Report No: 2012R057

Lowland Cornwall: the Hidden Landscape
Volume Two

The influence of additional factors

Historic Environment Projects

Lowland Cornwall: the Hidden Landscape
Volume Two

The influence of additional factors

Client	**Historic England**
Report Number	**2012R057**
Date	**August 2015**
Status	**Final version**
Report author(s)	**Andrew Young**

Historic Environment, Cornwall Council

Kennall Building, Old County Hall, Station Road, Truro, Cornwall, TR1 3AY

tel (01872) 323603 fax (01872) 323811 E-mail hes@cornwall.gov.uk

www.cornwall.gov.uk

Acknowledgements

This study was commissioned by English Heritage and carried out by Historic Environment Projects, Cornwall Council (HE). Valuable assistance in designing the spreadsheets used to create the predictive models was provided by Stuart Ladd, then an MA student at Bristol University on work placement with HE Projects. Analysis of the HE aerial photo database was carried out by Emma Trevarthen. Analysis of events record data was undertaken by Joanna Sturgess.

Freedom of Information Act

As Cornwall Council is a public authority it is subject to the terms of the Freedom of Information Act 2000, which came into effect from 1st January 2005.

Historic Environment, Cornwall Council is a Registered Organisation with the Institute for Archaeologists

Contents

List of Figures

Abbreviations

BGS	British Geological Survey
CAU	Cornwall Archaeological Unit
CC	Cornwall Council
CUCAP	Cambridge University Committee for Aerial Photography
EH	English Heritage
GIS	Geographic Information System
HER	Cornwall and the Isles of Scilly Historic Environment Record
HE	Historic Environment, Cornwall Council
HLC	Historic Landscape Characterisation
NGR	National Grid Reference
NMP	National Mapping Programme
NMR	National Monument Record
OD	Ordnance Datum
OS	Ordnance Survey
PRN	Primary Record Number in Cornwall HER
RAF	Royal Air Force

Summary

Cornwall's lowland areas probably have the highest archaeological potential in the county, but are poorly understood and increasingly subject to the impacts of major change in land use and development. The lowland Cornwall project attempts to address this issue by developing a method for predictive modelling of the lowland prehistoric and Romano-British landscape. The models produced by the project will better inform future management and land use decisions.

The project consisted of four stages: preparation of datasets and high level predictive models; deepening or refinement of Historic Landscape Characterisation (HLC); further analysis of the archaeological resource and the preparation of predictive models using the refined HLC; and the presentation of final results.

This report is the second of four volumes presenting the results of the Lowland Cornwall project. Volume 1 presents high level predictive models created by correlating the distribution of selected site types with HLC Types. This volume outlines an assessment of the extent to which additional factors, such as soils and geology, may influence known distribution patterns of below ground archaeology. This was achieved by creating three zone high level models based on correlations between cropmark distribution and Agricultural Land Classification data, geology and soil types. The geology and soil land classes were then joined with the pattern of aerial reconnaissance in Cornwall to produce a visibility map showing where below-ground archaeology is most likely to occur and where it is most likely to have been identified and recorded.

Agricultural Land Classification data proved ineffective for model building: the polygons representing each grade of land are highly generalised and, consequently, too schematic. There were also problems with the bedrock geology data derived from the British Geological Survey (BGS), in that in places where the data tiles meet, the rock types from neighbouring tiles do not always correspond.

Models were created using both the soils and the bedrock data and the two models are similar in that much of their high probability zones lie in central and western areas of the county whilst large areas to the east of the Fowey and Camel rivers are ranked in the medium or low probability zones. Both models perform with similar levels of accuracy, capturing around 70% of the cropmarks in their high probability zones, but neither is particularly precise.

The soils and bedrock data were combined and a complex new dataset was created. This produced a model more precise but less accurate than those produced by soils or bedrock on their own. The east/west divide was apparent in this model but with more of west Cornwall falling in the low probability zone than in either of the other two models.

A model was also produced correlating the distribution of known cropmark sites with the pattern of aerial reconnaissance. This was achieved by tracing the history of Cornwall's Historic Environment Service reconnaissance pattern. There are a number of weaknesses in the flight data used to create this model, but it does demonstrate that some areas have been flown many times and others not at all. The rarely flown areas are mostly in east and southeast Cornwall. The model produced by the reconnaissance pattern is precise but not very accurate; it clearly shows where cropmarks are most likely to have been recorded and depicts east Cornwall generally as an area of low probability.

A definitive cropmark visibility model was created by combining the soils/bedrock model with the aerial reconnaissance model. The visibility model is both accurate and reasonably precise and analysis of the way in which the polygons are ranked suggests that soils and geology are more influential than the aerial reconnaissance pattern. The cropmark visibility model provides a more nuanced version of the east/west divide, with an extensive area close to the southeast coast ranked in the high probability zone.

All five models were tested using events record data. Only events in which below ground features were recorded but which have not yet been input to the HER were used in the tests. All the models were rejected by the test sample to varying degrees. The final cropmark visibility model was the most clearly rejected, with the low probability zone capturing twice as many sites as predicted. So this model only shows those areas where cropmarks are most likely to form and where they are most likely to have been seen; it shows 'absence of evidence' rather than 'evidence of absence'.

One possible implication is that the known distribution of rounds and enclosures (of which more than half are recorded as cropmarks) may be biased towards the high probability zone of the cropmark visibility model. To explore this question further, models were made based on the correlation of rounds and enclosures with bedrock geology, soils and a combination of the two. Although it was possible to create a three zone model for rounds correlated with bedrock geology, it was difficult to define cut off points for medium and low probability zones for the models correlating rounds and enclosures with soils and with soils and bedrock combined: in effect these perform as two zone models with high and low probability zones. The soils model and the geology model both achieved good levels of accuracy and were reasonably precise; the soils and geology combined model was not as accurate but more precise.

All three models broadly resemble the cropmark models in that eastern parts of the county are generally classed as low or medium probability, although there are also differences. Unlike the cropmark models, however, when tested with events record data the test sample provided a close fit. Analysis of the form of survival of the enclosures captured in each of the probability zones shows that the percentage of extant enclosures closely resembles the percentage of all enclosures. Thus the great majority of extant enclosures are captured in the high probability zones of all three models, suggesting that if there is any bias towards cropmark enclosures it is only minimal. Also of interest is the distribution of rounds and enclosures over the two most widespread soil types, the Denbigh 1 and Denbigh 2 series. These are both very similar loams over shale and overlie similar rock types, yet three times as many enclosures are recorded from Denbigh 2 soils as from Denbigh 1. The main difference between the two soils is their regional distribution, with Denbigh 1 mainly confined to southeast Cornwall and Denbigh 2 predominantly occurring in central areas. Taken together these outcomes suggest that the high number of rounds and enclosures found on some rock and soil types reflects a deliberate preference for those locations and that the east/west disparity in the distribution of rounds and enclosures is real rather than being the result of factors influencing cropmark formation and visibility.

One weakness of the model presented in Volume 1, correlating rounds and enclosures with HLC Types is the lack of precision in its high probability zone due to its large size. To try to reduce this high probability zone by sub-division, the HLC layer was combined with the soils layer and the bedrock geology layer and new models were made by correlating the distribution of rounds and enclosures with these combined layers. A simplified model was also made by joining the soils/bedrock combined layer with the HLC layer. Analysis of the way in which the various polygons were ranked in these models suggests that HLC is the most influential factor in determining the rankings.

In all three of these models the high probability zone is defined with considerably more precision than in the rounds/HLC model, its extent ranging from 41-48% of the project area as opposed to 66%. Although this was achieved at the expense of some accuracy the models can be considered to attain an acceptable level of accuracy. The models also reflect regional disparities in enclosure distribution, with parts of eastern Cornwall classed as low or medium probability but in a more nuanced way than those based on soils and geology alone. The models based on bedrock and soils alone are notably broad brush, but when combined with HLC Types a much finer granularity is achieved.

All three models were verified to a greater or lesser extent when tested using events record data, with the high probability zone of the soils/HLC model in particular performing better than predicted.

1 Introduction

1.1 Project background

Cornwall's lowland areas probably have the highest archaeological potential in the county, but are poorly understood and increasingly subject to the impacts of major change in land use and development. The Lowland Cornwall project attempts to address this issue by developing a method for predictive modelling of the lowland prehistoric and Romano-British landscape. Predictive models will better inform future management and land use decisions and increase confidence in responses to development proposals in areas where the Historic Environment Record (HER) currently shows no below-ground features. The method may also have the potential for application in other parts of the country.

The project comprises an appraisal of currently available data from a range of sources in order to develop models of past land-use, settlement patterns and landscape development. Whilst the primary aim is to indicate areas of high archaeological potential, at the same time it addresses key research agenda and contributes towards developing our understanding of historic landscape character.

The idea for the project was developed from a series of discussions with the County Archaeologist and other senior officers within Historic Environment, Cornwall Council (HE), and with the English Heritage South West regional and Characterisation teams. The project was commissioned by English Heritage (EH) following the submission of a project design in early 2009 (Young 2009).

1.2 Aims and objectives

1.2.1 Aims

1. To demonstrate the potential and significance of below-ground archaeology in lowland Cornwall, in particular to develop a better understanding of the extent and character of the prehistoric and Romano-British landscape. This improved understanding will better inform both development control and management and land use decisions in lowland Cornwall, the latter by highlighting those areas with high archaeological potential and thus higher priority in terms of most effective targeting of agri-environment schemes and other landscape-scale management initiatives. On a strategic level the better understanding and predictive modelling resulting from the project will provide a more meaningful context in which to specify the scope of future PPS 5 work and to assess the results of such work.
2. To define models for prehistoric settlement patterns and landscape development in lowland Cornwall and by exploring the relationship between these patterns and the early medieval and medieval patterns of settlement and land use, gain a better understanding both of the development of Cornwall's early society and economy and of the character and patterning of the county's buried archaeological remains.
3. To test and review interpretations of the development and potential of historic landscape character types.

1.2.2 Objectives

1. To review currently available HER, National Mapping Programme (NMP) and Events Record data. In particular to examine the range of settlement types, evidence for field systems and land use, and evidence for phasing and change.
2. To propose models for prehistoric settlement patterns and landscape development by linking the results of this review with Historic Landscape Characterisation (HLC) data to identify patterns in settlement distribution, in the spatial relationships between settlements and field systems, and in the relationships between areas of intense activity and areas which are apparently blank.

3. To review current interpretations of the development and potential of historic landscape character types by better defining the extent of Anciently Enclosed Land and Recently Enclosed Land HLC Types.

1.3 Report layout

The project comprised three distinct stages and generated an enormous amount of data. In order to present the results of the project in an accessible format, the final report is published as five separate volumes.

1.3.1 Volume 1

During stage one data for selected site types was extracted from the Cornwall HER and correlated with the existing HLC Types in order to identify recurring distribution patterns and to create high level predictive models. Volume 1 presents the outcome of this work, describing the methodology used to create the models, the results of the modelling and a discussion and set of conclusions drawn from this research. Volume 1 also outlines the background to, and scope of, the whole project.

1.3.2 Volume 2

Also during stage one an assessment was made of the extent to which additional factors, such as soils and geology, may influence known distribution patterns of below-ground archaeology. Further high level models were built based on correlations between site distribution and geology and soil types. The distribution of geology and soils was then joined with the pattern of aerial reconnaissance in Cornwall to produce a cropmark visibility map showing where below-ground archaeology is most likely to occur and where it is most likely to have been identified and recorded. More than half of the rounds and enclosures in lowland Cornwall are recorded as cropmarks and additional models were made based on the correlation of their distribution with that of soil and geology types for comparison with the models for cropmarks generally. Further research was carried out into rounds and enclosures by combining soil and geology types with HLC Types and correlating these combinations with the distribution of rounds and enclosures to create combined models. Volume 2 (this volume) presents the results of this research.

1.3.3 Volume 3

Stage two involved refining or deepening HLC in four selected study areas. The HLC refinement comprised a more detailed analysis than that carried out for Cornwall's existing HLC. Specifically, some HLC Types were broken down into sub-types and characterisation was carried out for a number of time slices. The results of HLC refinement are presented in Volume 3.

1.3.4 Volume 4

Stage three involved building predictive models based on correlations between site distribution and the refined HLC Types and sub-types, to see whether more accurate and precise models could be achieved using the refined HLC. A detailed analysis of the sites within each study area was also produced. Volume 4 presents the results of this work and contains the overall conclusions arising from the project.

1.3.5 Volume 5

Volume 5 presents a summary of the information contained in Volumes 1 – 4.

2 Background to Volume 2

The results presented in Lowland Cornwall Volume 1 include a series of high level predictive models for selected site types, built by correlating known site distribution against HLC Types. The models identify those areas where sites are most or least likely to occur as zones of high, medium or low probability.

One outcome of constructing predictive models for the distribution of known sites using HLC Types as the sole variable is that whilst some of the models accurately indicate the areas where sites are most likely to be found, they provide little in the way of precision. In many of the models the zone of high probability covers a substantial portion of the project area (e.g. Lowland Cornwall Volume 1, fig 21). One effect of this lack of precision is the failure of the models to reflect regional variations in the known distribution of certain site types. The known locations of rounds and enclosures, for instance, cluster around the Camel and Helford Estuaries and are relatively lacking in east Cornwall, but the high level model based on HLC Types suggests that parts of east Cornwall are within the zone of high probability.

A very high percentage of the enclosures forming the clusters mentioned above have no above-ground remains surviving and are recorded in the HER as cropmarks. Not all types of geology and soils are conducive to cropmark production and a likely explanation for the apparent failure of the predictive models to reflect regional variations in site distributions is that the distribution patterns are skewed by variations in underlying geology, soils and land use. It is also plain that the pattern of aerial reconnaissance will have influenced the likelihood of cropmarks being seen just as variations in geology, soils and land use influence the likelihood of cropmarks being formed.

For these reasons a cropmark visibility map was produced showing those areas where cropmarks are most likely to form and where they are most likely to have been identified. As a separate procedure, predictive models were built based on the distribution of known enclosures correlated with the land classes mentioned above and, as a further step, models were built based on a combination of these land classes and HLC Types to create a more complex set of variables. The results of these models are presented in this volume.

3 Prehistoric and Romano-British cropmark sites

The assessment was carried out by displaying HER point data for cropmark sites in GIS and spatially correlating this with the relevant land classes.

In total the dataset contains records for 1,759 prehistoric or Romano-British sites listed as cropmarks. Their overall distribution of in Lowland Cornwall is shown in Fig 1 below.

The distribution is characterised by significant concentrations in parts of central Cornwall and the northern part of the Lizard peninsula; cropmark distribution in east and southeast Cornwall by contrast is sparse.

The cropmark dataset is dominated by rounds and enclosures, whose substantial enclosing ditches are more likely to produce visible cropmarks than the relatively slight ditches forming field systems or the gullies of round houses, which are notoriously difficult to identify as cropmarks. This is confirmed by the fact that only 24 round houses are listed in the dataset. A full breakdown of site types in the cropmarks dataset is outlined in table 1 below.

Site type	No. of sites	% of total
Round/enclosure	1,047	59.5%
Barrow	398	23%
Field system	286	16%
Hut circle/round house	24	1.3%
Hillfort	4	0.2%
Total	**1,759**	

Table 1. Breakdown of site types contained in the cropmark dataset.

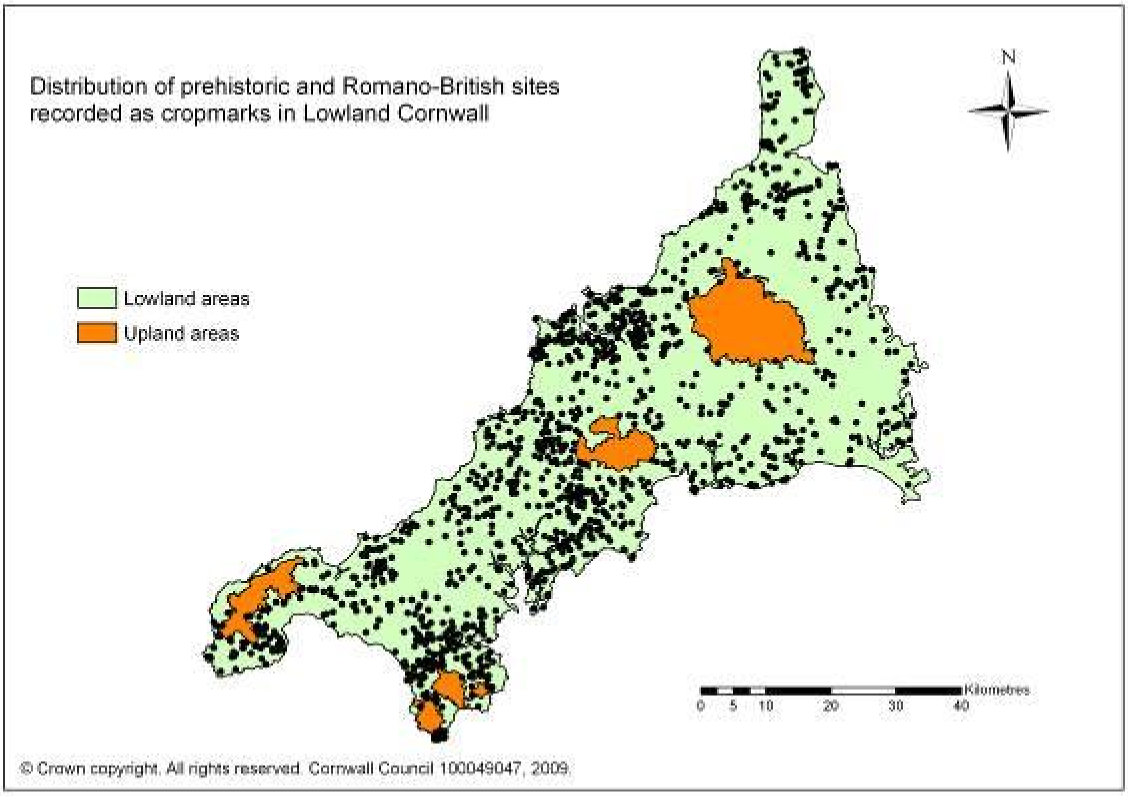

Fig 1. Overall distribution of prehistoric cropmark sites in Lowland Cornwall.

4 Land classes used as variables

Three land classes were used as the variables for model building: Agricultural Land Class (ALC), bedrock geology and soil types. At the outset of the project the intention was also to use the 1995 Land Cover data produced by the Cornwall Wildlife Trust, showing areas of arable and improved grassland. There were however, technical problems attempting to join HER data with this layer and ultimately it was not used. It should be pointed out that the Land Cover layer may not have been as useful as might be expected because arable in Cornwall does not just equate to cereal crops but includes large acreages of potatoes, cabbage, cauliflower, daffodils and maize, none of which produce cropmarks.

4.1 Agricultural Land Class (ALC)

This data is derived from the Agricultural Land Classification of England and Wales (Ministry of Agriculture, Fisheries and Food 1988). The layer contains 11 different categories described below.

ALC Grade	Definition
Grade 1	Excellent quality agricultural land
Grade 2	Very good quality agricultural land.
Grade 3	Good to moderate quality agricultural land
Grade 3A	Good quality agricultural land
Grade 3B	Moderate quality agricultural land
Grade 4	Poor quality agricultural land
Grade 5	Very poor quality agricultural land
Non-agricultural	'Soft' uses where most of the land could be taken back into agriculture (e.g. golf courses, parkland)
Not surveyed	Land not surveyed
Other	Includes woodland and open water
Urban	Built-up or 'hard' uses with little potential for a return to agriculture

Table 2. Categories contained in the ALC layer.

The distribution of ALC classes is shown below in Fig 2. More than 83% of the project area is covered by only two of the categories (Grade 3, covering 64.7% and Grade 4, covering 18.8%). Furthermore the polygons forming each category are highly generalised and their boundaries are consequently somewhat schematic (Fig 3). This strongly implies that any model derived from this layer must be judged as approximate only.

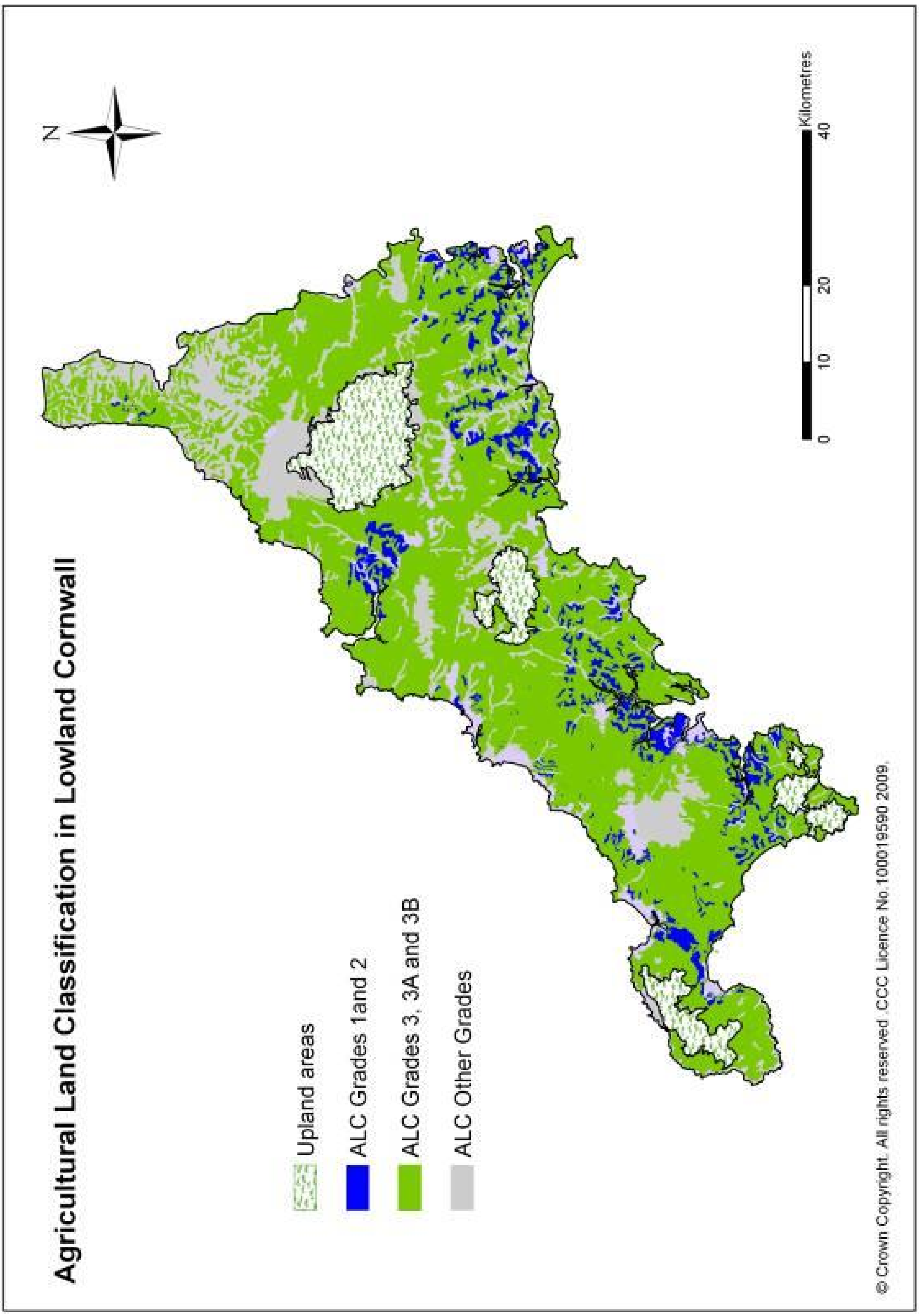

Fig 2. Agricultural Land Classification in Lowland Cornwall.

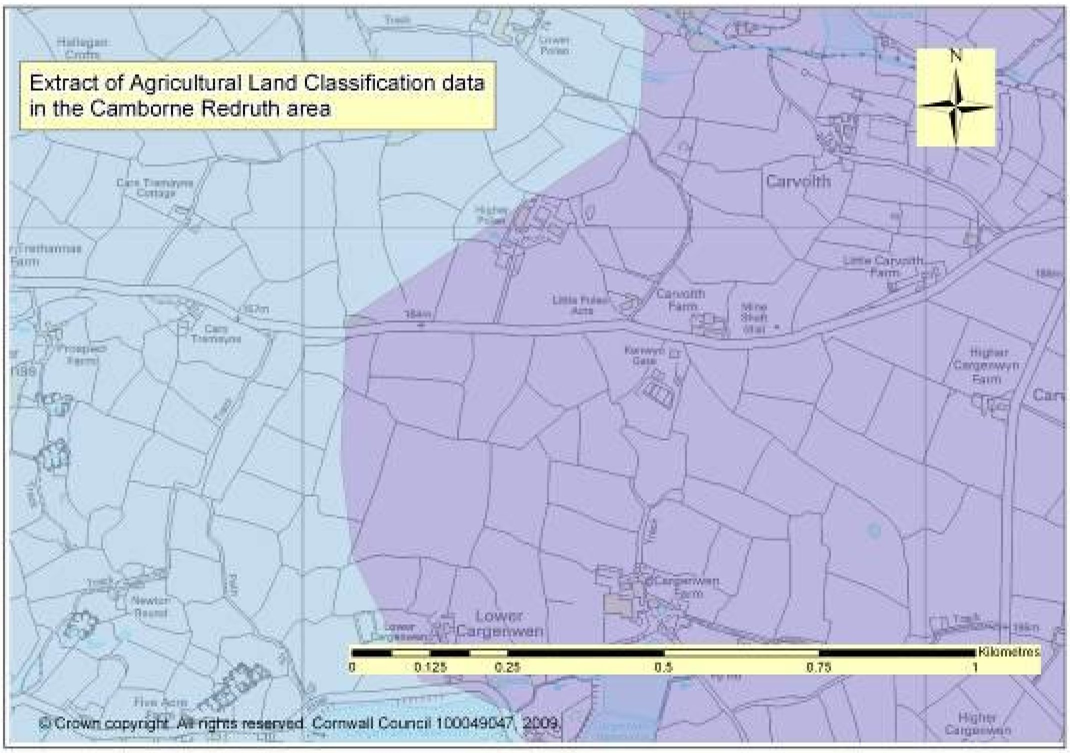

Fig 3. An extract of ALC data showing its generalised nature. Here Grade 4 land (shown in purple) on the right meets Grade 3 land (in blue) on the left. The boundary between the two Grades is not at all 'map sensitive', sweeping broadly through the landscape, and should be regarded as broad brush and schematic.

