA_xy.160519.csv"
source_file <- paste(path_source,"/",PAS_source_file_name,sep="")
# Truncate source file name for use in output file names
PAS_source_file_name <- substr(PAS_source_file_name,1,(nchar(PAS_source_file_name)-4))

PAS_xy_data <- read.csv(source_file, as.is = FALSE)

# Load shape file for England and English Counties
shape_file_name <- "England_ct_1991_area.shp"
file_name <- paste(path_shape,"/",shape_file_name,sep="")
# Truncate shape file name to remove extension
layer_name <- substr(shape_file_name,1,(nchar(shape_file_name)-4))
map_outline <- readOGR(dsn=file_name,layer=layer_name)
# create map bounds of the map outline
map_bounds <- boundingbox(as.owin(map_outline))
plot(map_bounds)
plot(map_outline, add=TRUE)

# get projection string for current map
projection_string <- proj4string(map_outline)
crs_projection <- CRS(projection_string)

coordinates(PAS_xy_data) <- c("Easting", "Northing")
proj4string(PAS_xy_data) <- crs_projection

# read in hillfort data, limit to England
file_name <- paste(path_source,"/Confirmed_sites.Key_params.160610.csv", sep="")
xy_data_all <- fread(file_name, header=TRUE, select=c("AtlasId","X","Y","Country"))
xy_data_england <- subset(xy_data_all,Country=="EN")
xy_data <- subset(xy_data_england,select=c("AtlasId","X","Y"))

# Convert data to SpatialPointsDataFrame, using project4string from selected map file

# turn data into SpatialPointsDataFrame
coordinates(xy_data) <- c("X", "Y")
```

```
proj4string(xy_data) <- crs_projection
sites <- as.ppp(coordinates(xy_data), as.owin(map_bounds))

# PAS_sites is the same data, within the bounds of the map bounds
PAS_sites <- as.ppp(coordinates(PAS_xy_data), as.owin(map_bounds))
n_sites <- PAS_sites$n
# note: otherwise possibly removes some sites which are located close to the map outline
# NOTE: sites are still rejected but they are probably sites with no coordinates
#     for confidentiality/ security

# determine 'optimum' standard deviation (sigma) using Diggle
diggle_sigma <- bw.diggle(PAS_sites)
# Create kernel density plots for sites
PAS_sites_diggle <- density(PAS_sites, sigma=diggle_sigma, edge=TRUE, eps=2500)
PAS_sites_10000 <- density(PAS_sites, sigma=10000, edge=TRUE, eps=2500)
#PAS_sites_12500 <- density(PAS_sites, sigma=12500, edge=TRUE, eps=2500)
# crop to map outline
crop_PAS_s_diggle <- PAS_sites_diggle[e_map, drop=FALSE]
crop_PAS_s_10000 <- PAS_sites_10000[e_map, drop=FALSE]
#crop_PAS_s_12500 <- PAS_sites_12500[e_map, drop=FALSE]

map_name <- PAS_source_file_name
# plot Kernel Density for IA PAS finds
file_map_png <- paste(path_plots,"/",map_name,"_kD%02d.png",sep="")
png(file=file_map_png, units="cm", width=21, height=21, res=300)

plot(map_outline)
points(PAS_sites, main="", pch=19,cex=0.1,col="black")
title(paste("PAS Iron Age finds: ", n_sites),
      sub=paste("Sources: ",PAS_source_file_name))
mtext(paste("Shape file: ", shape_file_name))
plot(crop_PAS_s_diggle, col=topo.colors(100), main="", box=FALSE)
plot(map_outline, add=TRUE)
title(paste("PAS IA finds kernel density; Sigma value: ",round(diggle_sigma, digits=0)),
      sub=paste("Sources: ",PAS_source_file_name))
mtext(paste("Shape file: ", shape_file_name))
plot(crop_PAS_s_10000, col=topo.colors(100), main="", box=FALSE)
plot(map_outline, add=TRUE)
title("PAS IA finds kernel density; Sigma value: 10000",
```

```
      sub=paste("Sources: ",PAS_source_file_name))
mtext(paste("Shape file: ", shape_file_name))

#plot(crop_PAS_s_12500, col=topo.colors(100), main="", box=FALSE)
#title("PAS IA finds kernel density; Sigma value: 12500",
#     sub=paste("Sources: ",PAS_source_file_name))
#mtext(shape_file_name)
#plot(crop_PAS_s_15000, col=topo.colors(100), main="", box=FALSE)
#title("PAS IA finds kernel density; Sigma value: 15000",
#     sub=paste("Sources: ",PAS_source_file_name))
#mtext(shape_file_name)
dev.off()

# Nearest neighbour analysis
PAS_sites_dist <- nndist(sites)
file_map_png <- paste(path_plots,"/",map_name,"_nearest_neighbour%02d.png",sep="")
png(file=file_map_png, units="cm", width=21, height=21, res=300)
hist(PAS_sites_dist, xlab="nearest neighbour distance",
      main="PAS IA finds: nearest neighbour histogram")
title("", sub=paste("Sources: ",PAS_source_file_name))
hist(PAS_sites_dist, breaks=240, xlim=c(0, 10000), prob=TRUE, xlab="nearest neighbour distance",
      main="PAS IA finds: nearest neighbour histogram")
title("", sub=paste("Sources: ",PAS_source_file_name))
hist(PAS_sites_dist, breaks=240, prob=TRUE, xlab="nearest neighbour distance",
      main="PAS IA finds: nearest neighbour histogram")
title("", sub=paste("Sources: ",PAS_source_file_name))
dev.off()

# Create marked ppp object for PAS sites
mPAS_sites <- (PAS_sites)
marks(mPAS_sites) <- as.factor("PAS")
# Create marked ppp data object for Hillforts
msites <- sites
marks(msites) <- as.factor("Hillforts")

# combine into single marked object
comb_sites <- superimpose(mPAS_sites, msites)

# Calculate point of neutrality
```

```
neutral_point <- signif(sites$n/comb_sites$n, digits=2)
# Create density surface for combined sites
comb_sites_10000 <- density(comb_sites, sigma=10000, edge=TRUE, eps=2500)
crop_comb_s_10000 <- comb_sites_10000[e_map, drop=FALSE]

# Create relative risk surface
rrd_comb_sites <- relrisk(comb_sites, sigma= 10000, edge=TRUE, eps=2500)
crop_rrd_comb_s <- rrd_comb_sites[e_map, drop=FALSE]

map_name <- paste(PAS_source_file_name, "_density_plots")

file_map_png <- paste(path_plots,"/",map_name,"_%02d.png",sep="")
png(file=file_map_png, units="cm", width=21, height=21, res=300)

#create breaks for colour palette, centred on the point of neutrality
lbrk <- seq(0, neutral_point, by=(neutral_point/50))
ubrk <- seq(neutral_point, 1, by=((1-neutral_point)/49))
brk <- c(lbrk, ubrk)

plot(map_outline, main="")
points(comb_sites, pch=19,
      cex=(ifelse(comb_sites$marks=="Hillforts",0.4,0.2)),
      col=(ifelse(comb_sites$marks=="Hillforts","red","gray40")))
title("PAS Iron Age and Hillforts",
      sub=paste("Sources: ",PAS_source_file_name,"; ",source_file_name))
mtext("PAS Iron Age finds: gray, Hillforts: red")

plot(crop_rrd_comb_s, main="", col=topo.colors(100), breaks=brk, box=FALSE)
title("Relative Risk combined PAS and Hillforts",
      sub=paste("Sources: ",PAS_source_file_name,"; ",source_file_name))
mtext(paste("Point of neutrality: ", neutral_point))
plot(map_outline, add=TRUE)
#points(sites, pch=19, cex=0.3, col="red")

rrd_comb_sites01 <- crop_rrd_comb_s
rrd_comb_sites01[as.matrix(crop_comb_s_10000)<.1e-06] <- NA
plot(rrd_comb_sites01, main="", col=topo.colors(100), breaks=brk, box=FALSE)
title("Relative Risk combined PAS and Hillforts: 0.1e-06",
      sub=paste("Sources: ",PAS_source_file_name,"; ",source_file_name))
```