4.2 Soil types

The GIS soils layer is derived from the Soil Survey of England and Wales. The classification system includes 'Simple description', comprising 25 different classes. Of these, types described as *loam over shale* are by far the most extensive, covering 61% of the project area; the next largest category, *loam over granite*, covers only 10%. These classes are sub-divided into 33 soil types listed under their 'Map unit name'. Map unit name, being more detailed, are the most appropriate for using in model construction and these, along with the appropriate simple descriptions are listed below in table 3. The soils map for lowland Cornwall is shown in Fig 4; for simplicity this map shows the simple descriptions rather than map unit names.

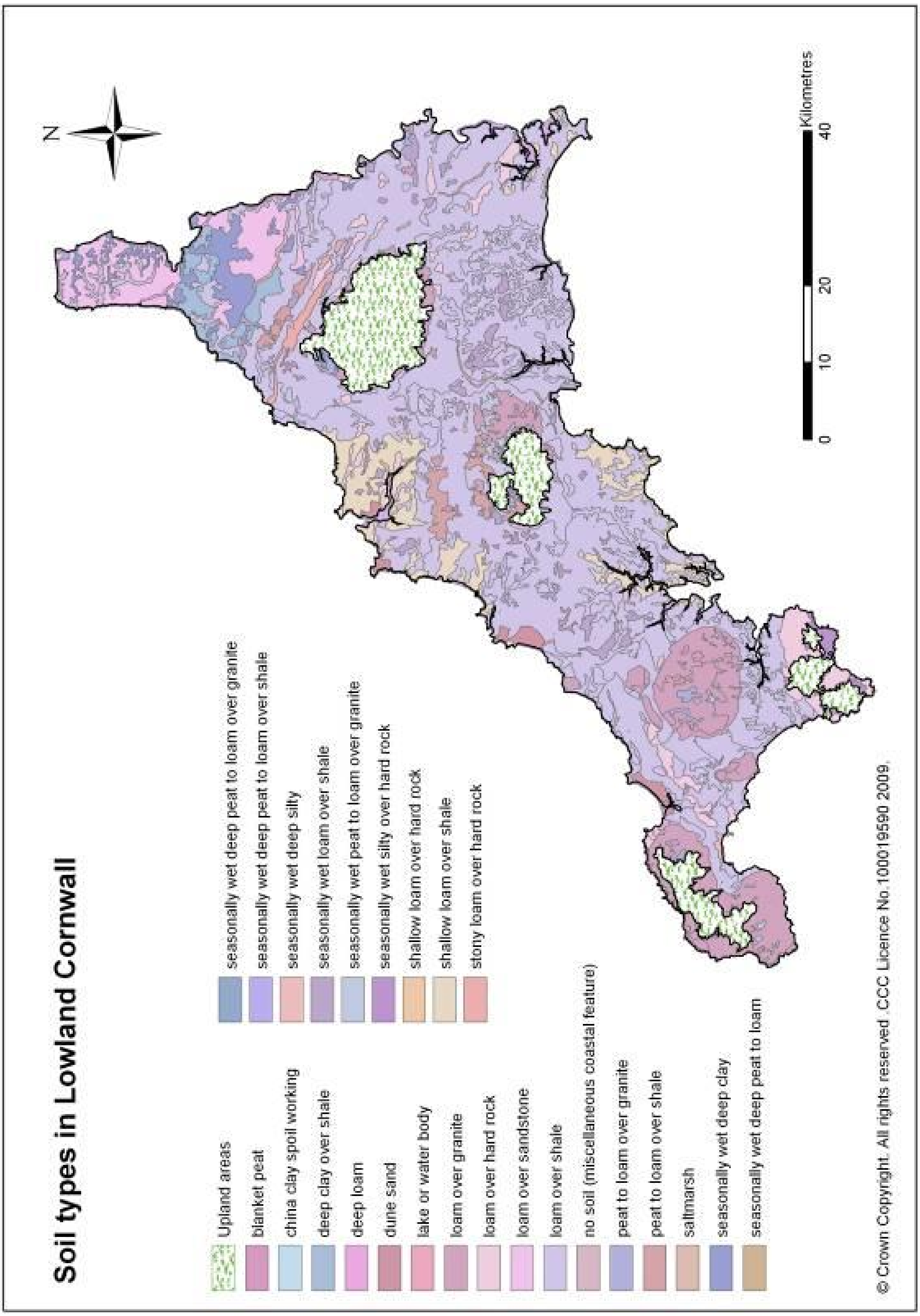

Fig 4. The soils map for lowland Cornwall.

Map unit name	Simple description	Area km^2	% of area
Denbigh 2	Loam over shale	940.205	29.48%
Denbigh 1	Loam over shale	686.088	21.51%
Manod	Loam over shale	334.462	10.49%
Moretonhampstead	Loam over granite	234.218	7.34%
Powys	Shallow loam over shale	192.717	6.04%
Neath	Loam over sandstone	144.519	4.53%
Trusham	Loam over hard rock	126.781	3.97%
Moor gate	Loam over granite	89.175	2.80%
Hallsworth 1	Seasonally wet deep clay	61.607	1.93%
Sportsmans	Seasonally wet loam over shale	58.756	1.84%
Halstow	Deep clay over shale	57.202	1.79%
Hafren	Peat to loam over shale	46.807	1.47%
Yeollandpark	Seasonally wet loam over shale	33.278	1.04%
Sandwich	Dune sand	25.91	0.81%
Laployd	Seasonally wet peat to loam over granite	23.579	0.74%
Hexworthy	Peat to loam over granite	18.732	0.59%
Croft Pascoe	Seasonally wet silty over hard rock	16.196	0.51%
Sea	Sea	14.658	0.46%
Crowdy 2	Blanket peat	7.71	0.24%
Princetown	Seasonally wet deep peat to loam over granite	1.771	0.06%
Raw china clay spoil	China clay spoil working	1.489	0.05%
Hallsworth 2	Seasonally wet deep clay	20.168	0.63%
Malvern	Stony loam over hard rock	18.697	0.59%
Conway	Seasonally wet deep silty	15.358	0.48%
Lake	Lake or water body	3.863	0.12%
Teme	Deep loam	6.989	0.22%
Larkbarrow	Loam over sandstone	5.528	0.17%
Onecote	Seasonally wet deep peat to loam over shale	1.199	0.04%
Wick 1	Deep loam	1.186	0.04%
Hense	Seasonally wet deep peat to loam	0.524	0.02%
Saline 1	Saltmarsh	0.438	0.01%
Mcf	No soil (miscellaneous coastal feature)	0.001	0.00%
Dunwell	Shallow loam over hard rock	0.001	0.00%
Total		3189.8	

Table 3. List of the soil types contained in the GIS soils layer, showing their map unit name, simple description and their total extent in km^2.

4.3 Bedrock geology

The GIS geology layer is based on data from the British Geological Survey (BGS). This comprises two principal elements; bedrock (solid geology) and superficial (drift geology) data. Superficial geology consists predominantly of alluvium and head deposits and is largely confined to the river valleys (although there are a few more extensive alluvium deposits, for instance at Goss Moor, and some extensive deposits of blown sand, as at Perranporth). Much of lowland Cornwall is devoid of superficial deposits and therefore only the bedrock data was used for building the models.

The BGS bedrock data is more detailed than that for either soils or Agricultural Land Classification. In the classification system there are 107 different rock types listed under the classification LEX_D. Another category, RCS_D, simplifies these into 79 types and was the classification used to build the models.

A significant problem with BGS data is that in places where the data tiles meet, the geology types from neighbouring tiles do not always correspond. The most obvious examples are in the central Cornwall area (Fig 5).

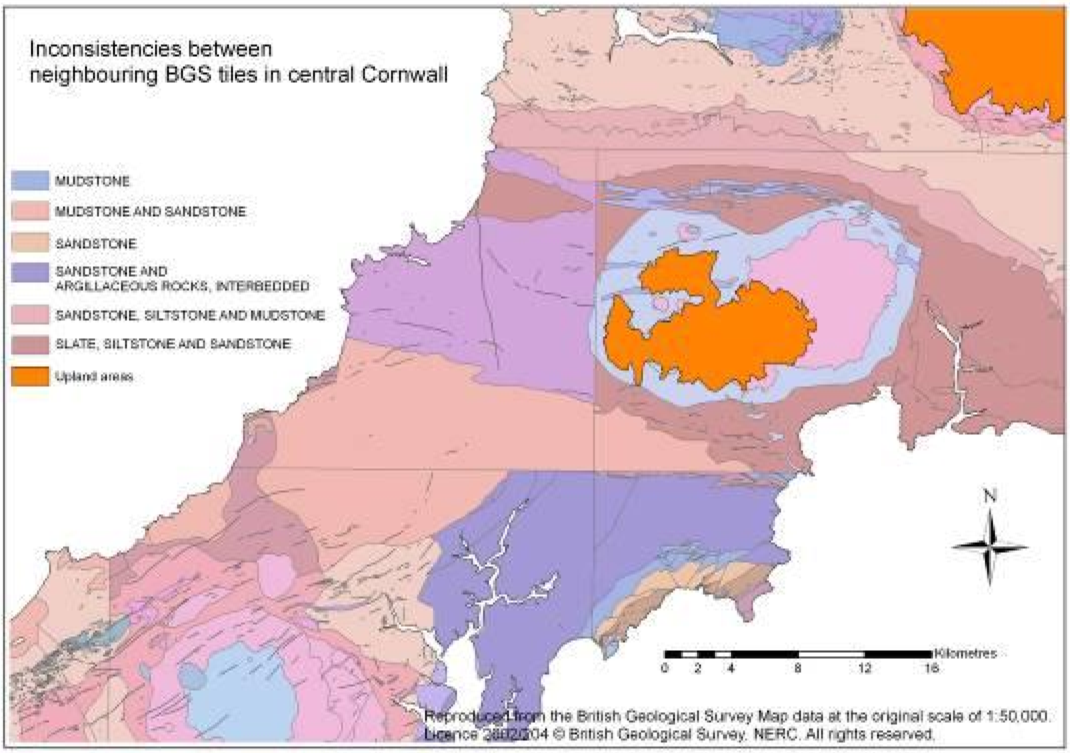

Fig 5. Extract of BGS bedrock mapping illustrating how adjoining data tiles do not always correspond.

In the extract shown in Fig 5 Sandstone and argillaceous rocks (shown in purple) in the two southernmost tiles do not continue into the tiles immediately to the north, but are replaced by Sandstone. Similarly data shown in the eastern and western tiles in the middle portion of this map do not correspond. Despite these inconsistencies, BGS data provides a reasonably good base layer for predictive modelling purposes.

5 The predictive models

5.1 Methodology

The methodology used for model construction was exactly the same as for the HLC models. This is outlined fully in Lowland Cornwall Volume 1, 37-48 and is summarised here.

The first step in creating viable models from the site data is to establish that the distributions apparent from the data analysis are statistically significant; that is that they are not merely representing by-chance patterns. In order to establish statistical significance the X^2 test (or Chi-Squared test) was used. X^2 is a standard statistical procedure (Lowry 2009). X^2 tests carried out for cropmarks found that when correlated with ALC classes, soil types and bedrock geology types, the null hypothesis was rejected in each case, indicating that the distributions are statistically significant (for results of X^2 testing, see Appendix 1).

Having established statistical significance, the Kj parameter was applied to measure the relative importance of each ALC class, soil type and rock type in site probability. The Kj parameter is defined as: √ (PS x (PS-PA)/PW). PS = the proportion of sites captured within each ALC, soil or rock type, PA = the proportion of lowland Cornwall taken up by each land type and PW is the proportion of the area that does not include sites. Because the high level predictive models for Lowland Cornwall used point data for sites rather than areas (polygons) this factor can be ignored (see Verhagen and Berger 2007).

The performance of the models is gauged using a number of gain measures. Foremost among these is Kvamme's gain: Gain = 1-(PA/PS). An important point about Kvamme's gain is that because PA/PS can never = 0, Kvamme's Gain can never reach the maximum 1: there is therefore always a maximum gain dependent on the model itself. A simpler measure is Relative Gain: PS-PA, resulting in theoretical values ranging from 1 to -1 (Wansleeben and Verhart 1992).

The ratio of proportion of sites (PS) to proportion of area (PA) is a straightforward way to measure importance. This formula, PS/PA, is known as the Indicative Value and it can be used internally to compare the performance of each probability zone against that of the others. An even simpler measure of site density is S/A – a calculation of the number of sites per square kilometre.

The concepts of accuracy and precision in terms of predictive modelling are important to grasp. Accuracy is a measure of correct prediction – are most of the sites captured in the high probability zone? Precision is a measure of how far the model has limited the high probability zone to as small an area as possible.

5.2 The cropmarks/ALC model

As described in section 4.1 above more than 83% of the project area is covered by only two of the ALC categories. As a result the model derived from ALC data lacks any precision: 81% of the cropmarks are located in these two categories. The high probability zone is formed by ALC Grades 2, 3 and 3B. Full details of the model are contained in the table below.

Cropmarks and ALC model: High probability zone					
ALC Grade	**Cropmarks**	**PA**	**PS**	**Cum Kj**	**Kvamme's gain**
G3	1283	0.6476	0.7306	0.2463	0.1136
G2	251	0.0882	0.1429	0.3469	0.1577
G3B	27	0.0092	0.0154	0.3576	0.1619
Totals	**1561**	**0.745**	**0.889**	**0.3576**	**0.1619**

Cropmarks and ALC model: Medium probability zone					
G4	148	0.1887	0.0843	0.1959	0.0405
G3A	13	0.0088	0.0074	0.1932	0.0388
Totals	**161**	**0.198**	**0.092**	**0.1932**	**-1.1541**
Cropmarks and ALC model: Low probability zone					
ALC Grade	**Cropmarks**	**PA**	**PS**	**Cum Kj**	**Kvamme's gain**
G5	9	0.0106	0.0051	0.1793	0.0331
Urban	10	0.0174	0.0057	0.1440	0.0211
Non-agricultural	12	0.0220	0.0068	0.0760	0.0058
G1	1	0.0009	0.0006	0.0741	0.0055
Other	2	0.0063	0.0011	0.0194	0.0004
Not surveyed	0	0.0002	0.0000	0.0000	0.0000
Totals	**34**	**0.057**	**0.019**	**0**	**-1.9544**

Table 4. Model for prehistoric and Romano-British cropmark sites correlated with Agricultural Land Classification data.

The individual PA and PS values for each ALC Grade are listed in the tables. In the high probability zone the PS values exceed the PA values for each ALC Grade. Thus Grade 3 covers 64.7% of the project area and contains 73% of the sites. For ALC Grades ranked in the medium and low probability zones the opposite is the case: Grade 4, for instance, covers 18.8% of the project area but only contains 8.4% of the sites.

Kj values are calculated for each ALC Grade to define the order in which the Grades are ranked in the model. The cumulative Kj values (these are listed as Cum Kj in the tables) are used to identify the cut off points for each of the three zones.

The performance of the model is summarised below.

Probability	PA	PS	Kvamme's gain	PS/PA
High	0.74	0.89	0.1614	1.19
Medium	0.20	0.09	-1.1541	0.46
Low	0.06	0.02	-1.9544	0.33

Table 5. Summary of the overall performance of the ALC model for cropmark sites.

The Kvamme's gain for the high density zone is low. Whilst the model is accurate in that 89% of the sites are captured in the high probability zone it lacks precision because this zone covers 74% of the project area.

Further assessment of model quality can be made by comparing the performance measures of the three zones using the ratio of the indicative value (PS/PA) for each zone (Verhagen 2007). The ratio of indicative values for the high and medium density zones is 1.19/0.46 = 2.6 and for the high and low density zones is 1.19/0.33 = 3.6. This means that the probability of encountering a site in the high density zone is 2.6 times higher than in the medium density zone and 3.6 times higher than in the low density zone. But the probability of encountering a site in the medium density zone is only 1.4 times higher than in the low density zone.

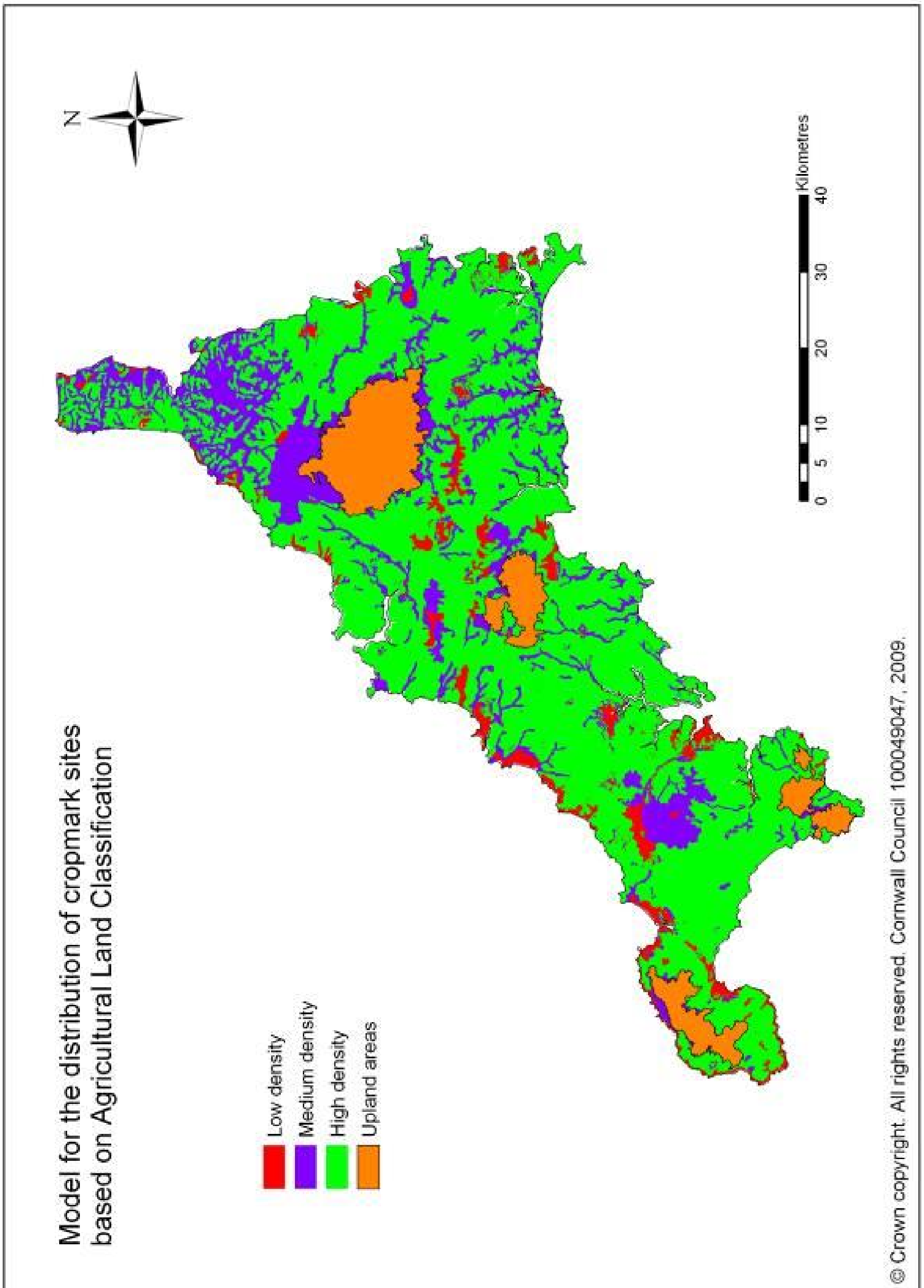

Fig 6. Probability map based on the ALC model for cropmark sites in Lowland Cornwall.

A probability map based on the model is shown in Fig 6. This shows that a spatial correlation of cropmark sites with the ALC layer suggests that cropmarks are likely to be found anywhere apart from urban areas, some coastal areas, steep sided river valleys and high ground. The most informative aspect of the model is the defining of high ground around Carnmenellis and in northeast Cornwall as being outside the high probability zone. Given the generalised nature of the ALC model it can be concluded that Agricultural Land Classification is unlikely to add much critical information to the question of cropmark visibility. For this reason ALC data was not included in the final visibility model.

5.3 The cropmarks/soil types model

When correlated with HER cropmark site data the soils layer provided a good model, details of which are outlined in table 6 below.

Cropmarks/soils. High probability zone				
Soil type	**Cropmarks**	**PA**	**PS**	**Cum Kj**
Denbigh 2	700	0.2948	0.3980	0.2027
Powys	238	0.0604	0.1353	0.3082
Moretonhampstead	159	0.0734	0.0904	0.3488
Trusham	110	0.0397	0.0625	0.3866
Sportsmans	48	0.0184	0.0273	0.4022
Totals	**1255**	**0.4867**	**0.7135**	**0.4022**
Cropmarks/soils. Medium probability zone				
Denbigh 1	221	0.2151	0.1256	0.3394
Neath	61	0.0453	0.0347	0.3326
Hallsworth 1	41	0.0193	0.0233	0.3423
Croft Pascoe	21	0.0051	0.0119	0.3535
Yeollandpark	23	0.0104	0.0131	0.3595
Hallsworth 2	13	0.0063	0.0074	0.3623
Totals	**380**	**0.3015**	**0.216**	**0.3623**
Cropmarks/soils. Low probability zone				
Moor gate	21	0.0280	0.0119	0.3433
Hafren	15	0.0147	0.0085	0.3363
Manod	59	0.1049	0.0335	0.2167
Halstow	12	0.0179	0.0068	0.1904
Larkbarrow	5	0.0017	0.0028	0.1936
Sandwich	6	0.0081	0.0034	0.1814
Crowdy 2	2	0.0024	0.0011	0.1779
Hexworthy	2	0.0059	0.0011	0.1642
Onecote	1	0.0004	0.0006	0.1649
Laployd	1	0.0074	0.0006	0.1427
Dunwell	0	0.0000	0.0000	0.1427
Mcf	0	0.0000	0.0000	0.1427
Saline 1	0	0.0001	0.0000	0.1422
Hense	0	0.0002	0.0000	0.1417
Wick 1	0	0.0004	0.0000	0.1403
Raw china clay spoil	0	0.0005	0.0000	0.1387
Princetown	0	0.0006	0.0000	0.1367
Lake	0	0.0012	0.0000	0.1321
Teme	0	0.0022	0.0000	0.1236

Sea	0	0.0046	0.0000	0.1033
Conway	0	0.0048	0.0000	0.0766
Malvern	0	0.0059	0.0000	0.0000
Totals	**124**	**0.2119**	**0.0703**	**0**

Table 6. Model for the distribution of cropmarks correlated with soils data.

All of the soil types ranked in the model's high probability zone are loams. This is to be expected because these soils are the most fertile and therefore the most likely to be currently under arable cultivation (although it should be noted that Sportsmans is described as a seasonally wet loam). Cereal production is most likely to take place on these soils – hence the high probability of cropmarks being observed. It is also possible that because the loams are the best soils prehistoric settlements were concentrated here.

The medium probability zone is made up of loams (Denbigh 1 and Neath), seasonally wet loam (Yeollandpark), clays (Hallsworth 1 and 2) and seasonally wet silty and fine loamy soils (Croft Pascoe).

The performance of the model is summarised below in table 7. The model is accurate, capturing 71% of the cropmarks in the high probability zone. It is not altogether precise in that the high probability zone covers 49% of lowland Cornwall, but does produce a reasonable Kvamme's gain. The medium probability zone performs well (containing less than a quarter of the sites), although its weakness is that it covers a relatively large area, and the low probability zone is defined precisely and accurately.

Indicative values (PS/PA) show that the chance of encountering a site in the high probability zone is twice as high as in the medium zone and four times higher than in the low probability zone. The chances of encountering a site in the medium probability zone are twice as high as in the low probability zone.

Probability zone	PA	PS	Kvamme's gain	PS/PA
High	0.49	0.71	0.3178	1.47
Medium	0.30	0.22	-0.3958	0.72
Low	0.21	0.07	-2.0028	0.33

Table 7. Summary of the cropmark/soils model performance.

The probability map derived from the model is shown in Fig 7. The most obvious feature is the contrast between east Cornwall and the rest of the county. Much of the county east of the rivers Fowey and Camel is classed as medium or low probability, whilst the greater part of Lowland Cornwall west of the Fowey contains extensive areas classed as high probability. Within this overall picture there are exceptions to the general trend: in the east the land around the fringes of Bodmin Moor are in the high probability zone; in the west areas overlying granite or metamorphic aureole are classed as zones of low probability. In central Cornwall the same is true of the land bordering the Hensbarrow granite, as well as the high ground of St Breock Downs. In west and central Cornwall there are some areas of medium probability (especially around Truro and Falmouth) but, for the most part, the medium probability zone here comprises the river valleys.

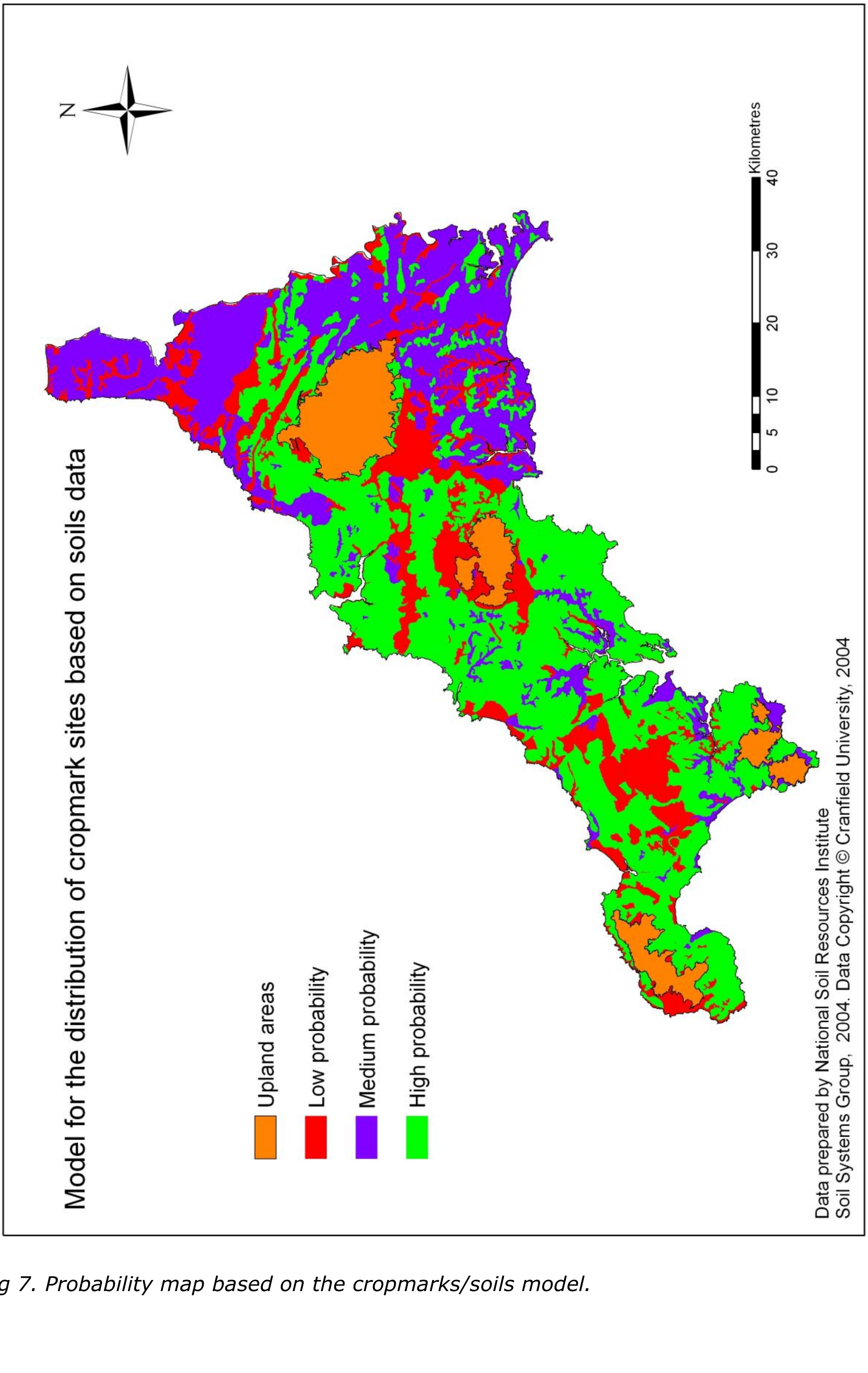

Fig 7. Probability map based on the cropmarks/soils model.

5.4 The cropmarks/bedrock geology model

Details of the model are shown in table 8 below.

Cropmarks/bedrock geology model. High probability zone				
Bedrock	**Cropmarks**	**PA**	**PS**	**Cum KJ**
Slate and siltstone	328	0.1531	0.1868	0.0793
Sandstone and [subequal/subordinate] argillaceous rocks, interbedded	213	0.0750	0.1213	0.1570
Mudstone	84	0.0132	0.0478	0.2020
Mudstone and sandstone	137	0.0643	0.0780	0.2360
Mudstone, siltstone and sandstone	80	0.0135	0.0456	0.2774
Granite	160	0.0921	0.0911	0.3016
Slate and sandstone, interbedded	99	0.0407	0.0564	0.3314
Sandstone	69	0.0449	0.0393	0.3361
Hornblende schist	34	0.0066	0.0194	0.3536
Totals	**1204**	**0.5034**	**0.6857**	**0.3536**
Cropmarks/bedrock geology model. Medium probability zone				
Slate, siltstone and sandstone	129	0.1059	0.0735	0.3374
Hornfelsed slate and Hornfelsed siltstone	64	0.0503	0.0364	0.3291
Slaty mudstone with sedimentary rock, metamorphic rock and igneous rock clasts	24	0.0066	0.0137	0.3404
Microgabbro	21	0.0048	0.0120	0.3514
Mudstone and siltstone	93	0.0896	0.0530	0.3153
Peridotite and Serpentinite	21	0.0060	0.0120	0.3257
Basaltic lava	10	0.0047	0.0057	0.3281
Totals	**362**	**0.2679**	**0.2063**	**0.3281**

Table 8. Model for the distribution of cropmarks correlated with bedrock geology: high and medium probability zones.

Probability zone	PA	PS	Kvamme's gain	PS/PA
High	0.50	0.69	0.2660	1.36
Medium	0.27	0.21	-0.2857	0.77
Low	0.23	0.10	-1.1155	0.43

Table 9. Performance summary of the cropmarks/bedrock geology model.

This model does not perform as well as the soils model, with a Kvamme's gain of 0.266 as opposed to 0.3178. A smaller proportion of sites (69% compared with 71%) is captured in the high probability zone of the geology model and the high probability zones of both models are similar in size (50% in the geology model and 49% in the soils model), hence the lower Kvamme's gain. A higher proportion of sites are captured in the low probability zone than in the soils model (10% against 7%).

The probability map based on bedrock geology (Fig 8) also shares some aspects with the map based on soils (Fig 7) – most notably that much of the high probability zone lies in western and central areas. There are, however, significant differences. In the bedrock model an extensive tract of land classed in the high probability zone runs in a band from the Camel Estuary to the south of Bodmin Moor, and as far east as the county boundary at Saltash: this area is in either the medium or low probability zone of the soils model. There is a strongly defined east to west band of land classed as low probability zone in the bedrock model (running from Tintagel to Launceston): in the soils model much of this area is classed as either medium or high probability. The area around Bude and Week St Mary, overlying Mudstone and siltstone, is classed as high probability whereas in the soils model the same area is placed in the low or medium probability zones. Further west there is a series of comparable differences between the two models.

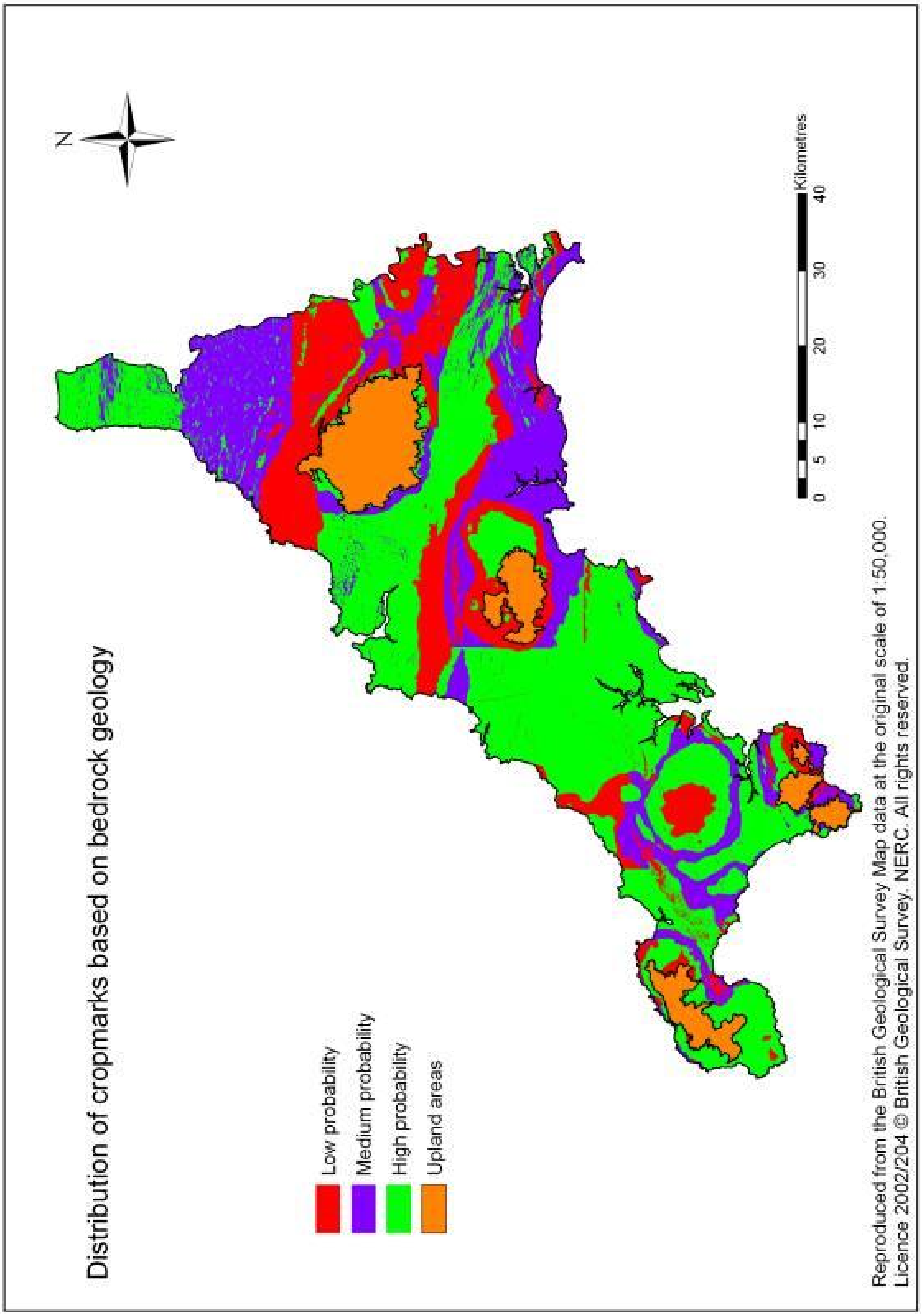

Fig 8. Probability map based on bedrock geology model for cropmarks in Lowland Cornwall.

On a broader level the bedrock probability map appears to be relatively broad brush in character and lacking subtlety, with a series of unbroken tracts of land falling into the various zones in contrast with the soils map in which the patterns of the three probability zones are far more inter-mingled.

An obvious flaw in the map is caused by the lack of agreement between some neighbouring data tiles (mentioned above and illustrated in Fig 5). This carries over into the probability map and results in abrupt boundaries to the zones in a number of places, most clearly the straight north–south boundary to the high probability zone west of the Hensbarrow uplands.

5.5 The cropmark/soils and bedrock geology model

The soils probability map and the geology map display broad similarities but also a number of differences. For example the rock type granite is frequently overlain by the soil type Moor gate and also by the soil type Moretonhampstead. Granite is classed as part of the high probability zone in the bedrock model: in the soils model Moretonhampstead is classed as part of the high probability zone but Moor gate is placed in the low probability zone. To reconcile these differences a model based on a combination of both sets of data was produced. To do this the soils and bedrock geology layers were amalgamated by performing a spatial union in GIS. A model was then produced based on the combined polygons. (In the bedrock and soils combined model granite/Moretonhampstead is placed in the high probability zone and granite/Moor gate is classed as medium probability).

The combined layer consisted of 558 different combinations of soils and bedrock geology of which 134 contain cropmark sites. The performance of the model is summarised below.

Probability	PA	PS	Kvamme's gain	PS/PA
High	0.3085	0.5369	0.4255	1.74
Medium	0.2820	0.2563	-0.1004	0.91
Low	0.4096	0.2068	-0.9803	0.50

Table 10. Performance of the cropmarks/soils and bedrock geology model.

This model produces a better Kvamme's gain than either the soils or geology models on their own. However the high probability zone has only captured 53% of the sites and the low probability zone has captured 20% of the sites - only 5% fewer than the medium probability zone. Thus whereas this model is quite precise, it is wanting somewhat in accuracy.

The probability map derived from this model is shown in Fig 10 and the combinations of geology and soil types forming the zones of high and medium probability are summarised in table 11 below.

Despite the weaknesses of this model it does best express the influence of all the geomorphological factors affecting the formation of cropmarks. This model was therefore taken forward and combined with the model for the pattern of aerial reconnaissance.

Cropmark/soils and bedrock geology model. High probability zone				
Bedrock description	**Soil type**	**Cropmarks**	**Kj Max**	**Kvamme's gain**
Sandstone and [subequal/subordinate] argillaceous rocks, interbedded	Denbigh 2	152	0.0662	0.5869
Slate and siltstone	Denbigh 2	144	0.1183	0.4952
Granite	Moretonhampstead	141	0.1571	0.4003
Mudstone and sandstone	Denbigh 2	116	0.1965	0.3913

Cropmark/soils and bedrock geology model. High probability zone				
Bedrock description	**Soil type**	**Cropmarks**	**Kj Max**	**Kvamme's gain**
Slate and sandstone, interbedded	Denbigh 2	81	0.2256	0.3923
Slate and siltstone	Powys	71	0.2582	0.4154
Mudstone	Powys	65	0.2928	0.4478
Mudstone, siltstone and sandstone	Powys	50	0.3179	0.4655
Slate, siltstone and sandstone	Denbigh 2	47	0.3261	0.4383
Hornblende schist	Trusham	34	0.3425	0.4477
Hornfelsed slate and hornfelsed siltstone	Denbigh 2	44	0.3502	0.4255
Totals		**945**	**0.3502**	**0.4255**
Cropmark/soils and bedrock geology model. Medium probability zone				
Slate, siltstone and sandstone	Denbigh 1	61	0.3449	0.3641
Slate and siltstone	Denbigh 1	53	0.3425	0.3241
Sandstone	Neath	42	0.3490	0.3113
Mudstone and siltstone	Hallsworth 1	39	0.3593	0.3078
Slate and siltstone	Trusham	29	0.3707	0.3115
Sandstone and [subequal/subordinate] argillaceous rocks, interbedded	Denbigh 1	28	0.3730	0.3008
Slaty mudstone with sedimentary rock, metamorphic rock and igneous rock clasts	Denbigh 2	23	0.3839	0.3067
Slate	Denbigh 1	30	0.3793	0.2853
Sandstone and [subequal/subordinate] argillaceous rocks, interbedded	Powys	20	0.3803	0.2777
Peridotite and serpentinite	Croft Pascoe	17	0.3884	0.2822
Microgabbro	Trusham	15	0.3969	0.2878
Slate and siltstone	Sportsmans	15	0.4048	0.2927
Mudstone	Denbigh 2	15	0.4121	0.2965
Mudstone and siltstone	Neath	19	0.4053	0.2788
Slate	Denbigh 2	17	0.3992	0.2638
Granite	Moor gate	16	0.3937	0.2506
Mudstone, siltstone and sandstone	Yeollandpark	12	0.4010	0.2556
Totals		**451**	**0.4010**	**-0.1004**

Table 11. The high and medium probability zones of the cropmarks/soils and bedrock geology model.

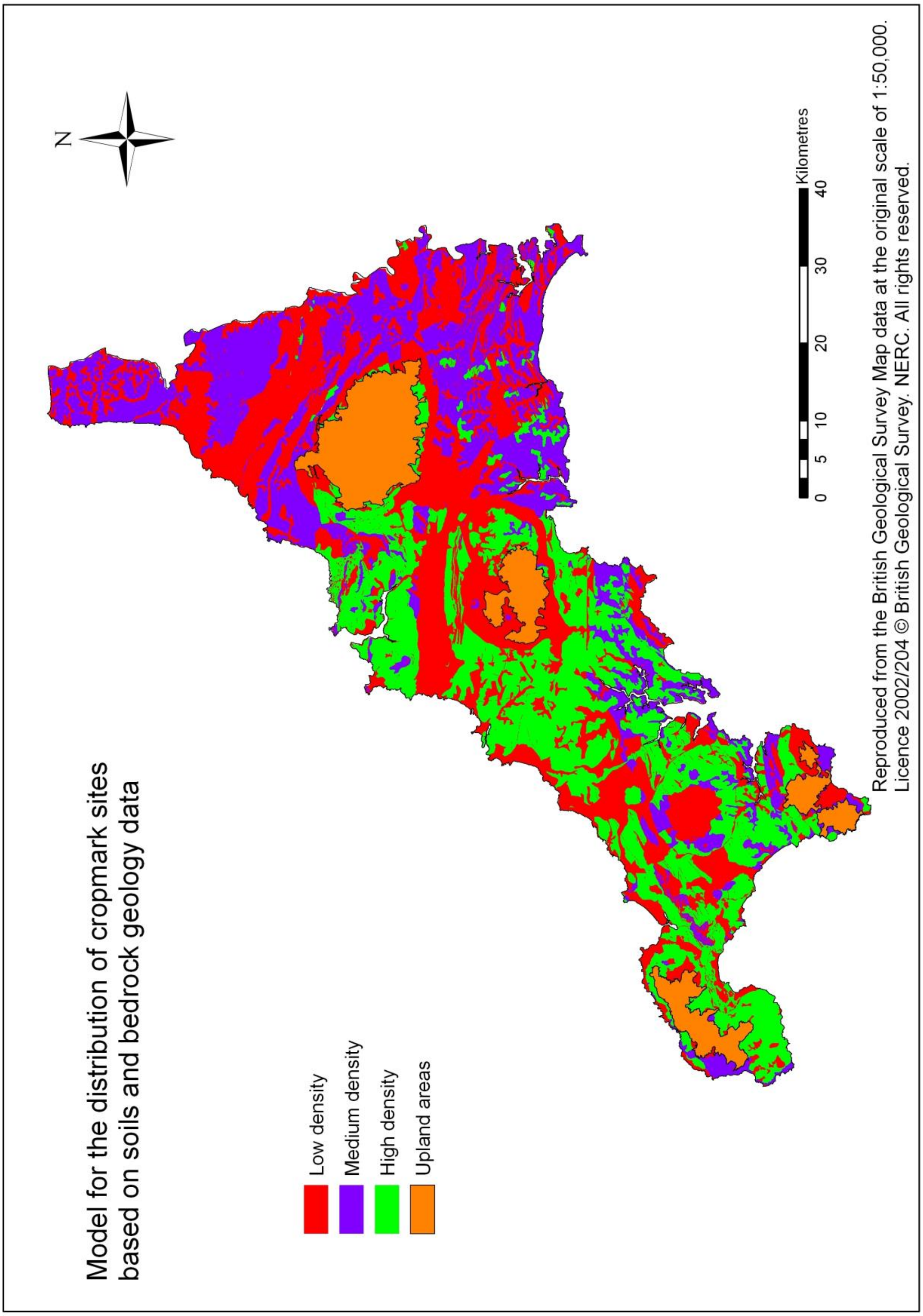

Fig 9. Model for the distribution of cropmark sites based on soils and Bedrock geology.

6 The pattern of aerial reconnaissance

Cropmarks of prehistoric and Romano-British sites in Cornwall have been identified and recorded from a range of aerial photographs, including both oblique and vertical images. Vertical photographs are not normally taken for archaeological purposes and those of Cornwall come from a variety of sources, most notably the RAF, Ordnance Survey (OS) and a Census flight commissioned by Cornwall County Council in 1995. The pattern of vertical reconnaissance can be regarded as even across the project area (the RAF and Census coverage is county-wide and OS coverage nearly so). Deviations in the reconnaissance pattern therefore result from differential levels of oblique photography resulting from specialist flying carried out by archaeological organisations. A limited amount of reconnaissance has been flown by the Cambridge University Committee for Aerial Photography (CUCAP) and by the NMR (notably in the early 1980s). These, however, amount to only a handful of flights and the bulk of specialist aerial reconnaissance in Cornwall has been undertaken by Historic Environment, Cornwall Council between 1985 and the present time. At the outset of this project HE reconnaissance had produced a total of 10,825 photographs (a mixture of black and white prints and colour digital images) from 87 different flights. Tracing the variations in levels of aerial reconnaissance therefore focused on the patterns of HE flying.

6.1 Extracting flight data

The first stage was to identify those flights during which cropmarks had been photographed. This was done by querying the Photo tables in the HER database to extract all site records where the Form = 'cropmark' and photo source = 'CAU'. The Flight number and date of flight for these site records were then tabulated to produce a list of all flights during which cropmarks were visible.

Flight no	Date	Cropmarks
F1	02/07/1985	14
F2	21/08/1985	1
F3	06/09/1985	1
F7	19/06/1986	1
F8	16/08/1986	1
F11	22/04/1987	4
F12	08/05/1987	1
F14	09/09/1987	4
F15	06/05/1988	2
F18	19/06/1989	9
F19	20/06/1989	4
F20	29/06/1989	18
F21	03/07/1989	20
F22	14/07/1989	4
F23	16/07/1989	22
F24	18/07/1989	53
F25	19/07/1989	29
F26	21/07/1989	15
F27	08/08/1989	4
F28	19/06/1990	1

Flight no	Date	Cropmarks
F33	24/05/1991	1
F34	26/05/1992	3
F36	26/06/1992	4
F38	17/07/1992	4
F39	24/07/1992	6
F40	05/05/1993	2
F44	08/06/1995	24
F45	25/07/1995	11
F47	17/08/1995	9
F48	17/07/1996	1
F49	25/07/1996	3
F50	26/07/1996	10
F51	07/04/1997	2
F55	01/08/2001	6
F62	14/08/2003	1
F65	23/07/2004	23
F70	11/07/2005	1
F73	14/07/2006	6
F74	25/07/2006	2
F80	11/09/2007	1

Flight no	Date	Cropmarks	Flight no	Date	Cropmarks
F29	02/07/1990	1	F82	16/04/2008	1
F30	20/07/1990	7	F84	22/07/2008	1
F32	27/11/1990	1	F85	23/07/2008	4
Total number of cropmarks					**343**

Table 12. Flights carried out by Historic Environment, Cornwall Council during which cropmarks were recorded.

A total of 343 cropmark features were recorded during 46 HE flights. Obviously conditions were more favourable in some years than others and from flight to flight. Annual totals are summarised below.

Year	No. of flights	Cropmarks
1985	3	16
1986	2	2
1987	3	9
1988	1	2
1989	10	178
1990	4	10
1991	1	1
1992	4	17
1993	1	2
1995	3	44
1996	3	14
1997	1	2
2001	1	6
2003	1	1
2004	1	23
2005	1	1
2006	2	8
2007	1	1
2008	3	6
Total	**46**	**343**

Table 13. The number of flights in which cropmarks were recorded from each year of EH aerial reconnaissance and the number of cropmarks recorded per flight.

Clearly 1989 was an outstanding year for cropmark prospecting in Cornwall, but 1995 and 2004 also produced above average results. Unsurprisingly July is the most productive month; 21 of the successful flights were carried out in July. Overall, apart from a single (anomalous?) flight in November 1990, the date range of the flights falls between 7th April and 11th September.

There are two weaknesses in the data used for this exercise. Firstly site entry into the HE Photo database appears to be inconsistent for some flights. The average number of photos resulting from each flight is 132, but for some this figure is much lower – only 14 photographs are recorded from F1, 55 from F2, 46 from F26, 25 from F31 and 41 from F27, etc. This suggests that not all the photographs from these flights have been recorded in the database and this in turn means that flight paths derived from the photo data (see below) will be incomplete. The second weakness is that the majority of cropmark sites were input to the HER as part of Cornwall's NMP project and in some cases only the photograph from which the NMP transcription was made was recorded in the database entry. If, for example, a site was visible on a HE photo and a CUCAP photo and the transcription was made from the CUCAP photo then the HE photo would not be cited in the HER record. As a result the number of cropmarks appearing on HE photography is likely to be understated and it is possible that some flights during which cropmarks were photographed have consequently been omitted from the list above.

6.2 Plotting flight paths

The second stage in defining the reconnaissance pattern involved plotting the flight paths of the flights in which cropmark features had been recorded. Although flight plans for all HE flights exist, their format in the case of the earlier flights (pen-drawn lines on OS base maps reduced to A4 size) precluded the creation of accurate digital versions. Instead the technique used was to plot all photographs taken during the relevant flights as point data (whether the photos were of cropmarks or not) and from the resulting plots to reconstruct the approximate route of each flight.

Point data for all photographs from all the flights listed above are shown in Fig 10 and two selected flight paths in Fig 11.

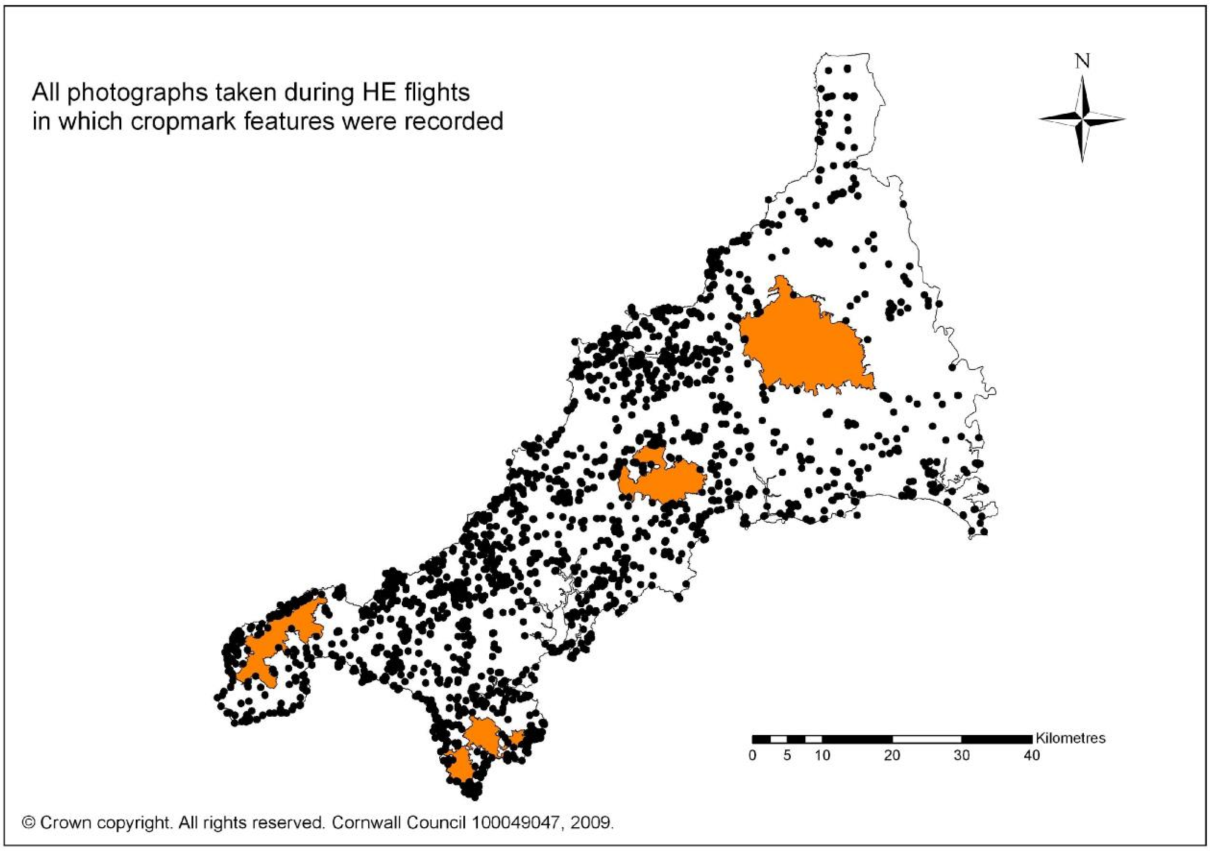

Fig 10. The location of all photographs taken during HE flights in which cropmarks were recorded.

The map shown in Fig 11 illustrates two weaknesses of the technique used for mapping the flight paths. Firstly, whilst the route of Flight 24 can be fairly confidently reconstructed from the pattern of photography, Flight 64 is focused on two separate areas. It is unclear from the pattern of photography whether these two areas were deliberately targeted and the land in between was simply flown over, or whether the land in between was subjected to reconnaissance but no sites were observed. A third possibility is that the flight moved from one area to the other by flying over the sea. So we don't know whether the land in between was overflown.