```
mtext(paste("Point of neutrality: ", neutral_point))
plot(map_outline, add=TRUE)

rrd_comb_sites015 <- crop_rrd_comb_s
rrd_comb_sites015[as.matrix(crop_comb_s_10000)<(.15e-06)] <- NA
plot(rrd_comb_sites015, main="", col=topo.colors(100), breaks=brk, box=FALSE)
title("Relative Risk combined PAS and Hillforts: 0.15e-06",
      sub=paste("Sources: ",PAS_source_file_name,"; ",source_file_name))
mtext(paste("Point of neutrality: ", neutral_point))
plot(map_outline, add=TRUE)

rrd_comb_sites02 <- crop_rrd_comb_s
rrd_comb_sites02[as.matrix(crop_comb_s_10000)<(.2e-06)] <- NA
plot(rrd_comb_sites02, main="", col=topo.colors(100), breaks=brk, box=FALSE)
title("Relative Risk combined PAS and Hillforts: 0.2e-06",
      sub=paste("Sources: ",PAS_source_file_name,"; ",source_file_name))
mtext(paste("Point of neutrality: ", neutral_point))
plot(map_outline, add=TRUE)

rrd_comb_sites05 <- crop_rrd_comb_s
rrd_comb_sites05[as.matrix(crop_comb_s_10000)<(.5e-06)] <- NA
plot(rrd_comb_sites05, main="", col=topo.colors(100), breaks=brk, box=FALSE)
title("Relative Risk combined PAS and Hillforts: 0.5e-06",
      sub=paste("Sources: ",PAS_source_file_name,"; ",source_file_name))
mtext(paste("Point of neutrality: ", neutral_point))
plot(map_outline, add=TRUE)

rrd_comb_sites07 <- crop_rrd_comb_s
rrd_comb_sites07[as.matrix(crop_comb_s_10000)<(.7e-06)] <- NA
plot(rrd_comb_sites07, main="", col=topo.colors(100), breaks=brk, box=FALSE)
title("Relative Risk combined PAS and Hillforts: 0.7e-06",
      sub=paste("Sources: ",PAS_source_file_name,"; ",source_file_name))
mtext(paste("Point of neutrality: ", neutral_point))
plot(map_outline, add=TRUE)

dev.off()
```

Raster elevation analysis

This script was written to generate data about elevations in the vicinity of each hillfort site. Although it worked for small data sets, such as the Isle of Man, it was far too inefficient for the very large British terrain model, and the work was not continued for lack of time. A different approach would probably be required to solve this.

```
# RASTER ANALYSIS SCRIPT
# Version 1: 10/5/16
# This code analyses elevations from rasters, given point data
# Raster is the DTM for a specified map, points are the coordinate locations of hill forts, in xy
# Details for selected source data (raster, shape file, point data) is specified in source_file.txt
# Version 1.1: 12/7/16

library(raster)
library(calibrate)
library(maptools)
library(rgdal)

setwd("/Users/Simon/Documents/Iron_Age_Hillforts/Percolation")
# paths for results and output maps
path_source <- paste(getwd(),"/source_data",sep="")
path_results <- paste(getwd(),"/working_data",sep="")
path_maps <- paste(getwd(),"/maps",sep="")
path_shape <- paste(getwd(),"/shape_files",sep="")
path_DTM <- paste(getwd(),"/DTM",sep="")

# File is the list of nodes and x y grid references
# Read in source file name defined in file source_file.txt
# Edit this file for different input file name
file_name <- paste(path_results,"/","source_file.txt",sep="")
source_files <- read.csv(file_name,header=TRUE,stringsAsFactors=FALSE)
source_file_name <- source_files$source_file
source_file <- paste(path_source,"/",source_file_name,sep="")
map_name <- source_files$map_name
shape_file_name <- source_files$shape_file
DTM_file_name <- source_files$DTM_file
```

```
# Load shape file
file_name <- paste(path_shape,"/",shape_file_name,sep="")
# Truncate shape file name to remove extension
layer_name <- substr(shape_file_name,1,(nchar(shape_file_name)-4))
print(map_name)
map_outline <- readOGR(dsn=file_name,layer=layer_name)

# Load raster file
file_name <- paste(path_DTM,"/",DTM_file_name,sep="")
map_raster <- raster(file_name)
# Note that this assumes raster set up with projections matching outline

# Read in nodes and grid coordinates
xy_data <- read.csv(source_file, as.is = FALSE)
# fields of interest:
# PlcIndex Easting Northing

# Convert data to SpatialPointsDataFrame, using project4string from selected map file
# get projection string for current map
projection_string <- proj4string(map_outline)
crs_projection <- CRS(projection_string)
# turn data into SpatialPointsDataFrame
coordinates(xy_data) <- c("Easting", "Northing")
proj4string(xy_data) <- crs_projection

# Extract elevations for sites from raster, and export
Plc_elevation <- extract(map_raster,xy_data,method='simple')
xy_data$elevation <- round(Plc_elevation, digits=1)

file_name <- paste(path_results,"/","elevations_",map_name,".csv",sep="")
write.csv(as.data.frame(xy_data),file=file_name,row.names=FALSE,quote=FALSE)

# Compute elevation statistics in buffer around sites, in metres
# Experient with these values
#buffer_size <- 500
for (buffer_size in seq(2000,3000, by=500))
{
	mean_elevs <- extract(map_raster,xy_data,buffer=buffer_size,fun=mean,na.rm=TRUE)
	col_name <- paste("mean_elev_",buffer_size,sep="")
```

```
	xy_data$col_name <- mean_elevs
	names(xy_data)[ncol(xy_data)] <- col_name
	max_elevs <- extract(map_raster,xy_data,buffer=buffer_size,fun=max,na.rm=TRUE)
	col_name <- paste("max_elev_",buffer_size,sep="")
	xy_data$col_name <- max_elevs
	names(xy_data)[ncol(xy_data)] <- col_name
}