The second weakness lies in attempting to make reliable estimates for the visible area covered by each flight. Each dot represents a single site in the landscape and the field of vision from the aeroplane might reasonably be estimated as a 1km radius from each site. However there is no way of knowing whether the flight proceeded directly from site to site or by a more circuitous route. This means that in joining up the dots to create a flight path it is difficult to gauge the field of vision for the whole flight. For instance in the Camel Estuary area (Flight 65, Fig 11), we don't know whether the flight proceeded from one site to another or whether the whole area was circled again and again. In general the rule of thumb adopted was that if the photo points follow a recognisable string the flight path was plotted as a fairly well-defined linear polygon (Flight 24), but if there is a cluster of photo points it was assumed that a wide area has been overflown (Flight 65). In following this policy it is likely that the area covered by HE reconnaissance has been somewhat overstated.

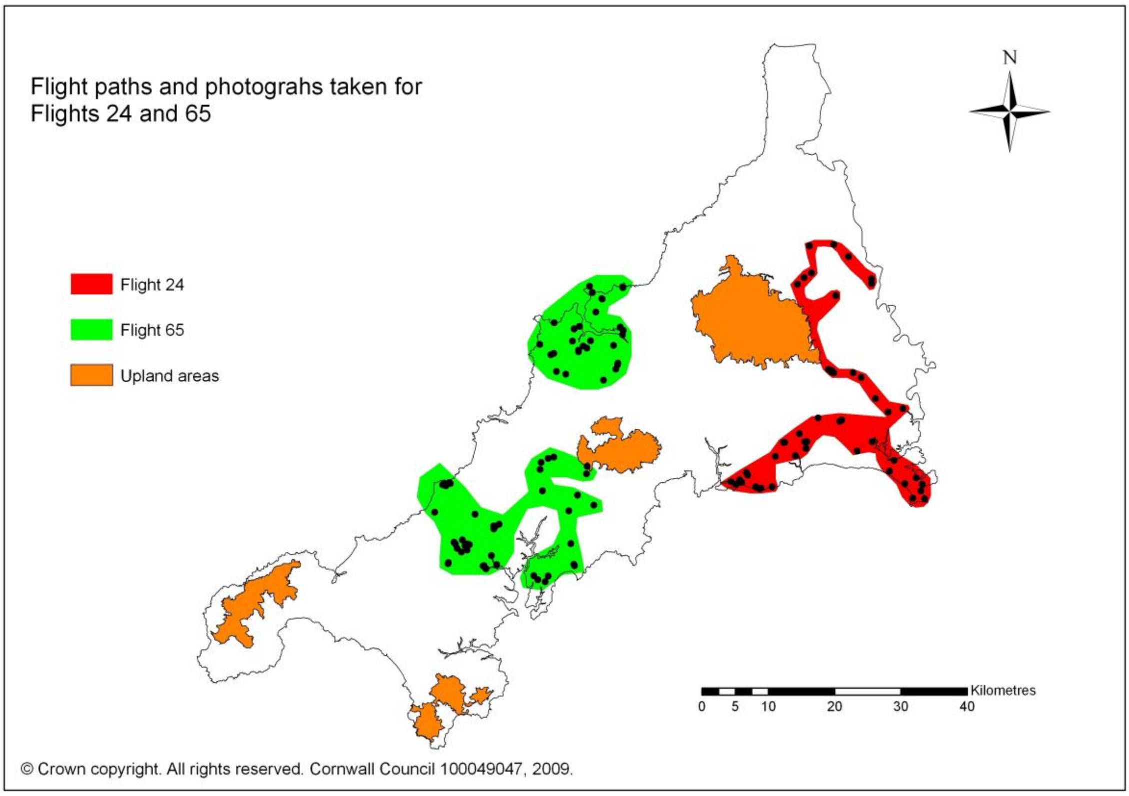

Fig 11 Example of flight path mapping. Flights 24 and 65, showing the location of photographs taken and the approximate flight paths.

Having defined an approximate route for each of the 46 flights the resulting flight paths were then intersected with one another in GIS and new polygons created for the zones of intersection. The polygons were categorised by the number of flight paths overlapping at any one location. Where two flights overlapped the polygon was given an attribute of 'Overlap count 2'; where four flights overlapped the polygon was given the attribute 'Overlap count 4' and so forth.

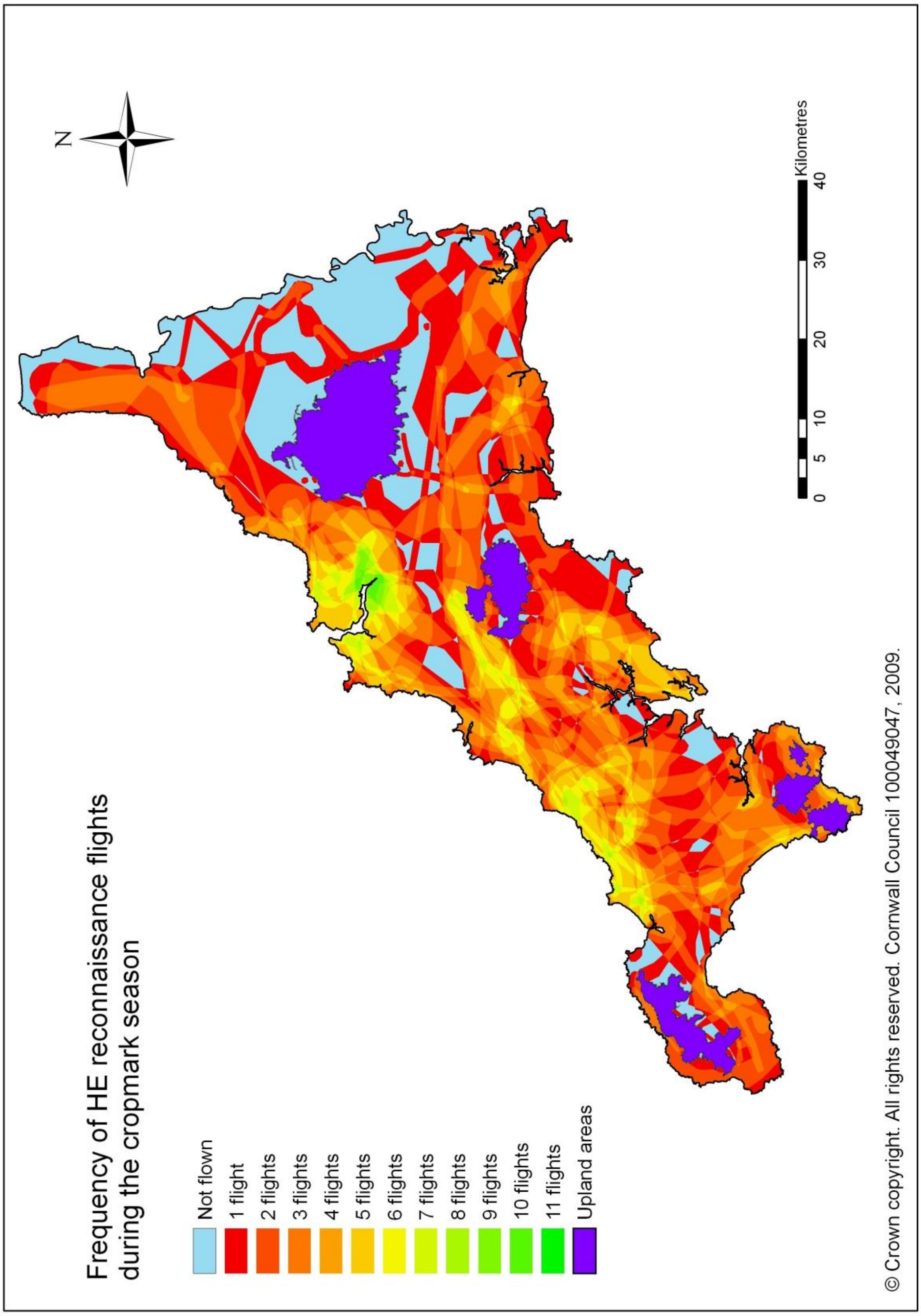

Fig 12 Frequency of HE cropmark reconnaissance showing the incidence of overlap between flights during which cropmark sites were observed

The overlap count ranged from one (with only a single flight) to 11 (the polygon had been overflown 11 times). In addition parts of the project area have never been flown during the prime cropmark periods (Fig 12). To produce a model based on the differential intensity of reconnaissance flights the overlap count was simplified into three categories representing well-flown, medium-flown and rarely-flown parts of the project area. The categorisation was defined by considering not just the overlap count but also the proportion of the project area taken up by each overlap count.

Overlap count	Area km^2	PA	Cumulative PA
0	602.74	18.87%	18.87%
1	691.57	21.65%	40.53%
2	666.73	20.88%	61.40%
3	563.37	17.64%	79.04%
4	294.25	9.21%	88.26%
5	212.57	6.66%	94.91%
6	100.43	3.14%	98.06%
7	38.59	1.21%	99.27%
8	10.64	0.33%	99.60%
9	8.24	0.26%	99.86%
10	4.53	0.14%	99.99%
11	0.02	0.01%	100.00%

Table 14. Summary of the overlap count for HE reconnaissance flights.

In table 14 the left hand column shows the number of times a given location has been overflown; the second column shows the size of the location that has been overflown that many times; the PA column shows the percentage of the project area taken up by that location and the final column shows the cumulative percentage of the project area taken up by those locations. For example, where the overlap count = 4, the table shows that an area of 294.25 square kilometres has been overflown four times during the cropmark season, that this area covers 9.21% of lowland Cornwall and that an area covering 88.26% of lowland Cornwall has been overflown four times or fewer.

Very clearly the areas which have been flown many times are much smaller than those where only a few flights have taken place. In view of this the categorisation was defined as follows.

Category	Overlap count
Well flown	4-11
Medium	2 & 3
Rarely flown	0 & 1

As a means of testing this categorisation the actual distribution of all cropmark sites was correlated with the three categories with the following result.

Category	PA	PS	Kvamme's gain	PS/PA
Well flown	0.2096	0.4269	0.5092	2.04
Medium	0.3852	0.3343	-0.1522	0.87
Rarely flown	0.4053	0.2388	-0.6973	0.59

The model is quite precise in that the well-flown zone covers only 20% of the project area but almost 43% of the known cropmark sites are captured within it. As a model it can be said to perform reasonably well, judging by the Kvamme's gain of 0.5092, although it does lack accuracy. The Indicative Values (PS/PA) for each category show that cropmarks are more than three times more likely to be encountered in the well flown area than in the rarely flown area.

The resulting probability map based on reconnaissance history is shown in Fig 13.

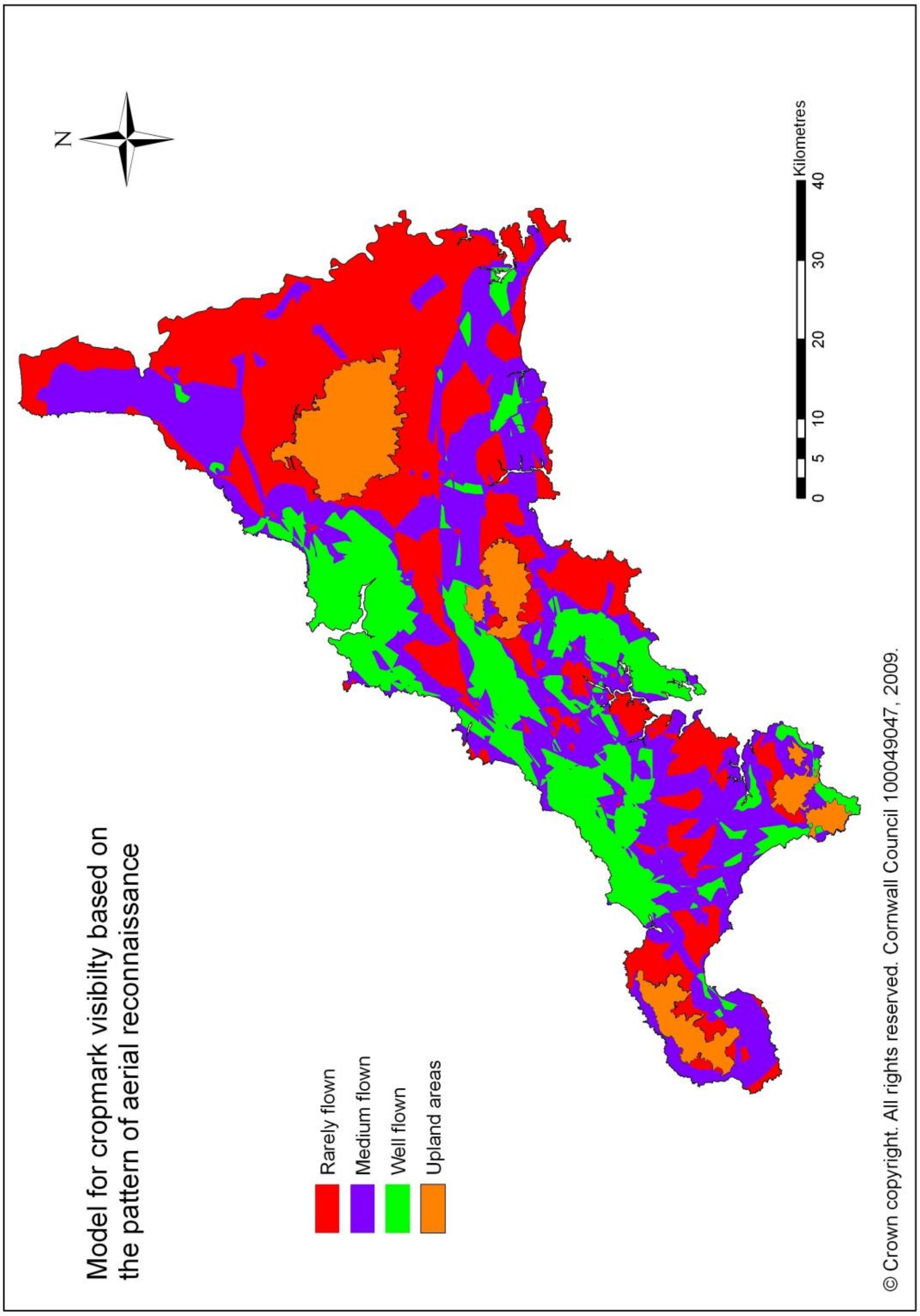

Fig 13. Cropmark visibility model based on patterns of aerial reconnaissance.

This map very clearly shows those locations where cropmarks are most likely to have been recorded – parts of the north coast (most notably the area around the Camel estuary), an east-west band in central Cornwall, parts of the Lizard peninsula and the

Roseland peninsula. On a broader level, parts of east Cornwall have been rarely flown during the cropmark season (see also Fig 12, which shows that large parts of east Cornwall have not been flown during the cropmark season). The pattern of aerial reconnaissance is similar in this respect to the models for cropmarks in relation to geology, soils and soils/geology, with east Cornwall depicted as a general area of low probability.

7 The final cropmark visibility map

The soils and geology model shows those areas of lowland Cornwall where cropmarks are most likely to form; the aerial reconnaissance model shows those areas where cropmarks are most likely to have been seen. In order to create a definitive cropmark visibility map, the two were combined.

To do this the zones of high probability in the geology/soils and aerial reconnaissance models were coded as H, the medium probability zones as M and the low probability zones as L. An additional code of N was included for those areas never flown during the cropmark season. To produce a definitive visibility model, both the geology/soils and the aerial reconnaissance models were then combined by means of a spatial union in GIS. The combinations of codes were used to define the zones of high, medium and low probability in the resulting model. Details of this model are shown in the table below. The first letter of the code combinations is derived from the geology/soils model and the second from the aerial reconnaissance model (e.g. the combination HL represents a combination of the high probability zone from the geology/soils model and the low probability zone from the reconnaissance model).

Visibility model. High probability zone					
Weighting	**Cropmarks**	**PA**	**PS**	**Kj Max**	**Kvamme's gain**
HH	547	0.105	0.311	0.2059	0.6621
MH	107	0.032	0.061	0.2351	0.6323
HL	121	0.052	0.069	0.2519	0.5717
HM	253	0.128	0.144	0.2681	0.4587
MM	175	0.101	0.099	0.2662	0.3893
Totals	**1203**	**0.418**	**0.684**	**0.2662**	**0.3893**
Visibility model. Medium probability zone					
LH	97	0.0723	0.0551	0.2491	0.3371
LM	160	0.1552	0.0910	0.1849	0.2228
MN	88	0.0747	0.0500	0.1602	0.1820
Totals	**345**	**0.3022**	**0.1961**	**0.1602**	**-0.5406**
Visibility model. Low probability zone					
ML	80	0.0738	0.0455	0.1319	0.1425
HN	25	0.0232	0.0142	0.1229	0.1308
LL	60	0.0904	0.0341	0.0665	0.0683
LN	46	0.0909	0.0262	0.0018	0.0018
Totals	**211**	**0.2783**	**0.12**	**0.0018**	**-1.3354**

Table 15. Results of the final cropmark visibility model

The performance of this model is summarised below.

Probability	PA	PS	Kvamme's gain	PS/PA
High	0.4177	0.6839	0.3893	1.64
Medium	0.3022	0.1961	-0.5406	0.65
Low	0.2801	0.1200	-1.3354	0.43

The model performs well, with 68% of the sites captured in a high probability zone covering 42% of the project area; with 19% captured in a medium probability zone covering 30% of the area and a low probability zone covering a similar sized area but containing only 12% of the cropmark sites. The Kvamme's gain of 0.3893 is modest, but better than some of the other models produced during this project.

Analysis of the model suggests that geology and soils are stronger predictors than the pattern of aerial reconnaissance (table 15). Of the five code combinations making up the high probability zone, three are derived from the high probability zone and two from the medium probability zone of the geology/soils model (HH, MH, HL, HM and HL). By contrast the five combinations comprise two from the well flown, two from the medium flown and one from the rarely flown zones of the aerial reconnaissance model.

An apparent contradiction to this pattern is the position in the low probability zone of the combination HN (the high probability zone of the geology/soils model combined with the never flown zone of the reconnaissance model). It should be pointed out, however, that this combination covers a very small area (2.3% of lowland Cornwall), but still contains 25 cropmark sites. This contrasts with, for example, the combination LM (the low probability zone of the geology/soils model combined with the medium flown zone of the reconnaissance model). This covers the largest area of any of the combinations (15.5%), but only contains 9% of the sites.

One important result of creating this model is the fact that 159 cropmark sites are recorded from locations which have not been flown by HE during the cropmark season. This may reflect inaccuracies arising from the weaknesses in the method used to recreate the pattern of reconnaissance described above in section 6.2. On the other hand it is more likely that these 159 cropmarks were identified from photographs not taken by HE, such as the various sources of vertical photographs. If this is the case it suggests that future programmes of HE reconnaissance in these areas would lead to the identification of more previously unrecorded cropmark sites.

The cropmark visibility map produced by this model (Fig 14) reinforces the trend (in line with the models discussed so far in this volume) for east Cornwall generally to be characterised as an area of low probability. However, by joining the soils/geology and reconnaissance models and then remodelling by correlating cropmark distribution with the combination probability zones a more nuanced version of this broad trend was achieved. For instance, in the soils/geology model (Fig 9) and in the reconnaissance model (Fig 13) the north eastern tip of the county, around Stratton and Bude, is classed as an area of either medium or low probability. In the cropmark visibility model a portion of this area is placed within the zone of high probability. In the same way, an extensive area close to the southeast coast (between Fowey and St Germans) is classed as a high probability zone in this model whereas in the other models much of the southeast coast is ranked in the low or medium probability zones.

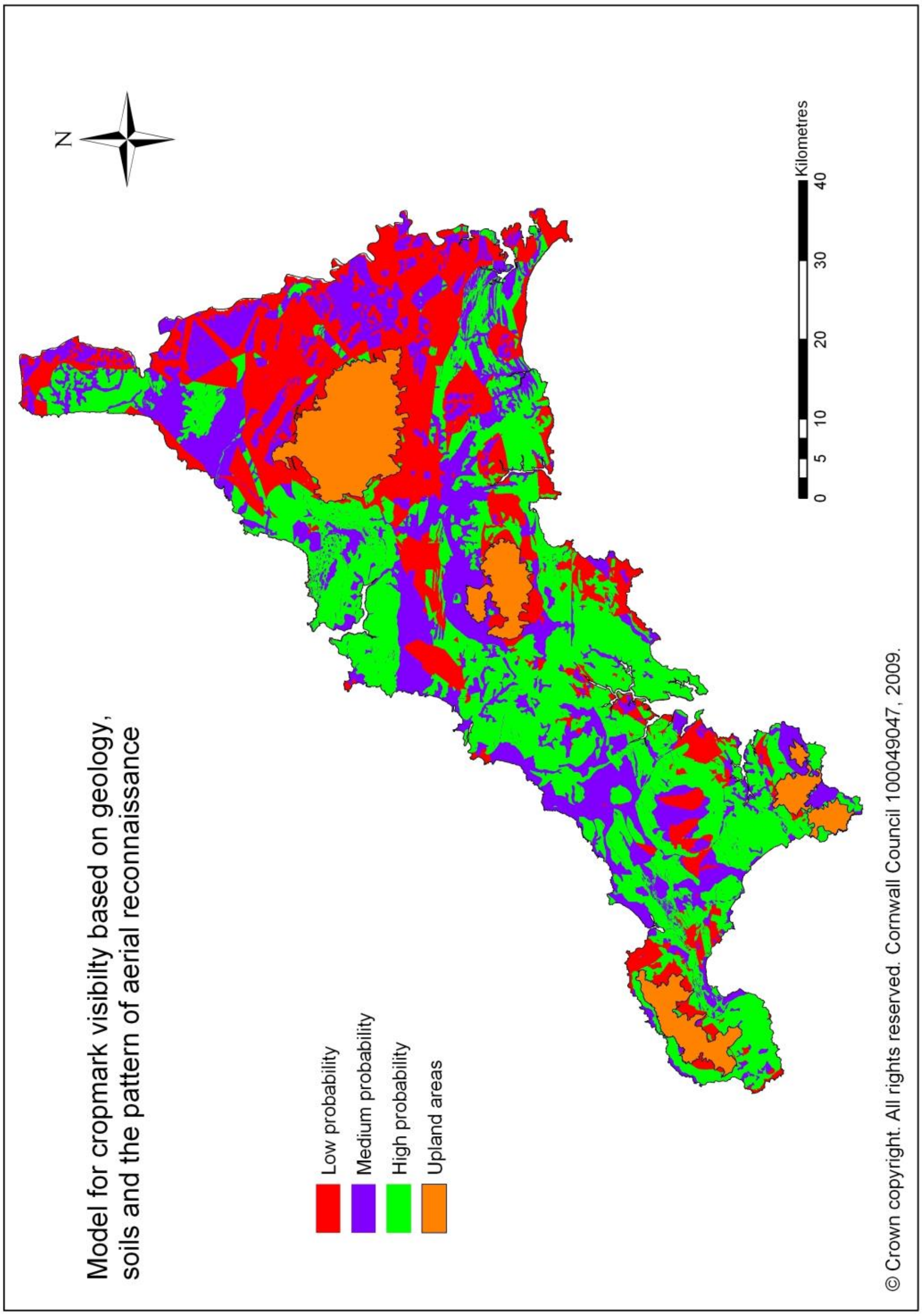

Fig 14. Model for the visibility of cropmarks based on geology, soils and aerial reconnaissance patterns

8 Testing the models

8.1 Methodology

As part of the project an analysis of the Cornwall events record was carried out with the aim of quantifying the extent and character of below-ground prehistoric and Romano-British archaeology identified by archaeological interventions, particularly those resulting from development-led work. Much of this archaeology is not recorded in the

HER and therefore provides a useful independent data sample with which to test the high level models developed during the project.

A full description of how the events dataset was filtered to include only those events relevant to the project is contained in Volume 1 of the Lowland Cornwall project report. The filtered dataset includes 424 events comprising the following event types.

Excavation (including test pits and trial trenches)	153
Watching briefs	121
Geophysical survey	141
Other (environmental, field walking, bore hole survey, etc)	9

The report or publication for each event record was studied to extract details of all the sites listed within it. New polygons for each event were created using the field boundaries marked on current OS maps as the polygon boundaries. The reasoning behind this is that although an event may examine only a portion of a field (for instance a single trench in a field) material found during the event is indicative of archaeological activity likely to extend beyond the confines of the immediate area examined. If, for example, pottery sherds were found in one trench, it is likely that further trenches dug elsewhere in the same field would uncover more sherds.

The events were then categorised in order to distinguish at a general level between the types of the recorded remains. These categories were defined as follows:

1. No features or finds in an area where archaeological levels were reached and no later disturbance had occurred.
2. Unstratified finds.
3. Discrete archaeological features with no dating evidence (such as gullies, pits and apparently random post holes, which are potentially prehistoric).
4. Discrete archaeological features with dating evidence.
5. Coherent arrangements of structural features with dating evidence and site plan.
6. A site where palaeoenvironmental evidence has been found.

In practice although some polygons created during the analysis represent a single field, many enclose a number of adjoining fields where archaeological material had been recorded. It will be appreciated that the analysis of events produced many more polygons than the number of actual events. The route of a pipeline (a single event) might cut through 100 fields and be represented by 30 polygons which may fall into several of the categories outlined above. Furthermore some fields might contain more than one category of event. An excavation of an Iron Age enclosure may also reveal Neolithic pottery and a Bronze Age pit: in this instance three sets of attributes would be created for the field in question (it would effectively be represented by three separate polygons). In total 833 polygons were created for the 424 individual events making up the Lowland Cornwall dataset.

One issue to be addressed when using the events data as a test sample is that in a number of cases more than one event has taken place at the same location. The most frequent occurrence of this is where a geophysical survey is followed by a watching brief and/or excavation. At Tremough, Penryn, for instance a total of nine individual events were carried out over a period of several years at this extensive multi-period settlement site (Fig 15).

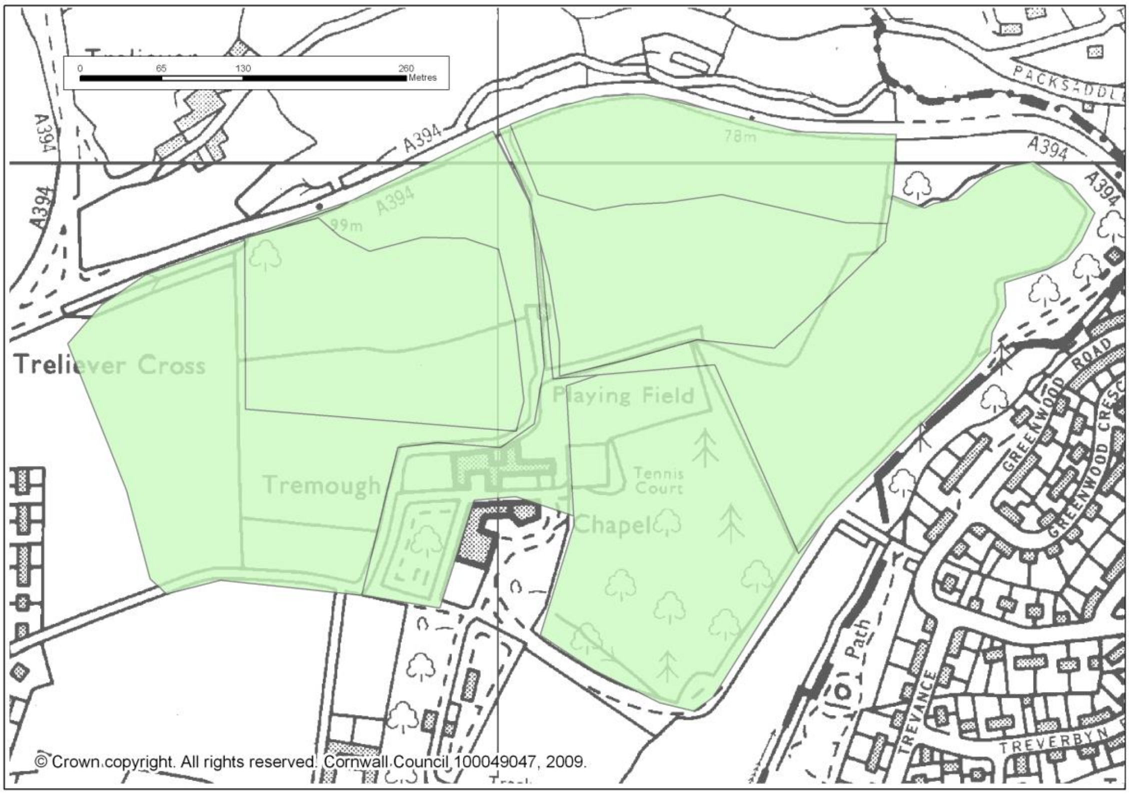

Fig 15. A series of overlapping polygons defining numerous events carried out at Tremough, Penryn.

The polygons defining these events overlap and overlay each other and this makes it difficult to calculate the extent of the site. The sum of the area covered by all nine polygons in the case of Tremough is 44.5ha but, because of the degree of overlap, this is much greater than the actual area surveyed. Therefore to create a more accurate events dataset for testing purposes, all overlapping or overlying polygons were deleted to produce a simplified dataset. In the case of Tremough, the nine polygons were reduced to four which accurately encompass the extent of the site and whose sum comes to 26ha. As a result of this process the events test dataset contains 702 polygons and covers a total area of 54.36km^2. Centroid points were created for each of the polygons so that the data can be analysed by numbers of sites (points) or by site area (polygons).

A spatial union was then performed in GIS linking the lowland event polygons with the visibility model. In many cases the events polygons intersect more than one visibility model polygon; as a result the events/visibility layer comprises 2,750 individual polygons. The layer can be interrogated on the basis of event category, site type, period, geology type, soil type, flight overlap count, weighting, or any combination of these attributes.

The cropmark visibility model is based on the known distribution of cropmark sites and, therefore, only events falling into categories 3, 4 and 5 (see above) were included in the test sample (these categories indicate the existence of structural features which might be expected to be visible as cropmarks). The test sample was reduced further by including only previously unrecorded sites discovered by the events. This was to ensure that only sites with no surviving above-ground remains were considered. In total 230 such sites are recorded in the events record dataset. In the events/visibility layer these sites are represented by 629 polygons covering a total area of approximately 12.25km^2.

One problem faced when using the events record data as a test sample is that the area surveyed by the events is not representative of the Lowland Cornwall project area as a whole. The proportion of the events record survey area made up by each probability zone of the visibility model differs from that within the overall project area. For example the low probability zone makes up approximately 18% of the events record survey area but covers 28% of the overall Lowland Cornwall project area (table 16).

Probability zone	% of project area	% of events area
High	42%	40%
Medium	30%	42%
Low	28%	18%

Table 16. Comparison of the proportion of area taken up by each probability zone in the overall project area and the area covered by events.

Whilst the proportions of the high probability zone in the Lowland Cornwall project area and the area covered by archaeological events are roughly equal, the proportion of the medium probability zone covered by events is significantly greater than the proportion of Lowland Cornwall made up by that zone, and the proportion of the low probability zone covered by events is significantly smaller. To use the events record data to test the visibility model, compensation for these variances must be made otherwise it is likely that more sites than predicted would be captured in the medium probability zone and fewer than predicted in the low probability zone.

The simplest way to compensate for this variance is to calculate the S/A value - the number of sites per km^2 - for each of the model's probability zones. The area (in km^2) making up each zone in the test survey area is then multiplied by the S/A value from the original model to arrive at a notional predicted number of sites for each zone. From these notional figures the predicted PS value for each zone of the test survey area can be defined. This is illustrated below using the visibility model as an example; the S/A values are calculated from the model as follows:

Probability	AREA	SITES	S/A
High	1333.96	1203	0.90
Medium	965.04	345	0.36
Low	894.68	211	0.24
Total	3193.69	1759	0.55

So, for the high probability zone in the test sample area the expected density of sites is 0.9 per km^2 and, given that this zone covers 21.6km^2, we can predict the theoretical number of sites to be captured in this zone is 21.6 x 0.9 = 19.45. In total we can expect, in theory, 29.9 sites to be recorded in the test sample and the proportion of these falling within the high probability zone will be 19.45/29.9 = 0.65 or 65% (table 17 below).

Zone	Area km^2	S/A	Notional sites	Predicted PS
High probability	21.6135	0.90	19.45	0.65
Medium probability	22.8321	0.36	8.22	0.27
Low probability	9.9156	0.24	2.38	0.08
Total	54.3612	0.55	29.90	1.00

Table 17. Testing the visibility model: predicted proportion of sites for each probability zone in the events record survey area.

8.2 Test results

8.2.1 The cropmark visibility model

In testing the cropmark visibility model with events record data, only previously unrecorded sites newly discovered by the events were included in the test sample. This was to ensure that only sites with no surviving above-ground remains were considered. The test sample was further filtered by removing early medieval and medieval sites, leaving only those that were prehistoric, Romano-British, or of unknown date (and therefore potentially prehistoric). In total 230 such sites are recorded in the events record dataset. In the events/visibility layer these sites are represented by 629 polygons covering a total area of approximately 12.25km^2. The distribution of the sites is shown in Fig 16 below.

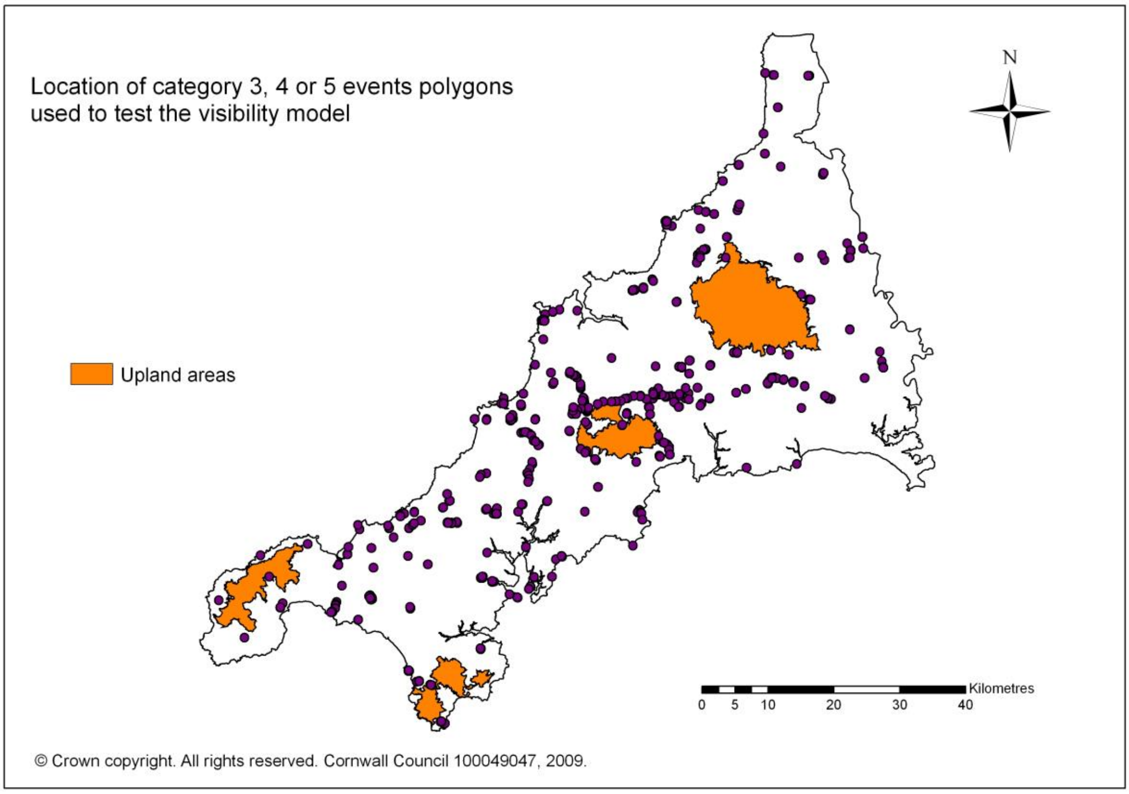

Fig 16. The location of events record polygons containing category 3, 4 or 5 features used to test the cropmark visibility model.

The most striking result of testing with events record data is the degree to which the cropmarks dataset fails to indicate the likely extent of below-ground archaeology in lowland Cornwall. Whilst the cropmark visibility model suggests that we might expect 30 sites to be recorded in the area surveyed by the events, the actual figure is almost eight times greater than this. The distribution of the sites, whilst not even, reflects a good spread of sites across the whole project area (Fig 16). The results of the test are shown in tables 18 and 19 below.

Zone	Area km²	SA	NS	Predicted PS	PS	Predicted Sites	Sites
High probability	21.6135	0.90	19.45	0.65	0.45	150	103
Medium probability	22.8321	0.36	8.22	0.27	0.39	62	90
Low probability	9.9156	0.24	2.38	0.08	0.16	18	37
Totals	54.3612	0.55	29.90			230	230

Table 18. Results of events record testing of the cropmark visibility model: test based on numbers of sites. NS = theoretical number of sites predicted.

Zone	Area km²	Predicted PS	PS	Predicted site area	Site area
High probability	21.6135	0.65	0.39	7.97	4.75
Medium probability	22.8321	0.27	0.44	3.32	5.43
Low probability	9.9156	0.08	0.17	0.96	2.07
Totals	54.3612			12.25	12.25

Table 19. Results of events record testing of the cropmark visibility model: test based on area of events polygons.

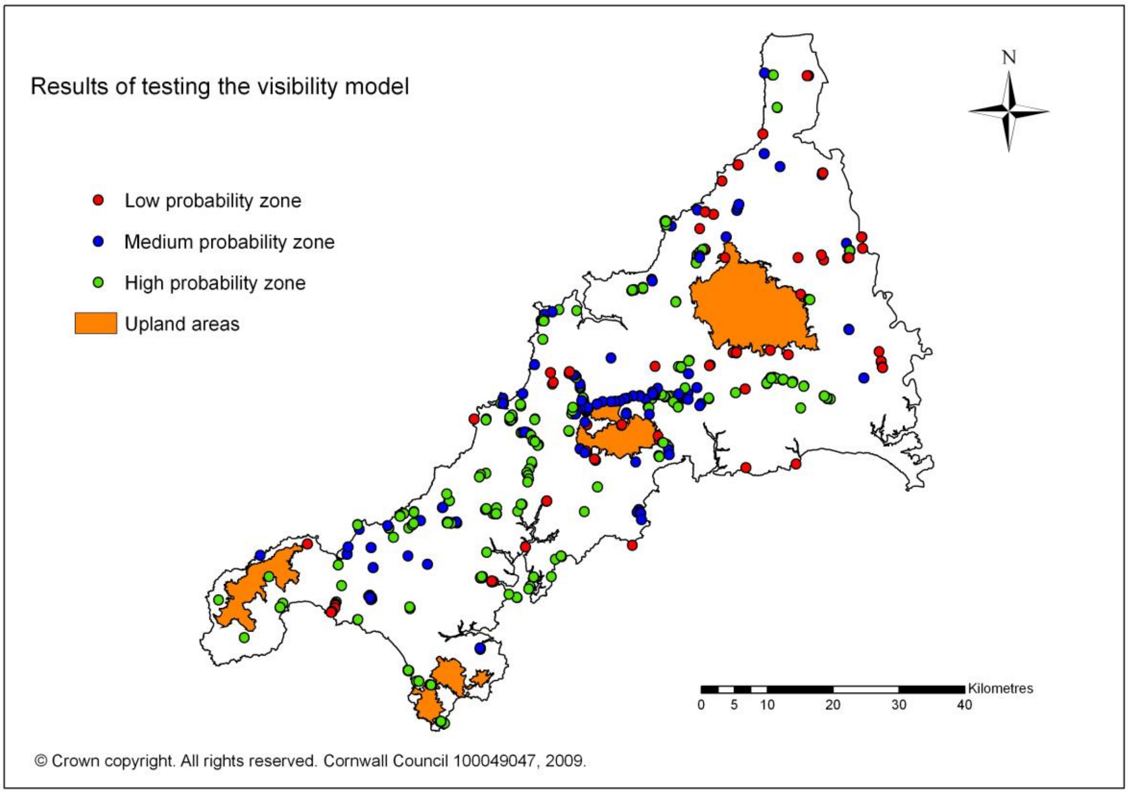

Fig 17. The location of events record polygons within the probability zones of the cropmark visibility model.

Fig 17 shows the distribution of the events used in the test as a series of points colour coded to indicate in which probability zone they are captured. It is clear from these results that, whether the test is based on site numbers or site area, the cropmark visibility model is rejected by the test sample. Both the medium and low probability zones perform better than expected, with the low probability zone in particular capturing more than twice as many sites as predicted and whose polygons cover more than twice the predicted area. The performance of the high probability zone is especially poor when the test is based on polygon area – with a PS value of only 0.39, as opposed to the predicted PS value of 0.65. Furthermore 44% of the area containing below-ground archaeology is captured in the medium probability zone – a greater area than is captured in the high probability zone.

8.2.2 The cropmarks and soils model

Bearing in mind that the cropmark visibility model is built from a range of variables, it is useful to examine whether any particular one of these is a stronger factor than the others in the failure of the model to accurately predict the location of below-ground archaeological features. For this reason each of the variables – soils, bedrock geology, soils and geology combined, and aerial reconnaissance patterns – were tested individually using the same events test sample. Results of testing the cropmarks and soils model are presented in tables 20 and 21 below. In table 20 S/A = sites per km^2 (according to the original model); NS = notional number of predicted sites; Predicted PS = the predicted proportions of sites captured in each probability zone. Sites = the actual number of sites captured, and Predicted Sites = the predicted number of sites from each probability zone, given the actual total number of sites recorded.

Zone	Area km^2	S/A	NS	Predicted PS	PS	Predicted Sites	Sites
High probability	28.2045	0.81	22.85	0.76	0.60	174	139
Medium probability	11.7566	0.40	4.70	0.15	0.20	36	47
Low probability	14.4002	0.18	2.59	0.09	0.19	20	44
Totals	54.3613		27.64			230	230

Table 20. Results of events record testing of the cropmarks and soils model: test based on numbers of sites.

Zone	Area km^2	Predicted PS	PS	Predicted site area	Site area
High probability	28.2045	0.76	0.57	9.29	6.94
Medium probability	11.7566	0.15	0.16	1.91	1.96
Low probability	14.4002	0.09	0.27	1.05	3.35
Totals	54.3613			12.25	12.25

Table 21. Results of events record testing of the cropmarks and soils model: test based on area of events polygons.

This test cannot be said to validate the model, especially when performance is measured by area. It does, however, perform slightly better than the cropmark visibility model – 16% fewer sites than predicted (a PS value of 0.60 as opposed to the predicted 0.76) are captured in the high probability zone: in the visibility model 20% fewer sites than expected are captured. When performance is measured by area the high probability zone of the soils model contains 19% less site area than predicted (in the cropmark visibility model this figure is 26%). A significant difference between the two models is that in the soils model the medium probability zone performs more or less as predicted or better (5% more sites as expected and a 1% larger area than predicted) so that the poor performance of the high probability zone is mainly a result of the low probability zone capturing more than twice as many sites as predicted. In the cropmark visibility model the shortfall of sites captured in the high probability zone is divided more equally between the medium and low probability zones. However, it should be noted that the proportion of area taken up by each probability zone in the two models is quite different. In the soils model the medium and low probability zones cover a broadly similar sized area (11.7 – 14.4km^2) and the high probability zone is twice as large: in the visibility model the high and medium probability zones are roughly the same size (21-22km^2) and the low probability zone is less than half this. So in actual fact the result in both cases is the much better than predicted performance of the low probability zone.

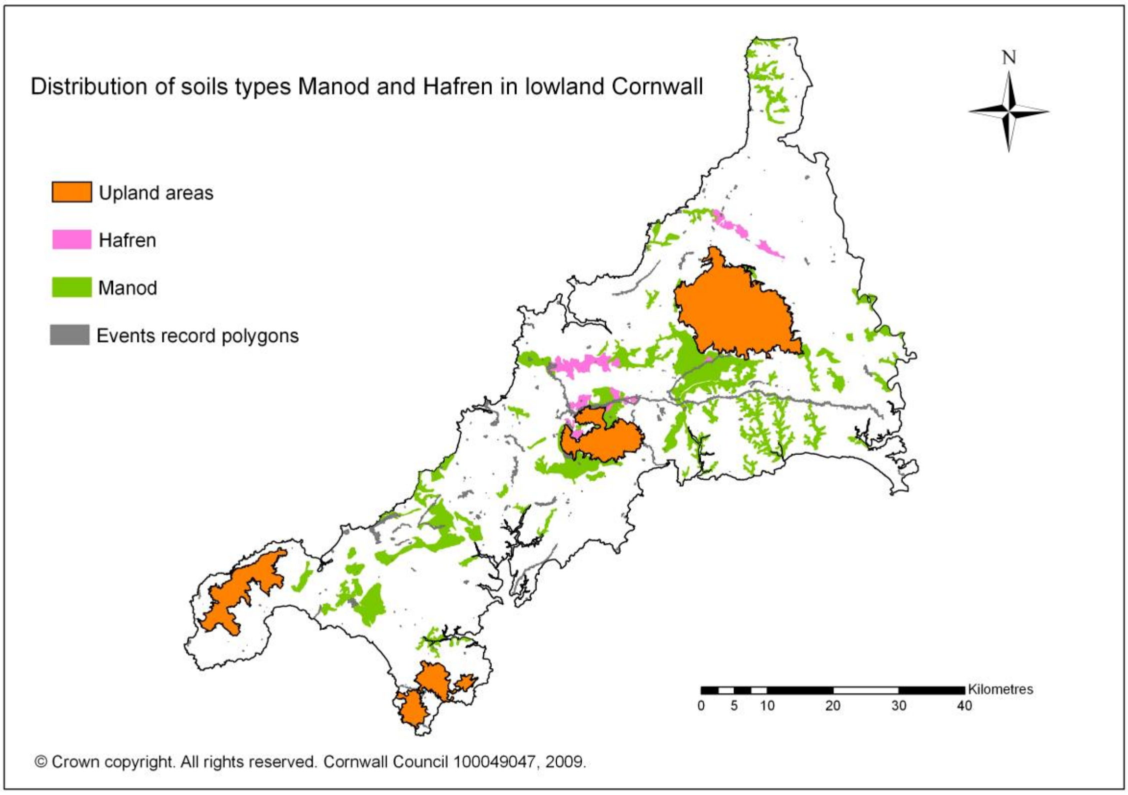

Fig 18. Map showing the distribution of Manod and Hafren soils in lowland Cornwall.

It is of interest to analyse further the soil types making up the low probability zone to see whether any light can be shed on its better than predicted performance. The two best performing soil types in the zone are Manod, which captures 19 event record sites and 1.8km^2 of site polygons, and Hafren, which captures 12 sites and 1.1km^2 of the polygons. So together, these two soils capture 70% of the sites in the low probability zone, and 87% of the total site area in this zone. The two soils are quite different: Hafren is described as 'peat to loam over shale, a loamy permeable upland soil with a wet peaty surface horizon and bleached subsurface horizon, often with a thin ironpan' it is typical of moorland and grassland habitats of moderate grazing value but is suitable for stock rearing and dairying on improved ground. In the cropmarks/soils model it is ranked thirteenth, it covers 1.4% of lowland Cornwall and contains only 0.85% of all recorded cropmarks. Sites recorded in the events record on Hafren soils include the metalworking enclosure at Little Quoit Farm. Manod soils are far more extensive, and are the third most widespread soil type in lowland Cornwall, covering 10.5% of its total area. They are typically well-drained fine loamy or fine silty soils, and support dairying and cereals in Cornwall. In the cropmarks/soils model they are ranked fourteenth. Despite being so widespread Manod soils only capture 3% of all recorded cropmarks in lowland Cornwall. Given that Manod soils are suitable for cereal growing and that they are so widespread, it is perhaps surprising that more cropmarks have not been recorded from them. Perhaps this is partly explained by the fact that they are frequently found in steep-sided river valleys (Fig 18).

8.2.3 The cropmarks and geology model

The results of testing the cropmarks and bedrock geology model are presented below in tables 22 and 23.

Zone	Area km²	S/A	NS	Predicted PS	PS	Predicted Sites	Sites
High probability	26.7496	0.75	20.03	0.69	0.51	160	118
Medium probability	9.976	0.42	4.22	0.15	0.15	34	34
Low probability	17.6356	0.26	4.58	0.16	0.34	37	78
Totals	54.3612	0.55	28.83			230	230

Table 22. Results of events record testing of the cropmarks and bedrock geology model: test based on numbers of sites.

Zone	Area km²	Predicted PS	PS	Predicted site area km²	Site area km²
High probability	26.7496	0.69	0.48	8.51	5.86
Medium probability	9.976	0.15	0.13	1.79	1.61
Low probability	17.6356	0.16	0.39	1.95	4.78
Totals	54.3612			12.25	12.25

Table 23. Results of events record testing of the cropmarks and bedrock geology model: test based on area of events polygons.

Again testing suggests the model is not convincing; the high probability zone performs worse than that of the soils model, but better than that of the visibility model, with 18% fewer of the sites captured than predicted. The performance of the medium probability zone is as predicted but that of the low probability zone is far better than expected. This is the most striking aspect of the test results: the low probability zone captures more than twice the number of sites predicted, or an 18% larger share of the sites than expected. In this test the enhanced performance of the low probability zone is at the expense of the high probability zone.

This is even clearer when the testing is based on sites area – the low probability zone now covers 23% more of the survey area than predicted. In common with test results from the other models, the high probability zone performs particularly badly when the test is based on site area – in this case the events polygons containing sites cover an area 21% less than predicted in the high probability zone.

Further examination of the test results indicates that by far the most important rock type making up the low probability zone is Hornfelsed slate and Hornfelsed sandstone: 37 of the previously unrecorded sites in the events record are located on this geology, and the polygons containing these sites cover almost 3km^2, which equates to roughly three fifths of the site area in the low probability zone. Whilst many of the new features recorded from this geology are groups of undiagnostic ditches and pits, there are a number of category 4 or 5 events, some of which are extensive and important sites. These include the small double-ditched enclosure at Little Quoit Farm; enclosures, pits and field system (a complex of features dating from the Bronze Age to Romano-British period) at Trenowah and enclosures and field system at Scarcewater. In the HER, only 22 cropmark sites are recorded from this rock type, but its occurrence is localised and limited to the land surrounding the St Austell granite (Fig 19). This is an area where there have been several major events (e.g. St Austell North East Distributor road, Scarcewater mitigation, A30 Bodmin to Indian Queens road improvement) and some 8km^2 (14%) of the events polygons overlie Hornfelsed slate and Hornfelsed sandstone. This is in contrast to the 2.4% of lowland Cornwall taken up by this rock type, and this variance may help explain the better than predicted performance of the type in this model.

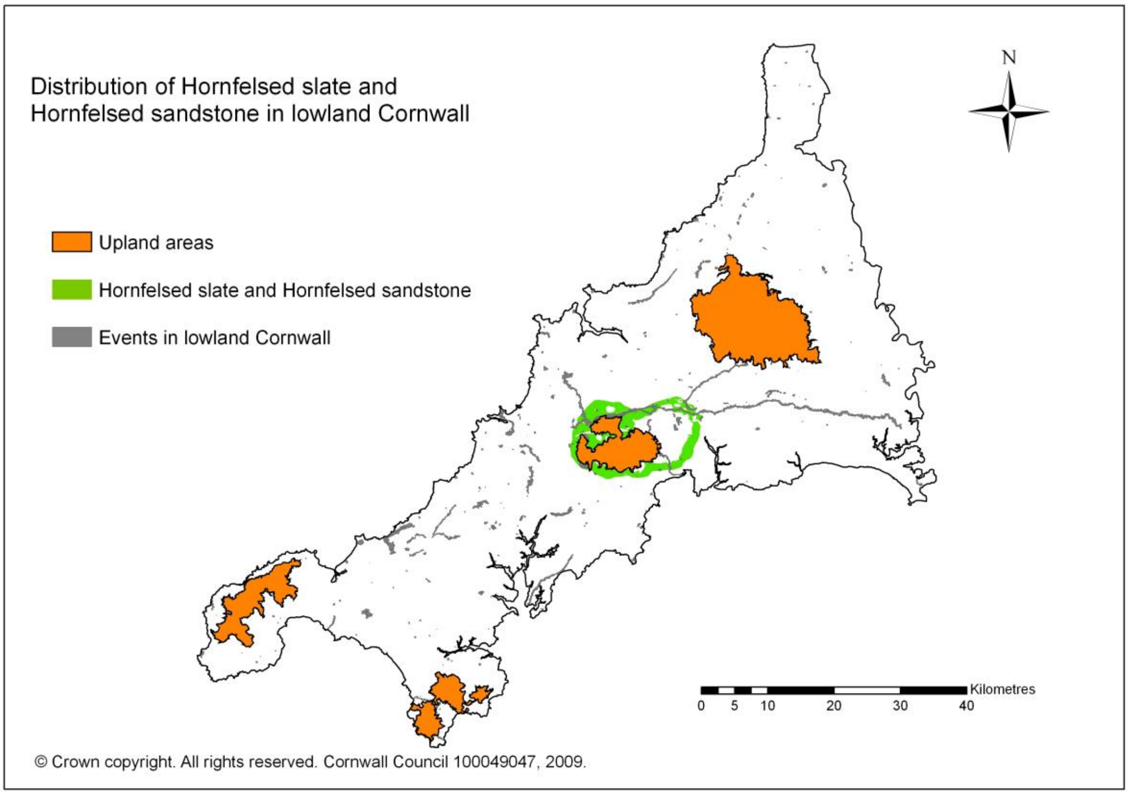

Fig 19.Map showing the localised occurrence of Hornfelsed slate and Hornfelsed sandstone in lowland Cornwall.