file_name <- paste(path_results,"/","elevations_",map_name,".csv",sep="")
write.csv(as.data.frame(xy_data),file=file_name,row.names=FALSE,quote=FALSE)
```

References

Airborne Remote Sensing, https://historicengland.org.uk/research/approaches/research-methods/airborne-remote-sensing/, Accessed: August 2016

Arcaute, E., Ferguson, P., Brookes, S. & Reynolds, A. (2014) Natural regional divisions of places in Domesday Book. *The Connected Past.* Imperial College London, unpublished.

Arcaute, E., Molinero, C., Hatna, E., Murcio, R., Vargas-Ruiz, C., et al. (2016) Cities and regions in Britain through hierarchical percolation. *Royal Society Open Science,* 3.

ArchAtlas, Version 4.1, http://www.archatlas.org/Home.php, Accessed: 2016

An Atlas of Hillforts in Britain and Ireland, http://www.arch.ox.ac.uk/hillforts-atlas.html, Accessed: 12/1/2014

Baddeley, A. & Turner, R. (2005) spatstat: An R Package for Analyzing Spatial Point Patterns. *Journal of Statistical Software,* 12.

Baddeley, A., Turner, R. & Rubak, E. *Package 'spatstat': Spatial Point Pattern Analysis, Model-Fitting, Simulation, Tests*, https://cran.r-project.org/web/packages/spatstat/spatstat.pdf, Accessed: 2016

Batardy, C., Buchsenschutz, O., Gruel, K. & Levéry, M. (2008) Analyse spatiale et statistique de l'âge du Fer en France. L'exemple de la "BaseFer". *Revue archéologique du Centre de la France [En ligne],* 47: 1-28.

Bradford, J. (1940) The Excavation of Cherbury Camp, 1939. *Oxoniensia,* V: 13-20.

Bradley, R. (2007) *The Prehistory of Britain and Ireland,* Cambridge, Cambridge University Press.

Brookes, S. & Reynolds, A. (2011) The Origins of Political Order and the Anglo-Saxon State. *Archaeology International,* 13/14: 84-93.

Brown, I. (2002) The Iron Age Hillforts of Wales and the Marches, their Location and Social Significance. unpublished MLitt thesis, University of Oxford.

Brown, I. (2008) *'Beacons' in the landscape : the hillforts of England and Wales,* Bollington, Windgather.

Burl, A. (2000) *The Stone Circles of Britain, Ireland and Brittany*, Yale University Press.

Clarke, D. (1968) *Analytical Archaeology,* London, Methuen.

Collis, J. (1981) A theoretical study of Hilforts In: Guilbert, G. (Ed.) *Hill-fort studies: Essays for AHA Hogg.* p66-76.

Consortium for Spatial Information, http://srtm.csi.cgiar.org/SELECTION/inputCoord.asp, Accessed: 2016

Cunliffe, B. (1971) Some aspects of hill-forts and their cultural environments. In: Hill, D. & Jesson, M. (Eds.) *The Iron Age and its hill-forts : papers presented to Sir Mortimer Wheeler on the occasion of his eightieth year at a conference held by the Southampton University Archaeological Society, 5th-7th March 1971.* Southampton, Southampton University Archaeological Society: 53-70.

Cunliffe, B. (1982) Britain, the Veneti and beyond. *Oxford Journal of Archaeology,* 1: 39-68.

Cunliffe, B. (Ed.) (1991) *Iron Age Communities in Britain,* London and New York, Routledge.

Cunliffe, B. (2013) *Britain Begins*, Oxford University Press.

Cunliffe, B. & Poole, C. (1991) *Danebury : an Iron Age hillfort in Hampshire Vol.4, The excavations, 1979-1988: the site*, Council for British Archaeology.

Cunliffe, B. & Rowley, T. (Eds.) (1976) *Oppida : the beginnings of urbanisation in Barbarian Europe : papers presented to a conference at Oxford, October 1975,* Oxford, British Archaeological Reports

Cunliffe, B.W. (1993) *English heritage book of Danebury,* London, Batsford/English Heritage.

DBSCAN, https://en.wikipedia.org/wiki/DBSCAN, Accessed: August 2016

Diggle, P.J. (1985) A kernel method for smoothing point process data. *Applied Statistics (Journal of the Royal Statistical Society, Series C),* 34: 138–147.

Digimap Resource Centre, http://digimap.edina.ac.uk/webhelp/resources/, Accessed: 2016

The Domesday Book Online, http://www.domesdaybook.co.uk, Accessed: 7th August 2016