8.2.4 The cropmarks model based on soils and bedrock geology

The results of testing the cropmarks and soils/bedrock geology model are presented below in tables 24 and 25.

Zone	Area km^2	S/A	NS	Predicted PS	PS	Predicted Sites	Sites
High probability	16.417	0.96	15.77	0.56	0.31	128	72
Medium probability	8.4759	0.50	4.25	0.15	0.21	35	48
Low probability	29.4685	0.28	8.21	0.29	0.48	67	110
Totals	54.3614	0.55	28.24			230	230

Table 24. Results of events record testing of the cropmarks and soils/bedrock geology model: test based on numbers of sites.

Zone	Area km^2	Predicted PS	PS	Predicted site area	Site area
High probability	16.417	0.56	0.31	6.84	3.77
Medium probability	8.4759	0.15	0.12	1.84	1.47
Low probability	29.4685	0.29	0.57	3.56	7.01
Totals	54.3614			12.25	12.25

Table 25. Results of events record testing of the cropmarks and soils/bedrock geology model: test based on area of events polygons.

Of the three models based on geomorphological factors this, using a combination of soils and bedrock geology, is the most strongly rejected. The most significant factor is the performance of the low probability zone, which has captured 19% more sites than predicted and the area covered by these sites is 28% greater than predicted. This is similar to the cropmarks/geology model but in this test the low probability zone actually captures many more sites than the high probability zone. One mitigating factor may be that in the original cropmarks and soils/geology model (table 11, section 5.5) the high probability zone is defined precisely at the expense of accuracy – with 53% of sites within 30% of the project area.

8.2.5 The cropmarks model based on patterns of aerial reconnaissance

The results of testing the cropmarks and aerial reconnaissance model are presented below in tables 26 and 27.

Zone	Area km^2	S/A	NS	Predicted PS	PS	Predicted Sites	Sites
High probability	19.5217	1.12	21.86	0.60	0.36	137	82
Medium probability	22.4851	0.48	10.79	0.29	0.41	68	94
Low probability	12.3545	0.32	3.95	0.11	0.23	25	54
Totals	54.3613		36.61			230	230

Table 26. Results of events record testing of the cropmarks and aerial reconnaissance model: test based on numbers of sites.

Zone	Area km^2	Predicted PS	PS	Predicted site area	Site area
High probability	19.5217	0.60	0.36	7.31	4.36
Medium probability	22.4851	0.29	0.44	3.61	5.35
Low probability	12.3545	0.11	0.21	1.32	2.53
Totals	54.3613			12.25	12.25

Table 27. Results of events record testing of the cropmarks and aerial reconnaissance model: test based on area of events polygons.

In terms of the performance of the high probability zone this model is clearly rejected. From the point of view of predicted numbers of sites, 24% fewer than expected are captured in this zone and they cover a 24% smaller area than predicted. The results of this test also differ from the others in that the performance of the medium probability zone has been enhanced as well as that of the low probability zone. In this respect the performance of this model is most similar to that of the cropmark visibility model, suggesting that the pattern of aerial reconnaissance has a greater influence on the cropmark visibility model than soils, geology or geology/soils combined.

8.3 Discussion and further analysis

The most striking aspect of all the tests carried out on these models is the far higher number of below-ground prehistoric or Romano-British features recorded in the events record than might be expected. Using the known distribution of cropmarks as a guide, and based on numbers of cropmarks per square kilometre, the predicted numbers of sites for each of the models ranged from 28 to 37. The actual number of new sites recorded in the events record is 230 – between six and eight times more than predicted.

This might be partly explained by the fact that the majority of the 230 new sites are classed as category 3 events – comprising in the main undiagnostic groups of ditches and/or pits – and that many of the category 4 and 5 events consist of unsubstantial features, such as pits, gullies and post holes, which rarely form clearly visible cropmarks. In fact the cropmark dataset is dominated by rounds and enclosures, and to a lesser degree, the surrounding ditches of barrows; in other words features characterised by substantial ditches. Analysis of the events record test sample shows that of the 230 sites, 176 (three quarters of the total) are slight features that would be unlikely to be visible as cropmarks. On the other hand, the test sample contains some extensive sites which one might expect to have been identified through the HE flying programme or from aerial photographs from other sources. These include, for instance, the complexes at Pennance, Trenowah and St Stephen in Brannel School.

Also striking is the fact that all the models were rejected to a greater or lesser extent by the test sample. It appears that the model based on the pattern of aerial reconnaissance bears most influence on the performance of the cropmark visibility model on the grounds that its performance most closely resembles that of the visibility model. The medium probability zone of these two models performs significantly better than predicted, whereas in the models based on the correlation of cropmarks with soils, with geology and with soils and geology, the medium zone performs more or less as expected and it is invariably the low probability zone whose performance is enhanced at the expense of the high probability zone. This is not to say that the models for cropmarks/soils, cropmarks/geology or cropmarks/geology and soils can be accepted as reliable.

The obvious conclusion to draw from this is that the cropmark visibility model is exactly that – a model showing the favoured distribution of known cropmark sites – and cannot be taken as a reliable guide to the likely location of previously undiscovered below-ground archaeology.

Another way of looking at the tests is to analyse the distribution of the events records within the probability zones of the models rather than predicting how many records will be captured in each. Taking the cropmarks/aerial reconnaissance model, for instance, the high probability zone covers 36% of the project area and it captures 82 sites. This equates to 36% of the 230 sites in the events record: suggesting a by-chance distribution. The same is true of the medium and low probability zones as shown in table 28 below (the by-chance distribution is confirmed by a Kvamme's gain of 0).

Zone	**Area km^2**	**Sites**	**PA**	**PS**	**Kvamme's gain**
High probability	19.5217	82	0.36	0.36	0
Medium probability	22.4851	94	0.41	0.41	0
Low probability	12.3545	54	0.23	0.23	0
Totals	54.3613	230			

Table 28. Distribution of events record sites in the cropmarks/aerial reconnaissance model probability zones, based on site numbers.

The low probability zone captures 23% of sites in 23% of the area and the medium probability zone captures 41% of sites in 41% of the area. The by chance nature of this test model is almost as clear when it is based on sites area, as shown in table 29.

Zone	**Area km^2**	**Site area**	**PA**	**PS**	**Kvamme's gain**
High probability	19.5217	4.36	0.36	0.36	0
Medium probability	22.4851	5.35	0.41	0.44	0.0682
Low probability	12.3545	2.53	0.23	0.21	-0.0952
Totals	54.3613	12.25			

Table 29. Distribution of events record sites in the cropmarks/bedrock geology model probability zones, based on site area.

The test sample used against the cropmarks/bedrock geology model also has a close to a by chance distribution, with 51% of the sites captured in a high probability zone covering 49% of the area.

Similarly the high probability zone of the cropmarks/bedrock and soils model covers 30% of the area and captures 31% of the sites.

When analysing the distribution of the test sample sites, the cropmarks and soils model produces the best result, with its high probability zone covering 52% of the area and capturing 60% of the sites and both the medium and low probability zones producing negative Kvamme's gain measures.

Zone	Area km^2	Sites	PA	PS	Kvamme's gain
High probability	28.2045	139	0.52	0.60	0.1333
Medium probability	11.7566	47	0.22	0.20	-0.10
Low probability	14.4002	44	0.26	0.19	-0.3684
Totals	54.3613	230			

Table 30. Distribution of events record sites in the cropmarks and soils model probability zones, based on site numbers.

Whilst testing of the cropmark and soils model produced results that cannot be regarded as convincing, looking at the distribution of the test sample sites within the three probability zones suggests that this model does have some currency but its low Kvamme's Gain should be noted. The test sample distribution in the other models indicates a by-chance distribution with no discernable patterning.

9 The cropmark visibility model: Summary and conclusions

In total the project dataset contains records for 1,759 prehistoric or Romano-British sites whose form is listed as cropmarks. Their distribution is not even, but characterised by significant concentrations in parts of central Cornwall and the northern part of the Lizard peninsula; cropmark distribution in east and southeast Cornwall by contrast is sparse.

The cropmark dataset is populated predominantly by rounds and enclosures, whose substantial enclosing ditches are more likely to produce visible cropmarks than the relatively slight ditches surrounding round houses or those forming field systems.

A number of different land classes were used in an attempt to build predictive models for below-ground prehistoric archaeology based on the distribution of cropmarks.

It quickly became apparent that Agricultural Land Classification (ALC) data is not suitable. The polygons forming each grade of land are highly generalised and their boundaries are consequently too schematic for predictive modelling purposes. Furthermore 83% of the project area is covered by only two of the ALC categories and the model produced from this data suggested below-ground archaeology is likely to be found anywhere apart from urban areas, some coastal areas, steep sided river valleys and high ground.

Polygons defining the soil types derived from Soil Survey mapping produced more satisfactory results. All the soils ranked in the model's high probability zone are loams. This is to be expected because loams are generally the most fertile soils and cereal production is most likely to take place here – hence the high number of cropmarks observed. It also seems more than likely that because the loams are the best soils prehistoric farmers would have established settlements here.

The most obvious feature of the soils model is the contrast between east Cornwall and the rest of the county. Much of the land east of the rivers Fowey and Camel is ranked as medium or low probability, whilst the greater part of Lowland Cornwall west of the Fowey contains extensive areas classed as high probability.

Bedrock geology mapping derived from BGS data is more detailed than that for either soils or Agricultural Land Classification. A significant problem with BGS data is that in some places where the data tiles meet, the bedrock types from neighbouring tiles do not always correspond.

The bedrock model shares some similarities with the soils model – most notably the tendency for much of the high probability zone to lie in western and central areas, but there are also some differences. Overall it does not perform as well as the soils model. It appears to be relatively broad brush in character and lacking subtlety, characterised by blocks of land falling into the various zones in contrast to the soils map in which the patterns of the three probability zones are far more inter-mingled.

The soils and bedrock models performed with similar levels of accuracy (capturing around 70% of sites in the high probability zone), but neither was particularly precise.

By combining the soils and bedrock data, a complex new dataset was created. This produced a model more precise but less accurate than those produced by either the soils or geology data on their own. The east/west divide characterising the other two models was also apparent in this model, but with more extensive tracts of land classed as the low probability zone occurring in western Cornwall.

Obviously the distribution pattern of known cropmarks is determined by the pattern of aerial reconnaissance. During the project, tracing this pattern focused on HE flight history only. This probably provides only a partial picture which may be a weakness in the modelling process.

In fact some sites in the cropmark dataset are recorded from locations which have not been flown by HE during the cropmark season. It is likely that these cropmarks were identified from photographs not taken by HE, such as the various sources of vertical photographs. This suggests that future programmes of HE reconnaissance in these areas would lead to the identification of more previously unrecorded cropmark sites.

The outstanding year for cropmark prospecting in Cornwall was 1989, but 1995 and 2004 also produced above average results. Regardless of year, the majority of successful flights were carried out in July.

There are three weaknesses in the flight data used for model building.

- Site entry into the HE Photo database appears to be inconsistent for some flights. This suggests that not all photographs from these flights have been recorded in the database and this in turn means that flight paths derived from the photo data are incomplete.
- The majority of cropmark sites were input to the HER as part of Cornwall's NMP project and in some cases only the photograph from which the NMP transcription was made was recorded in the database. As a result the number of cropmarks appearing on HE photographs is likely to be understated.
- For many of the earlier flights, precise flight path traces are not available (these were made by hand onto A4 map copies) leading to problems in defining the precise flight paths followed.

The flight data does demonstrate that some areas have been flown many times; others not at all, and that the rarely flown areas are, by and large, in east and southeast Cornwall.

The model based on the pattern of HE aerial reconnaissance is precise but not very accurate. It very clearly shows those locations where cropmarks are most likely to have been recorded – parts of the north coast, an east-west band in central Cornwall, and parts of the Lizard and Roseland peninsulas. On a broader level the model is similar to the models for cropmarks in relation to geology, soils and soils/geology, with east Cornwall depicted generally as an area of low probability.

A definitive cropmark visibility model was created by combining the soils/geology model with the aerial reconnaissance model. This cropmark visibility model clearly shows those locations where cropmarks are most likely to have been recorded. It is both accurate and reasonably precise.

Analysis of the model suggests that geology and soils are stronger predictors than the pattern of aerial reconnaissance.

The cropmark visibility map produced by this model reinforces the trend for east Cornwall generally to be characterised as an area of low probability. However, a more nuanced version of this broad trend was achieved by this model. The areas around Stratton and Bude, and an extensive area close to the southeast coast (between Fowey and St Germans) is classed as a high probability zone in this model whereas in the others much of the southeast coast is ranked in the low or medium probability zones.

Testing of the models was carried out using data extracted from the Cornwall events record. Only events in which below-ground features (such as ditches and pits) were recorded but which have not yet been entered into the HER were included in the test sample.

Testing was based both on numbers of sites (point data) and the area surveyed (polygons). In some cases the test results based on site numbers differ considerably from those based on site area. This suggests that the method used for defining the polygons (using present day boundaries of the fields in which each event took place) may have distorted the test results.

A further area of uncertainty inherent in the test sample was that geophysical anomalies which had not been further investigated by excavation were assumed to represent archaeological features, whereas experience shows that in some cases these anomalies turn out to be the result of natural agencies.

Notwithstanding these caveats the events record dataset does represent an independent test sample which can be expected to provide some measure of the veracity of the models. Furthermore, the distribution of the test events, whilst not even, reflects a good spread of sites across the whole project area.

With this in mind, an important result of the Lowland Cornwall project is the clear rejection of the cropmark visibility model by the test data. Both the medium and low probability zones perform better than expected, with the low probability zone in particular capturing more than twice as many sites as predicted and whose polygons cover more than twice the predicted area. The performance of the high probability zone is especially poor when the test is based on polygon area.

The conclusion to be drawn from the test is that the model does not accurately predict the location of below-ground prehistoric features. Had the model been strongly validated then it could be argued that it is a reliable indicator of the extent of the area in which below-ground archaeology is most likely to be found. Because the test sample represents previously unrecorded archaeology and the model was rejected it can be concluded that the model only shows those areas in which cropmarks are most likely to be visible. In other words it is exactly as described – a cropmark visibility model, which shows 'absence of evidence' rather than 'evidence of absence'. One possible implication is that the distribution of, for instance, rounds and enclosures (more than half of which are recorded as cropmarks in the HER) may be biased towards the high visibility zone of this model.

10 Rounds and enclosures and additional factors

The majority of rounds and enclosures in the Lowland Cornwall dataset are recorded as cropmarks and as a result the overall distribution of known rounds and enclosures is determined to a large extent by the distribution of the cropmark sites. As noted in Lowland Cornwall Volume 1 their distribution is not uniform across lowland Cornwall. Site densities are significantly higher in the western part of the county and there are notable concentrations or hot spots in central and western areas; for instance the Camel Estuary and the northern part of the Lizard Peninsula (Fig 20).

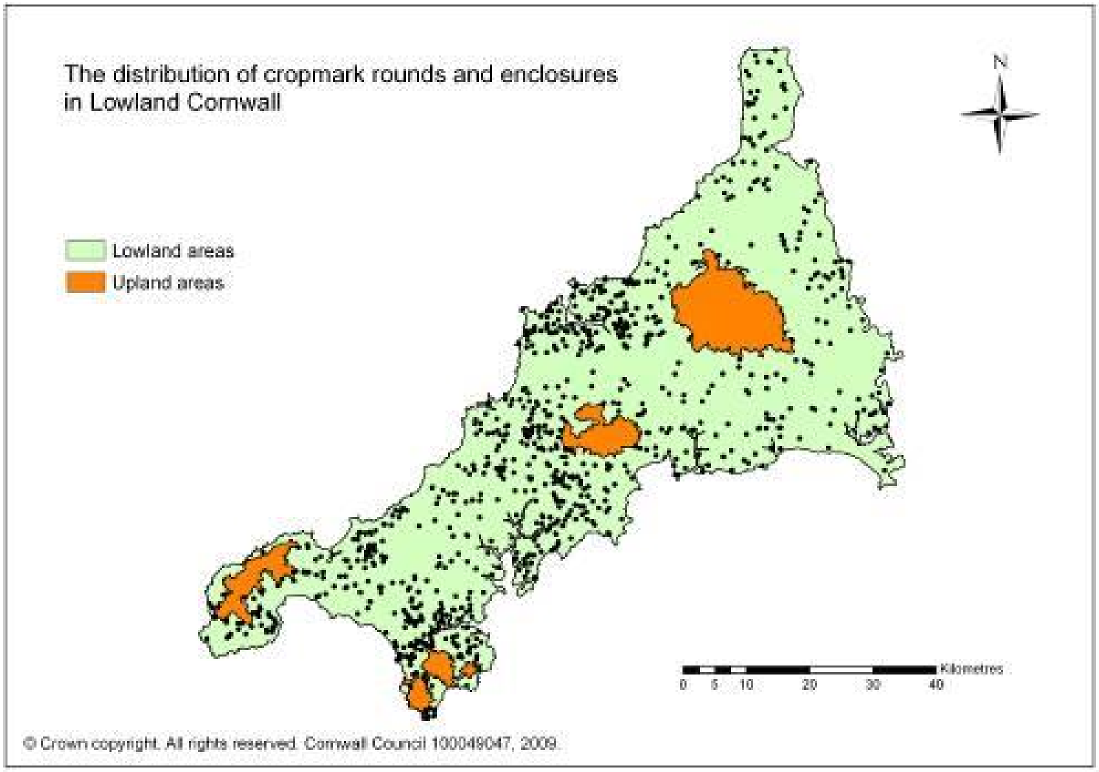

Fig 20. Map showing the distribution of cropmark rounds and enclosures in the Lowland Cornwall project area.

This distribution pattern is very much consistent with the cropmark models presented in this volume, with the eastern part of the county relatively bereft of sites. The fact that all these models were rejected to a greater or lesser extent when tested with events record data raises the question of to what extent is the distribution pattern of rounds and enclosures simply reflecting varying degrees of visibility across the project area. How many undiscovered enclosures are there in parts of the county where cropmarks do not readily form? On the other hand it should be emphasised that many of the buried features recorded in the events record are relatively slight (shallow ditches, pits, gullies, post holes, etc) compared with the often substantial ditches enclosing rounds, which might be expected to form cropmarks even in 'low visibility' areas.

To explore these questions further, models were made based on the correlation of the distribution of known rounds and enclosures with bedrock geology, soils and a combination of the two. These models were then compared and amalgamated with the high level model for rounds and enclosures correlated with HLC Types.

10.1 Rounds and enclosures against bedrock geology

The rounds and enclosures dataset was joined in GIS with the BGS bedrock geology data and a three zone model produced using the Kj parameter to measure the importance of each geology type and to define the cut off points between the three zones of the model. The result is set out in table 31 below.

Rounds and enclosures/bedrock geology model. High probability zone				
Bedrock	**Sites**	**PA**	**PS**	**Cum KJ**
Granite	298	0.0924	0.1524	0.0956
Sandstone and [subequal/subordinate] argillaceous rocks, interbedded	210	0.0752	0.1074	0.1547
Mudstone and sandstone	175	0.0644	0.0895	0.2023
Slate and siltstone	285	0.1535	0.1457	0.2327
Slate and sandstone, interbedded	104	0.0408	0.0532	0.2584
Hornfelsed slate and hornfelsed siltstone	113	0.0504	0.0578	0.2797
Mudstone, siltstone and sandstone	56	0.0135	0.0286	0.3025
Mudstone	49	0.0132	0.0251	0.3208
Slaty mudstone with sedimentary rock, metamorphic rock and igneous rock clasts	41	0.0066	0.0210	0.3405
Hornblende schist	33	0.0066	0.0169	0.3550
Peridotite and serpentinite	25	0.0060	0.0128	0.3649
Hornfelsed slate and hornfelsed sandstone	37	0.0243	0.0189	0.3644
Aplitic microgranite	17	0.0039	0.0087	0.3714
Totals	**1443**	**0.5508**	**0.738**	**0.3714**

Rounds and enclosures/bedrock geology model. Medium probability zone				
Bedrock	**Sites**	**PA**	**PS**	**Cum KJ**
Slate, siltstone and sandstone	123	0.1061	0.0629	0.3392
Gabbro	16	0.0040	0.0082	0.3459
Microgranite	20	0.0110	0.0102	0.3471
Sandstone	52	0.0450	0.0266	0.3299
Microgabbro	14	0.0048	0.0072	0.3343
Metabasaltic-rock	12	0.0021	0.0061	0.3406
Metamudstone and metasandstone	16	0.0106	0.0082	0.3392
Basaltic-rock	10	0.0020	0.0051	0.3441
Metabasalt	9	0.0026	0.0046	0.3476
Slate	53	0.0550	0.0271	0.3153
Basaltic lava	10	0.0048	0.0051	0.3167
Sandstone, siltstone and mudstone	43	0.0481	0.0220	0.2800
Chert	10	0.0076	0.0051	0.2767
Felsite	7	0.0034	0.0036	0.2776
Siltstone and mudstone, interbedded	4	0.0020	0.0020	0.2779
Tuff	3	0.0002	0.0015	0.2804
Metagabbro and metamicrogabbro	3	0.0008	0.0015	0.2819
Breccia	3	0.0016	0.0015	0.2819
Limestone and mudstone, interbedded	2	0.0001	0.0010	0.2836

Metalimestone and pelite	4	0.0043	0.0020	0.2802
Mica schist	2	0.0005	0.0010	0.2813
Totals	**416**	**0.3166**	**0.2125**	**0.2813**

Table 31. The high and medium probability zones of the rounds and enclosures/bedrock geology model. All other geology types make up the low probability zone.

The performance of the model is shown below.

Probability zone	PA	PS	Kvamme's gain	PS/PA
High	0.55	0.74	0.2534	1.34
Medium	0.32	0.21	0.0876	0.66
Low	0.13	0.05	-1.6788	0.38

The model is accurate, with three quarters of the enclosures captured in the high probability zone, but not very precise; the high probability zone takes up 55% of the project area. The model succeeds in predicted where enclosures are not located – with only 5% of the sites captured in the low probability zone. The main weakness of this model is the large size of the medium probability zone, which covers roughly a third of the project area – the model is indicating that this area is neither site likely nor site unlikely.

Given that so many of the rounds and enclosures are recorded as cropmarks it is not surprising that there are similarities between this model and the model for cropmarks correlated with bedrock geology. All the geology types making up the high probability zone of the cropmarks model are also in the high probability zone of this model, apart from sandstone (this forms part of the medium probability zone). However some of these types are ranked differently (granite, for example is the highest ranked type in the enclosures model, but is only ranked sixth in the cropmarks model). Also, there are more types making up the high probability zone of the enclosures model - Hornfelsed slate and Hornfelsed siltstone, Slaty mudstone with sedimentary rock, metamorphic rock and igneous rock clasts, Peridotite and serpentinite, Hornfelsed slate and Hornfelsed sandstone, and Aplitic microgranite are all highly ranked in this model but are placed in the medium or low probability zones of the cropmarks model. There are significant differences in the make up of the medium probability zones of the two models; most notably there are three times as many types making up the medium zone of the enclosures model.

There are also similarities between the probability map produced by this model (Fig 21) with that produced by the cropmarks model (Fig 8) – namely much of the high probability zone being in west and central areas, and the general broad brush nature of the map (caused by the nature of the geology dataset). There are differences, however: in the rounds and enclosures model there are hardly any areas of low probability zone in western Cornwall – for instance the Carnmenellis granite is classed as medium probability zone, but in the cropmarks model it is classed as low probability; around the Hensbarrow granite area the Hornfelsed slate and Hornfelsed sandstone forms part of the high probability zone, whereas in the cropmarks model this is ranked in the low probability zone; the area around Bude and Morwenstow forms part of the medium probability zone of the rounds and enclosures model but is part of the high probability zone of the cropmarks model and to the south of here the Culm measures are classed as low probability whereas they are medium probability in the cropmarks model (in fact the area of the Culm measures is the only extensive tract of low probability zone in the rounds and enclosures model).

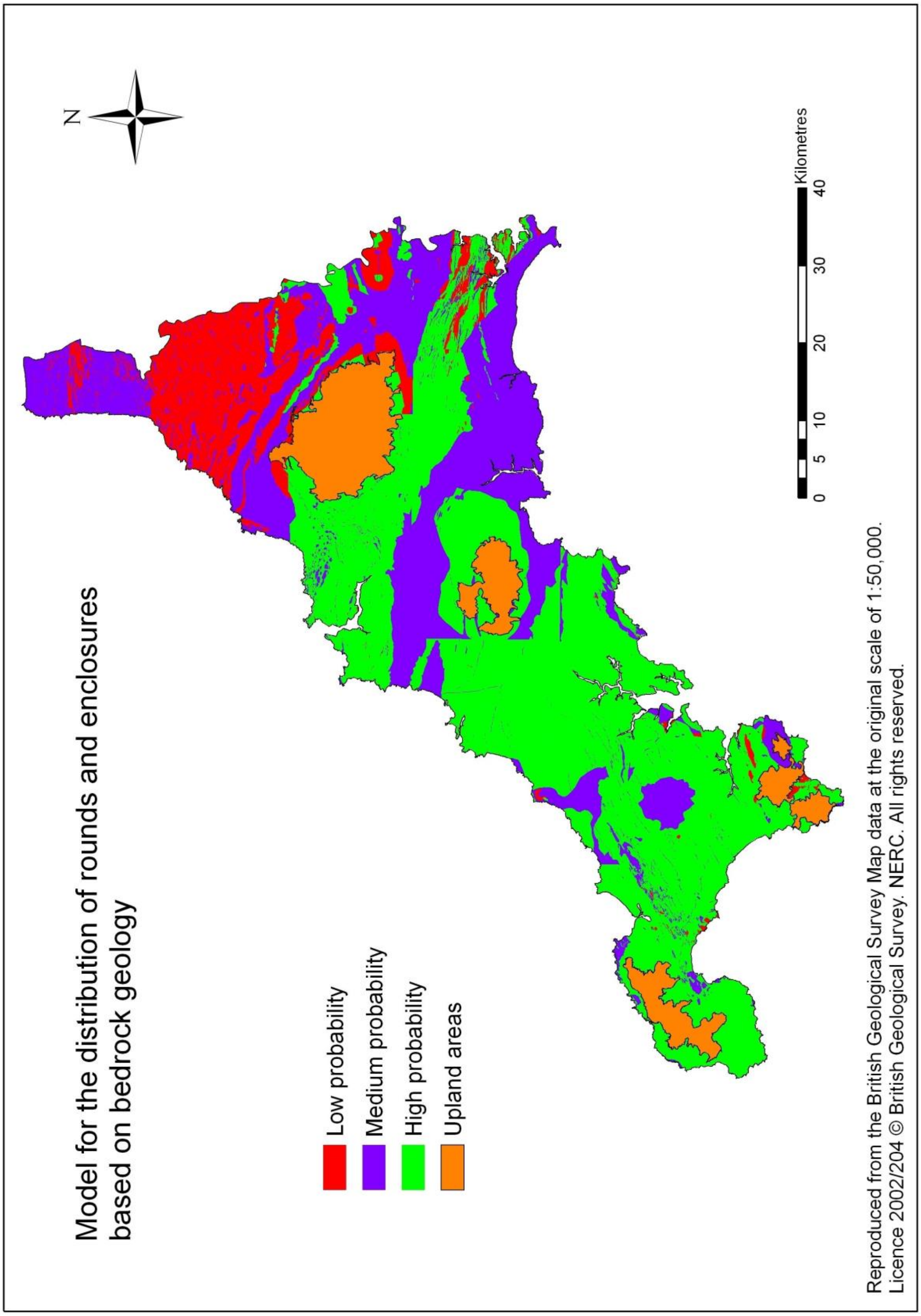

Fig 21. Map showing the model for rounds and enclosures based on bedrock geology.