Drennan, R.D. & Peterson, C.E. (2004) Comparing archaeological settlement systems with rank-size graphs: a measure of shape and statistical confidence. *Journal of Archaeological Science,* 31: 533-549.

Duchy of Cornwall, Frequently Asked Questions, http://duchyofcornwall.org/frequently-asked-questions.html#question_20, Accessed: 11/8/2016

Feachem, R.W. (1966) Hill-Forts of Northern Britain. In: Rivet, A.L.F. (Ed.) *The Iron Age in Northern Britain.* Edinbugh, Edinburgh University Press: 59-87.

Fox, A. (1958) South-Western Hill-Forts. In: Frere, S. (Ed.) *Problems of the Iron Age in Southern Britain : papers given at a C.B.A. conference held at the Institute of Archaeology, December 12 to 14, 1958,.* University of London, Institute of Archaeology: 35-60.

Fox, C.S. (1952) *The Personality of Britain: its influence on inhabitant and invader in prehistoric and early historic times,* Cardiff, National Museum of Wales.

fpc, https://cran.r-project.org/web/packages/fpc/fpc.pdf, Accessed: August 2016

Frere, S. (Ed.) (1958a) *Problems of the Iron Age in Southern Britain : papers given at a C.B.A. conference held at the Institute of Archaeology, December 12 to 14, 1958*, University of London, Institute of Archaeology.

Frere, S. (1958b) Some problems of the later Iron Age. In: Frere, S. (Ed.) *Problems of the Iron Age in Southern Britain : papers given at a C.B.A. conference held at the Institute of Archaeology, December 12 to 14, 1958,.* University of London, Institute of Archaeology: 84-92.

Gent, H. & Dean, C. (1986) Catchment Analysis and Settlement Hierarchy: a Case Study from Pre-Roman Britain. In: Grant, E. (Ed.) *Central Places, Archaeology and History.* University of Sheffield, Department of Archaeology and Prehistory: 27-36.

Google Earth. (2015). Google Inc.

Grant, E. (1986) Hill-forts, Central Places and Territories. In: Grant, E. (Ed.) *Central Places, Archaeology and History.* University of Sheffield, Department of Archaeology and Prehistory: 13-26.

Harding, D.W. (Ed.) (1976) *Hillforts : later prehistoric earthworks in Britain and Ireland,* London, Academic Press.

Harding, D.W. (2012) *Iron Age Hillforts in Britain and Beyond,* Oxford, Oxford University Press.

Hodder, I. (1971) The Use of Nearest Neighbour Analysis. *Cornwall Archaeology,* 10: 35-36.

Hodder, I. & Orton, C. (1976) *Spatial Analysis in Archaeology*, Cambridge University Press.

Hogg, A.H.A. (1971) Some applications of surface fieldwork. In: Hill, D. & Jesson, M. (Eds.) *The Iron Age and its Hillforts.* University of Southampton:

Hogg, A.H.A. (1972) The Size-Distribution of Hill-forts in Wales and the Marches. In: Lynch, F. & Burgess, C. (Eds.) *Prehistoric Man in Wales and the West. Essays in honour of Lily F. Chitty.* Bath, Adams and Dart: 293-306.

Hogg, A.H.A. (1975) *Hill-Forts of Britain,* London, Hart-Davis, MacGibbon.

Hogg, A.H.A. (1979) *British Hill-Forts, An Index,* Oxford, British Archaeology Reports.

Irish Townlands, www.townlands.ie/page/download, Accessed: 2015-2016

Isle of Man Government - Maps, https://www.gov.im/maps/, Accessed: April 2016

Jones, B. & Mattingly, D.J. (1990) An atlas of Roman Britain. Oxford, Blackwell Reference.

Landscapes of Governance Project, http://www.ucl.ac.uk/archaeology/research/projects/assembly, Accessed: 2016

Llobera, M. (2001) Building Past Landscape Perception With GIS: Understanding Topographic Prominence. *Journal of Archaeological Science,* 28: 1005-1014.

Llobera, M. (2012) *Geoprocessing: Topographic Prominence,* unpublished.

Lock, G.R., Gosden, C. & Daly, P. (2005) *Segsbury Camp: excavations in 1996 and 1997 at an Iron Age Hillfort on the Oxfordshire Ridgeway*, University of Oxford. School of Archaeology.

Lynch, F., Aldhouse-Green, S. & Davies, J.L. (2000) *Prehistoric Wales*, Stroud : Sutton.