10.1.1 Testing the rounds and enclosures/bedrock model

The rounds and enclosures/bedrock model was tested using events record data in the same way as the cropmark and visibility models (section 8). In the lowland events dataset 76 sites were interpreted as rounds or enclosures indicative of settlement evidence.

Forty three of the enclosures are already recorded in the HER and 33 are new sites. This data was used as a sample to test the rounds and enclosures/bedrock model in two ways. Firstly the whole dataset was used (as a largely internal test sample) and then only the new sites were used (as an independent test sample). Each test was carried out twice; first with sites represented by point data, and secondly with sites represented by polygons thereby basing the test on the area taken up by the sites as well as by site density.

Testing with all 76 sites

Zone	Area km^2	SA	NS	Predicted PS	PS	Predicted Sites	Sites
High probability	37.68	0.82	31	0.84	0.80	64	61
Medium probability	12.20	0.41	5	0.13	0.15	10	11
Low probability	4.48	0.23	1	0.03	0.05	2	4
Totals	54.36		37			76	76

Table 32. Results of events record testing of the rounds and enclosures/geology model: test based on numbers of sites. NS = notional number of sites predicted.

Zone	Area km^2	Predicted PS	PS	Predicted site area	Site area
High probability	37.68	0.84	0.82	347.03	338.28
Medium probability	12.20	0.13	0.12	56.37	49.95
Low probability	4.48	0.03	0.06	11.40	26.57
Totals	54.36			414.81	414.81

Table 33. Results of events record testing of the rounds and enclosures/geology model: test based on site area.

The results show that the distribution of rounds and enclosures contained in the events record is consistent with the model: the proportion of sites captured in the high probability zone is within 4% and 2% of that predicted whether the test is carried out using point or area data. When the test is based on point data both the medium and low probability zones perform slightly better than predicted at the expense of the high probability zone; when the test is carried out using area data the low probability zone performs better than expected at the expense of the other two zones. Nonetheless the results are a close fit with the model.

Testing with the 33 new sites

When the test is carried out with only the 33 new sites in the events record dataset, the results are more inconclusive. In the test based on point data (table 34) the proportion of sites captured in the high probability zone is 11% less than predicted and both the medium and low probability zones perform better than predicted, especially the low probability zone, which captures three times as many sites as expected.

By contrast, when the test is based on site area, the high probability zone performs almost exactly as predicted. However, the low probability zone still performs above expectations and the medium probability zone captures 4% less of the proportion of sites than expected.

Zone	Area km²	SA	NS	Predicted PS	PS	Predicted Sites	Sites
High probability	37.68	0.82	31	0.84	0.73	28	24
Medium probability	12.20	0.41	5	0.13	0.18	4	6
Low probability	4.48	0.23	1	0.03	0.09	1	3
Totals	54.36		37			33	33

Table 34. Results of events record testing of the rounds and enclosures/geology model: test based on numbers of sites. NS = notional number of sites predicted.

Zone	Area km²	Predicted PS	PS	Predicted site area	Site area
High probability	37.68	0.84	0.83	140.03	139.55
Medium probability	12.20	0.13	0.09	22.75	15.29
Low probability	4.48	0.03	0.07	4.60	12.53
Totals	54.36			167.38	167.38

Table 35. Results of events record testing of the rounds and enclosures/geology model: test based on site area.

One important result of the events record testing is the obvious importance of the rock type Mudstone and sandstone. Eighteen of the 76 rounds and enclosures in the events dataset and nine of the 33 new rounds are found on this geology (more than twice the number on any other rock type).

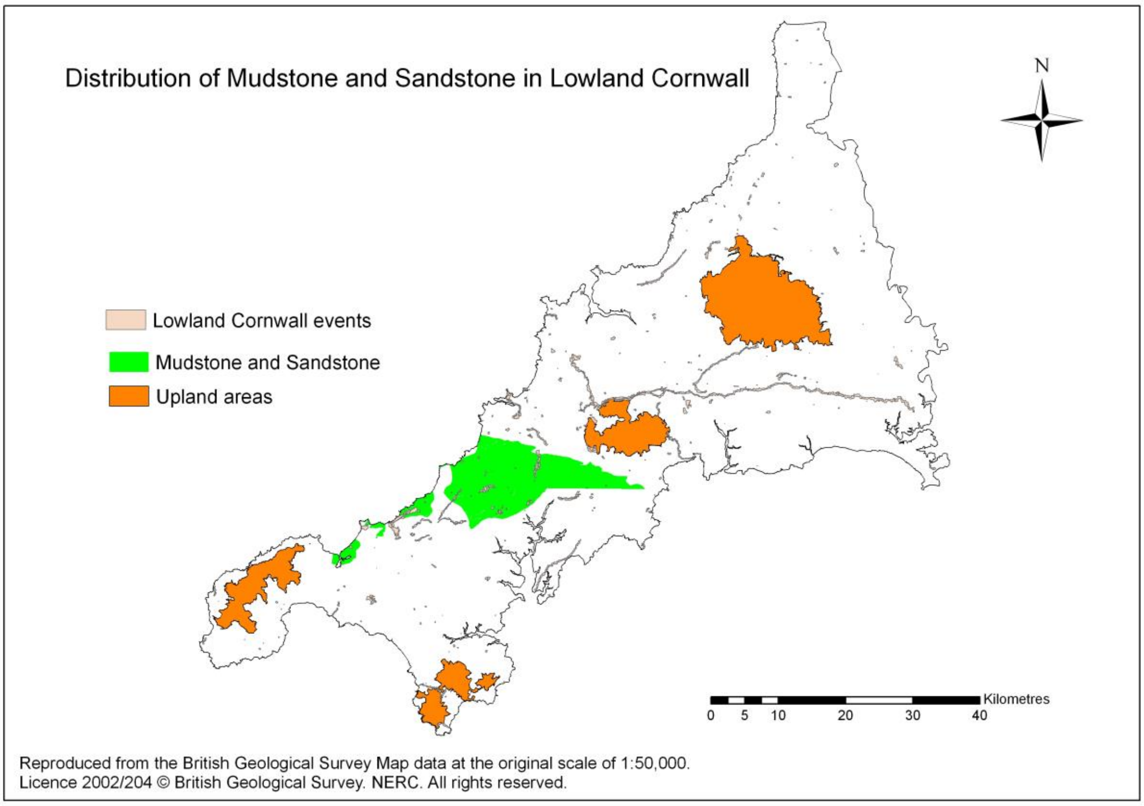

Fig 22. Map showing the distribution of the rock type Mudstone and Sandstone in lowland Cornwall.

Mudstone and Sandstone forms only 9.5% of the area surveyed by the events but 24% of the rounds and enclosures (and 27% of the new rounds and enclosures) are located on this rock type and if it was regarded as the high probability zone of a model, it would produce a Kvamme's gain of 0.6004. It does, however, have a limited distribution in lowland Cornwall (Fig 22).

10.2 Rounds and enclosures against soil types

The rounds and enclosures dataset was joined in GIS with the soils data and a three zone model produced using the Kj parameter to measure the importance of each soil type and to define the cut off points between the three zones of the model. The result is set out in table 36 below.

Rounds and enclosures/soils model. High probability zone				
Soil	**Sites**	**PA**	**PS**	**Cum KJ**
Denbigh 2	751	0.2948	0.3838	0.1848
Moretonhampstead	268	0.0734	0.1369	0.2818
Powys	183	0.0604	0.0935	0.3376
Trusham	125	0.0397	0.0639	0.3771
Moor gate	65	0.0280	0.0332	0.3911
Totals	**1392**	**0.4963**	**0.7113**	**0.3911**
Rounds and enclosures/soils. Medium probability zone				
Soil	**Sites**	**PA**	**PS**	**Cum KJ**
Denbigh 1	252	0.2151	0.1288	0.3288
Sportsmans	34	0.0184	0.0174	0.3308
Totals	**286**	**0.2335**	**0.1462**	**0.3308**

Table 36. The high and medium probability zones of the rounds and enclosures/soils model. All other soil types make up the low probability zone.

The performance of the model is shown below.

Probability zone	**PA**	**PS**	**Kvamme's gain**	**PS/PA**
High	0.4963	0.7113	0.3023	1.43
Medium	0.2335	0.1461	-0.5978	0.63
Low	0.2702	0.1426	-0.8953	0.53

The model performs well and produces a moderate Kvamme's gain. The high probability zone captures more than 70% of the sites so it is accurate but, because this zone covers almost half the project area the model is not very precise. The ratios of Indicative Values indicate that one is more than twice as likely to encounter an enclosure in the high probability zone as in the medium zone and almost three times more likely than in the low probability zone.

Although there was an obvious cut off point between medium and low probability zones (the highest ranked soil type in the low zone, Manod scored a cumulative Kj measure of 0.2715) the medium and low probability zones are virtually interchangeable – both capturing 14-15% of the sites and both covering a similar sized area (23-27% of the project area). Because of this the model could be regarded as a two zone model with a low probability zone covering 51% of lowland Cornwall and capturing around 29% of the sites.

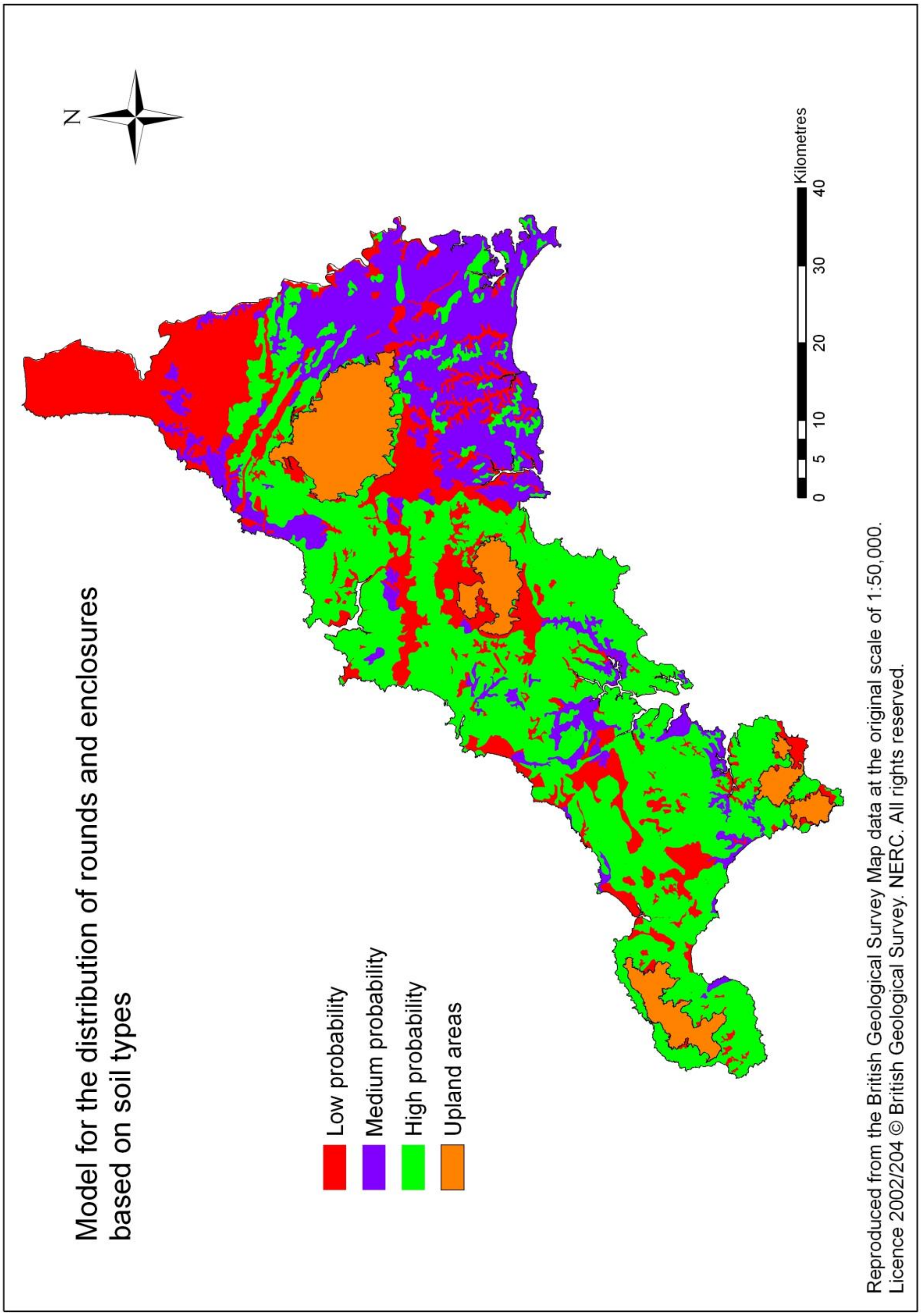

Fig 23. Map showing the rounds and enclosures and soils model.

There are similarities between this model and the cropmarks/soils model: the four soil types making up the high probability zone of the cropmarks model (Denbigh 2, Powys, Moretonhampstead and Trusham) are all in the high zone of this model. However, they are joined in this model by the soil type Moor gate, which was ranked in the low probability zone of the cropmarks model. The other difference is that the types Neath, Hallsworth 1 and Yeollandpark, which were in the medium zone of the cropmarks model, are ranked in the low probability zone of the rounds and enclosures model.

The maps produced by the two models are also very similar (Figs 7 and 23). The main difference is that the area around Bude and Morwenstow and the Culm measures are classed as low probability in the rounds model but as a mixture of medium and low probability zone in the cropmarks model. Another, less obvious difference is that some areas of west Cornwall, including parts of West Penwith, are classed as low probability in the cropmarks model but are in the high probability zone of the rounds model.

10.2.1 Testing the rounds and enclosures/soils model

The rounds and enclosures/soils model was tested using events record data in the same way as the cropmark and visibility models (section 8).

Testing with all 76 sites

Zone	Area km^2	SA	NS	Predicted PS	PS	Predicted Sites	Sites
High probability	25.46	0.88	22	0.69	0.71	52	54
Medium probability	14.06	0.38	5	0.16	0.11	12	8
Low probability	14.85	0.32	5	0.15	0.18	11	14
Totals	54.36		32			76	76

Table 37. Results of events record testing of the rounds and enclosures model: test based on numbers of sites. NS = notional number of sites predicted.

Zone	Area km^2	Predicted PS	PS	Predicted site area	Site area
High probability	25.46	0.69	0.63	285.98	263.05
Medium probability	14.06	0.16	0.10	68.19	43.09
Low probability	14.85	0.15	0.26	60.65	108.67
Totals	54.36			414.81	414.81

Table 38. Results of events record testing of the rounds and enclosures model: test based on site area.

As can be seen from these tables, when tested using point data the high probability zone performs somewhat better than predicted, which validates the model. However, the low probability zone performs better than expected at the expense of the medium probability zone. When the test is based on site area the low probability zone performs better than expected at the expense of both high and medium probability zones.

In fact the model results underline the fact that this is, in effect, a two zone model. If the model were correct one would expect to find 69% of the rounds in the high probability zone and 31% in the low zone, so the test based on point data produces a closer fit than that based on site area.

Testing with the 33 new sites

Zone	Area km^2	SA	NS	Predicted PS	PS	Predicted Sites	Sites
High probability	25.46	0.88	22	0.69	0.76	23	25
Medium probability	14.06	0.38	5	0.16	0.09	5	3
Low probability	14.85	0.32	5	0.15	0.15	5	5
Totals	54.36		32			33	33

Table 39. Results of events record testing of the rounds and enclosures model: test based on numbers of sites. NS = notional number of sites predicted.

Zone	Area km^2	Predicted PS	PS	Predicted site area	Site area
High probability	25.46	0.69	0.83	115.39	138.73
Medium probability	14.06	0.16	0.07	27.51	11.43
Low probability	14.85	0.15	0.10	24.47	17.22
Totals	54.36			167.38	167.38

Table 40. Results of events record testing of the rounds and enclosures model: test based on site area.

When the model is tested with only the 33 new rounds and enclosures the validity of the high probability zone is even more conclusively affirmed, with a significantly higher than predicted proportion of the sites captured within it, especially when the test is based on site area. This better than expected performance is mostly at the expense of the medium probability zone.

Denbigh 2 is the most important soil type in the tests, capturing 57% of all the rounds and enclosures and 73% of the new enclosures but covering only 33% of the area surveyed.

10.3 Rounds and enclosures against geology and soils

The rounds and enclosures dataset was joined in GIS with the combined soils and bedrock data. In total this layer consists of 558 separate combinations of intersected geology and soil types. A three zone model was then produced using the Kj parameter to measure the importance of each combination and to define the cut off points between the three zones of the model. The high probability zone comprises 14 different combinations of soils and bedrock types, the medium zone contains 20 different combinations and the low probability zone contains the remaining 524 combinations. The result is set out in table 41 overleaf and the overall performance of the model is outlined below.

Probability zone	PA	PS	Kvamme's gain	PS/PA
High	0.35	0.57	0.3860	1.63
Medium	0.31	0.22	-0.3801	0.72
Low	0.34	0.21	-0.6429	0.61

Table 41. Model for the distribution of rounds and enclosures correlated with bedrock geology and soils: high and medium probability zones.

Rounds and enclosures/Geology and soils. High probability zone					
Bedrock	**Soil**	**Rounds**	**PA**	**PS**	**Kj**
Granite	Moretonhampstead	240	0.0640	0.1225	0.0847
Sandstone and [subequal/subordinate] argillaceous rocks, interbedded	Denbigh 2	143	0.0357	0.0730	0.1369
Mudstone and sandstone	Denbigh 2	135	0.0424	0.0689	0.1799
Slate and siltstone	Denbigh 2	122	0.0492	0.0623	0.2104
Slate and sandstone, interbedded	Denbigh 2	85	0.0276	0.0434	0.2365
Hornfelsed slate and Hornfelsed siltstone	Denbigh 2	68	0.0257	0.0347	0.2546
Slate and siltstone	Powys	56	0.0153	0.0286	0.2742
Granite	Moor gate	47	0.0171	0.0240	0.2873
Slaty mudstone with sedimentary rock, metamorphic rock and igneous rock clasts	Denbigh 2	38	0.0051	0.0194	0.3047
Mudstone	Powys	36	0.0074	0.0184	0.3192
Mudstone, siltstone and sandstone	Powys	35	0.0074	0.0179	0.3330
Hornblende schist	Trusham	32	0.0060	0.0163	0.3462
Slate, siltstone and sandstone	Denbigh 2	42	0.0277	0.0214	0.3482
Sandstone and [subequal/subordinate] argillaceous rocks, interbedded	Denbigh 1	34	0.0182	0.0174	0.3530
Totals		**1113**	**0.3488**	**0.5682**	**0.3530**

Rounds and enclosures/Geology and soils. Medium probability zone					
Bedrock	**Soil**	**Rounds**	**PA**	**PS**	**Kj**
Slate and siltstone	Denbigh 1	50	0.0432	0.0255	0.3460
Slate, siltstone and sandstone	Denbigh 1	57	0.0550	0.0291	0.3308
Hornfelsed slate and Hornfelsed siltstone	Manod	27	0.0131	0.0138	0.3350

Rounds and enclosures/Geology and soils. Medium probability zone					
Bedrock	**Soil**	**Rounds**	**PA**	**PS**	**Kj**
Slate	Denbigh 1	32	0.0270	0.0163	0.3289
Slate and siltstone	Trusham	23	0.0089	0.0117	0.3347
Sandstone	Neath	28	0.0242	0.0143	0.3282
Sandstone and [subequal/subordinate] argillaceous rocks, interbedded	Powys	22	0.0136	0.0112	0.3284
Slate and siltstone	Manod	26	0.0279	0.0133	0.3157
Mudstone and sandstone	Sportsmans	16	0.0036	0.0082	0.3225
Aplitic microgranite	Moretonhampstead	15	0.0023	0.0077	0.3301
Mudstone and siltstone	Denbigh 1	19	0.0177	0.0097	0.3234
Peridotite and serpentinite	Croft Pascoe	14	0.0037	0.0071	0.3289
Hornfelsed slate and Hornfelsed sandstone	Manod	15	0.0089	0.0077	0.3291
Slate	Denbigh 2	17	0.0186	0.0087	0.3195
Hornfelsed slate and Hornfelsed sandstone	Denbigh 2	13	0.0069	0.0066	0.3207
Microgranite	Moor gate	13	0.0077	0.0066	0.3208
Sandstone, siltstone and mudstone	Denbigh 2	14	0.0114	0.0071	0.3171
Metamudstone and metasandstone	Denbigh 2	12	0.0052	0.0061	0.3195
Peridotite and Serpentinite	Trusham	11	0.0022	0.0056	0.3248
Slate and sandstone, interbedded	Powys	11	0.0052	0.0056	0.3265
Totals		**435**	**0.3063**	**0.2219**	**0.3265**

The gain measure of the high probability zone is higher than in either the soils or bedrock models and this model is quite precise, with the high probability zone covering only 35% of the project area. However the model lacks accuracy in that only 57% of the sites are captured in this zone and that the zones of medium and low probability are very similar in size and contain very similar proportions of sites – to all intents and purposes they are interchangeable. The probability map based on this model is shown in Fig 24.

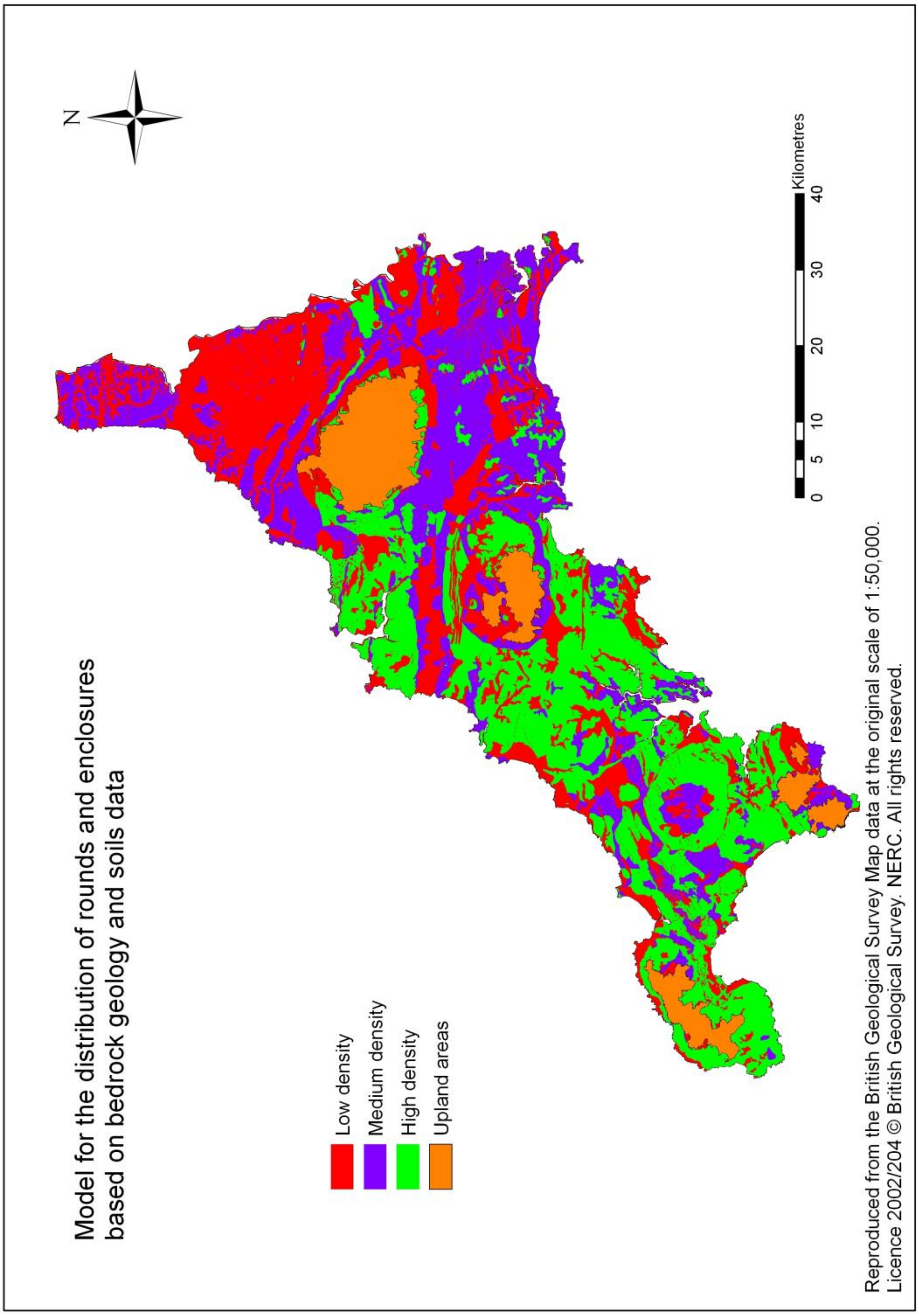

Fig 24. Map showing the rounds and enclosures and soil/bedrock model.

10.3.1 Testing the rounds and enclosures/soils and bedrock model

The rounds and enclosures/soils and bedrock model was tested using events record data in the same way as the cropmark and visibility models (section 8).

Testing with all 76 sites

Zone	Area km^2	SA	NS	Predicted PS	PS	Predicted Sites	Sites
High probability	17.22	1.00	17	0.51	0.57	39	43
Medium probability	17.73	0.45	8	0.23	0.22	18	17
Low probability	19.39	0.45	9	0.26	0.21	20	16
Totals	54.36		34			76	76

Table 42. Results of events record testing of the rounds and enclosures/soils and bedrock model: test based on numbers of sites. NS = notional number of sites predicted.

Zone	Area km^2	Predicted PS	PS	Predicted site area	Site area
High probability	17.22	0.51	0.51	210.56	211.88
Medium probability	17.73	0.23	0.23	97.56	96.00
Low probability	19.39	0.26	0.26	106.69	106.93
Totals	54.36			414.81	414.81

Table 43. Results of events record testing of the rounds and enclosures model: test based on site area.