Lynch, F. & Burgess, C. (Eds.) (1972) *Prehistoric man in Wales and the West : essays in honour of Lily F. Chitty,* Bath, Adams and Dart

Mackie, E.W. (1976) The Vitrified Forts of Scotland. In: Harding, D.W. (Ed.) *Hillforts : later prehistoric earthworks in Britain and Ireland.* London, Academic Press: 205-235.

Maga Kernow, Cornish Language Partnership, http://www.magakernow.org.uk/, Accessed: 11/8/2016

Moel-y-Gaer, Bodfari, Denbighshire, http://www.arch.ox.ac.uk/bodfari.html, Accessed: 2013-2016

Morris, I. (2010) *Why the West Rules—For Now,* London, Profile Books.

Nash-Williams, V.E. (1933) An early Iron Age hill-fort at Llanmelin, near Caerwent, Monmouthshire : with a note on the distribution of hill-forts and other earthworks in Wales and the Marches. *Archaeologia Cambrensis Offprint,* 88: 237-346.

National Monuments Service, https://www.archaeology.ie/archaeological-survey-ireland, Accessed: 2016

Newcomb, R.E. (1968) Geographical Location Analysis and Iron Age Settlement in West Penwith. *Cornwall Archaeology,* 7: 5-14.

Newcomb, R.E. (1970) The Spatial Distribution of Hill Forts in West Penwith. *Cornwall Archaeology,* 9: 47-52.

Newcomb, R.E. (1971) Comments on 'The Use of Nearest Neighbour Analysis'. *Cornwall Archaeology,* 10: 37.

Northern Ireland Sites and Monuments Record, http://apps.ehsni.gov.uk/ambit/Default.aspx, Accessed: 2016

Ordnance Survey (1962) Ordnance Survey Map of Southern Britain in the Iron Age. Chessington, Ordnance Survey.

OS Open Rivers, www.os.uk/opendata, Accessed: August-September 2016

Payne, A., Corney, M. & Cunliffe, B. (2006) *The Hillforts of Wessex Project*, English Heritage.

Peters, C. (2005) *The Archaeology of Cornwall*, Cornwall Editions Ltd.

Pierson-Jones, J. (1973) Undergraduate Part II Thesis. *Department of Archaeology.* unpublished, Cambridge University.

Piggott, S. (1966) A Scheme for the Scottish Iron Age. In: Rivet, A.L.F. (Ed.) *The Iron Age in Northern Britain.* Edinburgh, Edinburgh University Press: 1-16.

The Portable Antiquties Scheme, https://finds.org.uk/, Accessed: 2016

Quinnell, H. (1986) Cornwall during the Iron Age and the Roman Period. *Cornish Archaeology,* No. 25: 111-134.

The R project for Statistical Computing, https://www.r-project.org, Accessed: 2015, 2016

Ralston, I. (2006) *Celtic fortifications,* Stroud, Gloucestershire, Tempus.

Restormel Roman Fort, http://www.pastscape.org.uk/hob.aspx?hob_id=432777, Accessed: 2014, 2016

Rivet, A.L.F. (1958) Some of the Problems of Hill-Forts. In: Frere, S. (Ed.) *Problems of the Iron Age in Southern Britain : papers given at a C.B.A. conference held at the Institute of Archaeology, December 12 to 14, 1958,.* University of London, Institute of Archaeology: 29-34.

Rivet, A.L.F. (Ed.) (1966) *The Iron Age in Northern Britain,* Edinburgh, Edinburgh University Press.

Roman fort found in Cornwall (2007) *British Archaeology* #96.

Rozenfeld, H.D., Rybski, D., Gabaix, X. & Makse, H.A. (2011) The Area and Population of Cities: New Insights from a Different Perspective on Cities. *American Economic Review,* 101: 2205-2225.

Savory, H.N. (1976) Welsh Hillforts: A Reappraisal of Recent Research. In: Harding, D.W. (Ed.) *Hillforts : later prehistoric earthworks in Britain and Ireland.* London, Academic Press: 237-292.

Sherratt, A. (1996) Why Wessex? The Avon Route and River Transport in Later British Prehistory. *Oxford Journal of Archaeology,* 15: 211-234.

Stauffer, D. & Aharony, A. (1991) *Introduction to Percolation Theory,* Boca Raton, USA, CRC Press.

Stout, M. (1997) *The Irish Ringfort,* Dublin, Four Courts Press.

Streetmap, www.streetmap.co.uk, Accessed: 2016

Treasure Trove Scotland, http://www.treasuretrovescotland.co.uk/, Accessed: 11/8/2016

UK Grid Reference Finder, http://www.gridreferencefinder.com/batchConvert/batchConvert.php, Accessed: 2015, 2016