The test results are a very close fit to the model, especially when based on site area: in this case the test sample performs exactly as predicted. When the test is based on point data the high probability zone performs better than predicted at the expense of the other two zones, but primarily the low probability zone.

Testing with the 33 new sites

When the test is carried out using point data for the 33 new enclosures, a close fit is achieved for the high probability zone, but the medium probability zone performs better than predicted at the expense of the low probability zone. When the test is based on site area the high probability zone performs much better than predicted at the expense of the low probability zone, whilst the medium probability zone performs much as predicted.

Zone	Area km^2	SA	NS	Predicted PS	PS	Predicted Sites	Sites
High probability	17.22	1.00	17	0.51	0.48	17	16
Medium probability	17.73	0.45	8	0.23	0.33	8	11
Low probability	19.39	0.45	9	0.26	0.18	8	6
Totals	54.36		34			33	33

Table 44. Results of events record testing of the rounds and enclosures/soils and bedrock model: test based on numbers of sites. NS = notional number of sites predicted.

Zone	Area km²	Predicted PS	PS	Predicted site area	Site area
High probability	17.22	0.51	0.63	84.96	105.50
Medium probability	17.73	0.23	0.26	39.36	43.82
Low probability	19.39	0.26	0.11	43.05	18.06
Totals	54.36			167.38	167.38

Table 45. Results of events record testing of the rounds and enclosures model: test based on site area.

10.4 Discussion and further analysis

The distribution of rounds and enclosures closely resembles that of cropmarks, with relatively few in the east (not surprising in that rounds and enclosures make up almost 60% of the cropmark dataset).

It was possible to build models with good levels of accuracy using soil types and bedrock geology types as variables, and the model based on soil types was rather more precise than that based on geology. The model based on a combination of bedrock and soils was not as accurate but more precise.

All three models also bear similarities with the models for cropmark sites but there are differences. For instance the rock type Hornfelsed slate and Hornfelsed sandstone is ranked in the high probability zone of the rounds/bedrock model, but in the low probability zone of the cropmarks model. This is significant because the large number of below ground features recorded in the events record from this rock type was the main cause of the cropmarks/bedrock model being rejected when tested (section 8.2.3, tables 22 and 23).

Although there were clearly defined cut off points enabling the construction of a three zone model for rounds and enclosures using bedrock as the sole variable, it was difficult to identify cut off points for the medium and low probability zones in the models for soils and bedrock/soils combined. These models in effect perform as two zone models with high and low probability zones.

The most significant difference between the performance of the rounds/enclosures models and the cropmark sites models is that when tested with events record data the test sample provided a close or reasonably close fit to all the rounds models in contrast to the cropmark models, which were all rejected. This suggests that the high number of rounds and enclosures found on certain soil and rock types reflects a deliberate preference for those locations, rather than simply being the result of factors influencing cropmark formation and visibility. Rounds and enclosures are recorded from these areas because the soils are the most fertile or easily cultivated, not because they are most conducive to cropmark formation.

This suggestion is supported by analysis of the form of survival of the enclosures. This is summarised in the tables below.

ZONE	Cropmark	Extant	Documentary	Site of	Geophysics	TOTAL
HIGH	73%	68%	70%	60%	100%	**71%**
MEDIUM	14%	13%	17%	32%	0%	**15%**
LOW	13%	19%	13%	8%	0%	**14%**

Table 46. Analysis of form of survival of sites in each probability zone of the rounds and enclosures/soils model

ZONE	Cropmark	Extant	Documentary	Site of	Geophysics	TOTAL
HIGH	75%	68%	77%	79%	100%	**74%**
MEDIUM	20%	25%	19%	16%	0%	**21%**
LOW	5%	7%	4%	5%	0%	**5%**

Table 47. Analysis of form of survival of sites in each probability zone of the rounds and enclosures/bedrock model

ZONE	Cropmark	Extant	Documentary	Site of	Geophysics	TOTAL
HIGH	60%	52%	55%	45%	75%	**57%**
MEDIUM	21%	23%	24%	31%	25%	**22%**
LOW	19%	25%	21%	24%	0%	**21%**

Table 48. Analysis of form of survival of sites in each probability zone of the rounds and enclosures/bedrock and soils model

In these three tables the final column on the right shows the percentage of all rounds and enclosures captured in each of the probability zones, so 71% of the enclosures are captured in the high probability zone of the soils model, 15% in the medium probability zone and 14% in the low zone. The other columns show the percentage of enclosures in each probability zone according to their form as recorded in the HER. For instance, 73% of cropmark enclosures, 68% of extant enclosures, 70% of enclosures recorded from documentary evidence, 60% of enclosures recorded as 'site of' and all enclosures found by geophysical survey are captured in the high probability zone of the soils model (only four enclosures are recorded in the HER from geophysical survey so this figure is not significant).

Ignoring the enclosures recorded by geophysical survey, the percentages of captured enclosures according to form (68% - 73%) closely resemble the overall percentage of enclosures captured in the high probability zone, apart from those recorded as 'site of' (60%). The same is true of the medium probability zone of this model: 14% of the cropmarks, 13% of extant enclosures and 17% of documentary sites compared with 15% of all enclosures (interestingly 32% of 'site of' enclosures are captured in this zone). And in the low probability zone 13% of cropmarks, 19% of extant sites and 13% of documentary sites as opposed to 14% of all enclosures.

In other words the percentages of extant and documentary enclosures in each of the zones are broadly similar (very similar in some cases) to the percentage of all enclosures captured in each zone. The same pattern is apparent in the other two models. Whilst there is a trend for the percentage of cropmark enclosures in the high probability zones and the percentage of extant enclosures in the low probability zones to be slightly higher than the overall percentages, the variance is not large. The great majority of extant and documentary enclosures are captured in the high probability zones of all three models. It can be concluded from this analysis that if there is any bias in the models towards cropmark sites then it is minimal.

The possibility that there may be archaeologically significant regional variations in the distribution of rounds and enclosures is further suggested by variations in the density of enclosure distribution in areas overlying similar soil types. The best example is Denbigh 1 and Denbigh 2 soils, both of which are described as 'loam over shale' with the underlying geology being 'Palaeozoic slaty mudstone and siltstone'. These are by far the two most extensive soil types in Lowland Cornwall, covering 21% and 29% of the landscape respectively. Both overlie the rock types 'slate and siltstone' and 'slate, siltstone and sandstone': Denbigh 2 also overlies the important rock type 'mudstone and siltstone' whereas Denbigh 1 overlies 'mudstone and sandstone'. Denbigh 1 overlies slate, and Denbigh 2 overlies 'slate and sandstone interbedded'. So although there are some differences in the underlying geology there are similarities. However, virtually three times as many rounds and enclosures are recorded from Denbigh 2 soils than from Denbigh 1 (751 compared with 252). The PS to PA ratio for Denbigh 2 is 0.38 to 0.29 and for Denbigh 1 it is 0.13 to 0.21 and the Relative Gain for Denbigh 2 is 0.089 whilst for Denbigh 1 it is -0.086 and one is twice as likely to encounter an enclosure on Denbigh 2 soils as on Denbigh 1. There is a striking regional distinction in the distribution of the two soils, with Denbigh 2 predominantly occurring in central areas and Denbigh 1 mainly confined to southeast Cornwall (Fig 25).

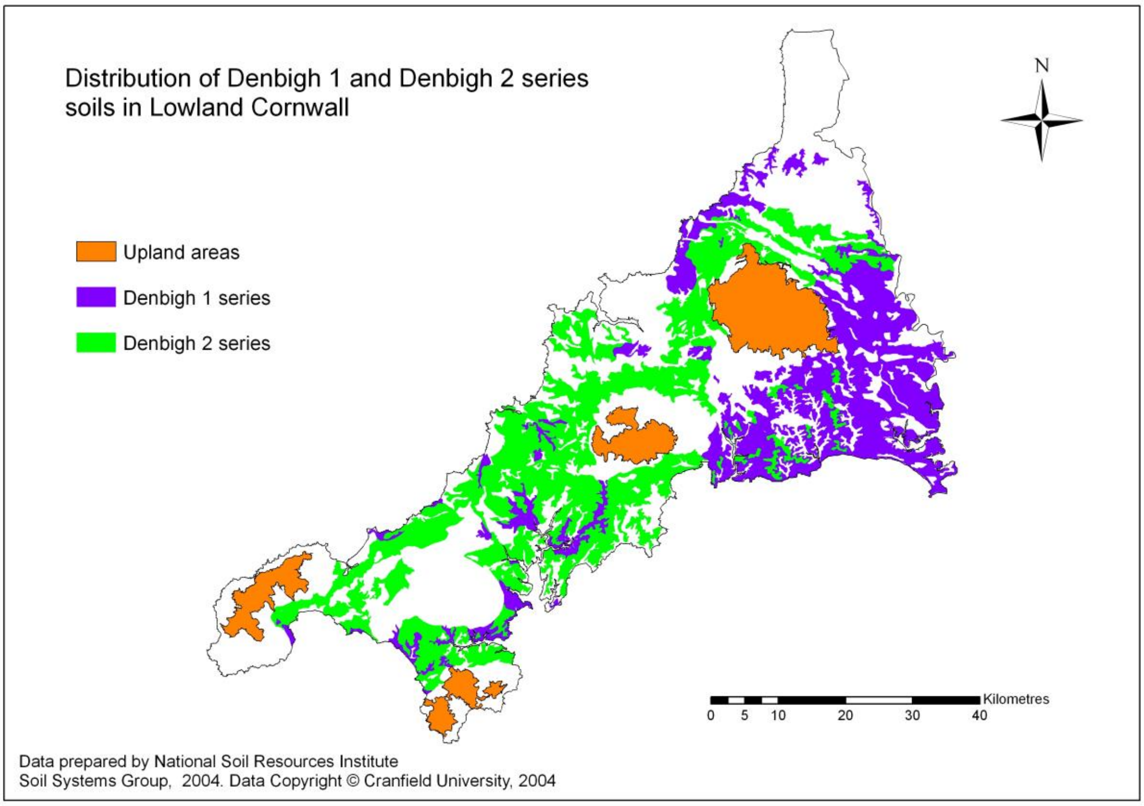

Fig 25. Map showing the distribution of Denbigh1 and Denbigh 2 soil types in lowland Cornwall.

The fact that there is such a clear disparity between the densities of enclosure distribution over two virtually identical soil types does hint at a genuine regional variation to the pattern, with fewer enclosures in eastern and south eastern areas of the county.

11 Predictive models using soils, geology and HLC Types as variables

The models presented in section 10, which correlate the distribution of rounds and enclosures with geology and soils, appear to accurately reflect the known distribution of rounds and enclosures in Cornwall with an apparent preference towards western and central areas. This contrasts with the model based on the distribution of rounds and enclosures correlated with HLC Types presented in Lowland Cornwall volume 1. The high probability zone of that model includes extensive areas in east Cornwall, whilst parts of west Cornwall (notably around Camborne, Redruth and Hayle) are classed as medium or low probability. A major issue with the HLC model discussed in volume 1 is its lack of precision, reflected in its Kvamme's gain of only 0.1715, due to the large area covered by the high probability zone. In an attempt to increase model precision, the HLC Types were combined with the bedrock geology and soils datasets, and the combined datasets were used as variables for modelling the distribution of rounds and enclosures. The results are presented in this section.

11.1 The rounds and enclosures/soils and HLC model

The soils and HLC Types polygons were combined by performing a spatial union in GIS and the resulting shapefile comprised 410 different combinations of soils and HLC Types. Rounds and enclosures were present in 131 of these combinations. A predictive model was created using the Kj parameter method and the result is shown in table 49 below; a probability map based on the model is shown in Fig 26.

When soils as well as HLC Types are used as variables, much of eastern Cornwall, which is largely classed as part of the high probability zone in the rounds/HLC model, is ranked in the medium and low probability zones. There are, however, scattered tracts of land in the southeast included in the high probability zone, as well as a more extensive band of land to the north of Bodmin Moor. The Carnmenellis area in west/central Cornwall is included in the low probability zone in the soils and HLC model, as is the land to the north of the Hensbarrow granite. St Breock Downs are very clear as a west – east band of land classed as low probability, as is much of the north coastal strip. In the rounds/HLC model these areas are characterised as a combination of medium and low probability, and much of the area to the north of Hensbarrow is classed as high probability.

The model is much less 'broad brush' than the rounds/soils model (Fig 23). For instance in the rounds/soils model virtually the whole of West Penwith is classed as part of the high probability zone, whereas in the combined model there are considerable tracts of medium and low probability land, particularly in coastal areas. The performance of the model is summarised below.

Probability zone	PA	PS	Kvamme's gain	PS/PA	Sites per km^2
High	0.41	0.62	0.3373	1.51	0.93
Medium	0.25	0.23	-0.1185	0.89	0.55
Low	0.34	0.15	-1.1981	0.45	0.28

This model is not as accurate as some others discussed in this volume, with only 62% of the enclosures captured in the high probability zone, but it is reasonably precise – the high probability zone covering only 41% of lowland Cornwall. The model's accuracy could be increased by including the highest ranked combination from the medium probability zone (HLC Type Farmland Medieval and soil type Denbigh 1; see table 49) in the high probability zone. If this were done 71% of the enclosures would now be captured in the high probability zone, but this would cover 54% of the project area and the zone would produce a Kvamme's gain of 0.2318 as opposed to 0.3373. For this reason the optimum performance of the model is that set out in the table above. The medium probability zone is quite well defined, taking up less than a quarter of the project area as is the low probability zone, which covers 34% of the project area but only captures 15% of the enclosures. One is 1.6 times more likely to encounter and enclosure in the high probability zone than in the medium zone and 3.3 times more likely than in the low probability zone.

Not surprisingly the highest ranked combinations are made up of high ranking types from the individual HLC and soils models: of the nine combinations making up the high probability zone, five contain the HLC Type Farmland Medieval, two contain Farmland C20 and one contains Farmland Prehistoric; three contain the soil type Denbigh 2, two each contain Moretonhampstead and Trusham and one contains Powys. The medium probability HLC Type Farmland Post medieval is classed in the high probability zone of this combined model where it overlays Denbigh 2 soils, and the low ranked soil type Manod is included in the high probability zone when it underlies the high ranked HLC Type Farmland Medieval. Neither of these combinations performs particularly well in terms of their Relative Gain (PS-PA) however: Farmland Post medieval combined with Denbigh 2 covers 5.5% of lowland Cornwall and captures 5.4% of the enclosures (a virtually by chance pattern), giving a Relative Gain of -0.0014; Manod soils combined with Farmland Medieval covers 4% of lowland Cornwall and captures 3% of the enclosures, giving a Relative Gain of -0.0089. By comparison, the highest ranked combination, Farmland Medieval and Denbigh 2 covers 17% of lowland Cornwall and captures 26% of the enclosures, resulting in a Relative Gain of 0.09.

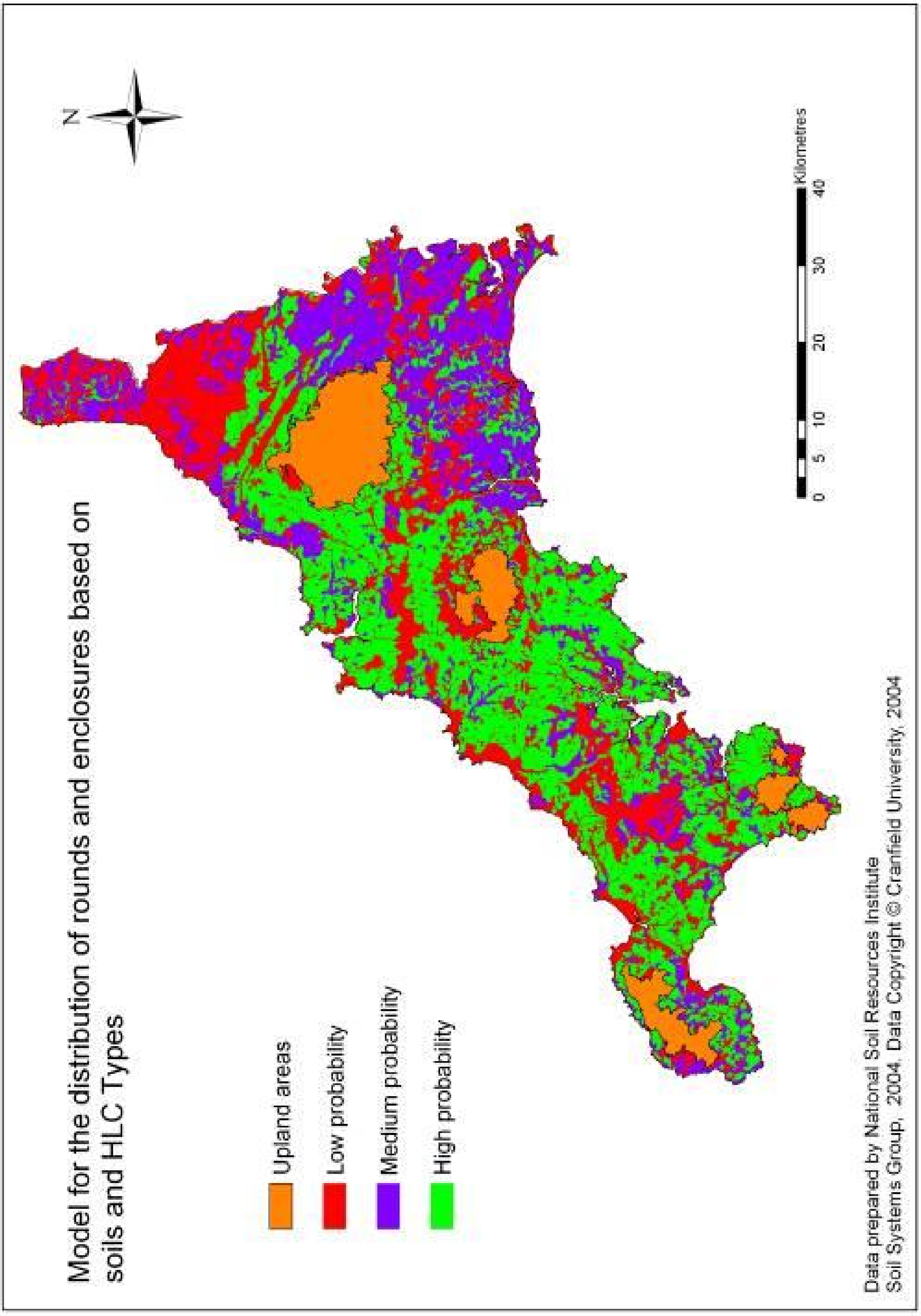

Fig 26. Map showing predictive model for rounds and enclosures using soils and HLC Types as variables.

Table 49. Model for the distribution of rounds and enclosures correlated with soils and HLC Types.

Rounds and enclosures model based on soils and HLC Types: High probability zone						
HLC Type	**Soil**	**Area (km2)**	**Rounds**	**PA**	**PS**	**Kj**
Farmland Medieval	Denbigh 2	548.52	515	0.1720	0.2632	0.1549
Farmland Prehistoric	Moretonhampstead	62.98	121	0.0197	0.0618	0.2081
Farmland Medieval	Powys	121.94	131	0.0382	0.0669	0.2520
Farmland Medieval	Trusham	86.63	94	0.0272	0.0480	0.2836
Farmland Medieval	Moretonhampstead	98.10	92	0.0308	0.0470	0.3114
Farmland Post medieval	Denbigh 2	177.26	106	0.0556	0.0542	0.3271
Farmland C20	Denbigh 2	77.84	74	0.0244	0.0378	0.3496
Farmland Medieval	Manod	127.83	61	0.0401	0.0312	0.3513
Farmland C20	Trusham	13.46	23	0.0042	0.0118	0.3612
Totals		**1314.56**	**1217**	**0.4122**	**0.6219**	**0.3612**

Rounds and enclosures model based on soils and HLC Types: Medium probability zone						
HLC Type	**Soil**	**Area (km2)**	**Rounds**	**PA**	**PS**	**Kj**
Farmland Medieval	Denbigh 1	420.94	169	0.1320	0.0864	0.3410
Farmland C20	Moretonhampstead	17.20	23	0.0054	0.0118	0.3504
Farmland C20	Powys	13.32	20	0.0042	0.0102	0.3591
Farmland C20	Denbigh 1	87.52	40	0.0274	0.0204	0.3568
Farmland Medieval	Moor gate	26.74	23	0.0084	0.0118	0.3631
Farmland Post medieval	Powys	18.76	20	0.0059	0.0102	0.3701
Farmland Medieval	Yeollandpark	14.98	15	0.0047	0.0077	0.3750
Farmland Post medieval	Moretonhampstead	25.89	17	0.0081	0.0087	0.3777
Farmland Post medieval	Sportsmans	26.52	17	0.0083	0.0087	0.3802

Rounds and enclosures model based on soils and HLC Types: Medium probability zone						
HLC Type	**Soil**	**Area (km2)**	**Rounds**	**PA**	**PS**	**Kj**
Coastal Rough Ground	Moor gate	6.89	13	0.0022	0.0066	0.3864
Farmland Medieval	Croft Pascoe	6.30	11	0.0020	0.0056	0.3916
Farmland Medieval	Neath	85.60	30	0.0268	0.0153	0.3831
Farmland Prehistoric	Moor gate	7.48	11	0.0023	0.0056	0.3879
Farmland C20	Manod	23.22	13	0.0073	0.0066	0.3887
Ornamental	Denbigh 2	14.06	10	0.0044	0.0051	0.3907
Farmland Prehistoric	Denbigh 2	8.61	9	0.0027	0.0046	0.3938
Totals		**804.03**	**441**	**0.2521**	**0.2253**	**0.3938**

All other combinations of soils and HLC Types make up the low probability zone.

11.1.1 Testing the rounds and enclosures/soils and HLC model

The rounds and enclosures/soils model was tested using events record data in the same way as the other models discussed so far (sections 8 and 10).

Testing with all 76 sites

Zone	Area km^2	SA	NS	Predicted PS	PS	Predicted Sites	Sites
High probability	23.54	0.93	22	0.65	0.67	49	51
Medium probability	11.80	0.55	6	0.19	0.14	15	11
Low probability	19.01	0.28	5	0.16	0.18	12	14
Totals	54.35		34			76	76

Table 50. Results of events record testing of the rounds and enclosures/soils and HLC model: test based on numbers of sites. NS = notional number of sites predicted.

Zone	Area km^2	Predicted PS	PS	Predicted site area	Site area
High probability	23.54	0.65	0.63	2.69	2.61
Medium probability	11.80	0.19	0.13	0.80	0.55
Low probability	19.01	0.16	0.24	0.65	0.99
Totals	54.35			4.15	4.15

Table 51. Results of events record testing of the rounds and enclosures/soils and HLC model: test based on site area.

The performance of the model when tested is very similar to that of the rounds and enclosures/soils model (tables 37 and 38). When tested using point data the high probability zone performs somewhat better than predicted, which validates the model. However, the low probability zone performs better than expected at the expense of the medium probability zone. When the test is based on site area the low probability zone performs better than expected at the expense of both high and medium probability zones.

Testing with the 33 new sites

Zone	Area km^2	SA	NS	Predicted PS	PS	Predicted Sites	Sites
High probability	23.54	0.93	22	0.65	0.76	21	25
Medium probability	11.80	0.55	6	0.19	0.06	6	2
Low probability	19.01	0.28	5	0.16	0.18	5	6
Totals	54.35		34			33	33

Table 52. Results of events record testing of the rounds and enclosures/soils and HLC model: test based on numbers of sites. NS = notional number of sites predicted.

When the test is carried out using only the new enclosures as the test sample the high probability zone performs significantly better than predicted (capturing 11% more sites than expected) at the expense of the medium probability zone. The low probability zone performs slightly better than predicted.

When the test is based on the area taken up by the new sites the high probability zone performs even better, capturing 81% of the enclosures at the expense of both medium and low probability zones. In terms of the high probability zone this model is certainly verified by the tests. The low probability zone performs more or less as expected, but the most noteworthy result is the much worse than predicted performance of the medium probability zone.

Zone	Area km^2	Predicted PS	PS	Predicted site area	Site area
High probability	23.54	0.65	0.81	1.09	1.35
Medium probability	11.80	0.19	0.05	0.32	0.09
Low probability	19.01	0.16	0.14	0.26	0.23
Totals	54.35			1.67	1.67

Table 53. Results of events record testing of the rounds and enclosures/soils and HLC model: test based on site area.

11.2 The rounds and enclosures/bedrock geology and HLC model

The bedrock geology and HLC Types polygons were combined by performing a spatial union in GIS and the resulting shapefile comprised 777 different combinations of rock and HLC Types. Rounds and enclosures were present in 185 of these combinations. A predictive model was created using the Kj parameter method and the result is shown in the table overleaf; a probability map based on the model is shown in Fig 27 and the performance of the model is summarised below.

Probability zone	PA	PS	Kvamme's gain	PS/PA	Sites per km^2
High	0.48	0.70	0.3111	1.4583	0.89
Medium	0.13	0.13	0	1.0	0.64
Low	0.39	0.16	-1.4375	0.4103	0.26

Although the Kvamme's gain produced by this model is less than that of the enclosures and rounds/soils and HLC model, its overall performance is better. It is more accurate, with 70% of the enclosures captured in the high probability zone (as opposed to 62% in the soils/HLC model); the medium probability zone is small (only 13% of lowland Cornwall) and is displaying a by chance distribution pattern (the medium probability zone is theoretically the zone of neutral probability) and the low probability zone covers a relatively large area at 39% of lowland Cornwall. One is half as likely again to encounter an enclosure in the high probability zone as in the medium zone, and 3.5 times more likely as in the low probability zone.

Of the 26 combinations of geology and HLC Types forming the high probability zone, 14 include the HLC Type Farmland Medieval, five contain Farmland C20 and two contain Farmland Prehistoric (these three Types form the high probability zone of the rounds and enclosures/HLC model). The other five combinations in the high zone include the HLC Type Farmland Post Medieval. The 11 highest ranked combinations all include either Farmland Medieval or Farmland Prehistoric. Similarly all the rock types which were highly ranked in the rounds and enclosures/bedrock model are represented in the high probability zone of this model. The one exception is Hornblende schist, whose highest rank in this model is 61st when combined with the HLC Type Farmland C20 in the low probability zone. The highest ranked rock type in this model is 'sandstone and [subequal/subordinate] argillaceous rocks, interbedded' when combined with Farmland Medieval. This rock type also appears in the high zone when combined with Farmland C20 (20th) and Farmland Post medieval (25th). Other rock types forming more than one high probability combination are granite (with Farmland Prehistoric, Medieval and C20), mudstone and sandstone (with Farmland Medieval, C20 and Post medieval), slate and siltstone (with Farmland Medieval, C20 and Post medieval), Hornfelsed slate and Hornfelsed siltstone (with Farmland Medieval and Farmland Post medieval) and slate and sandstone, interbedded (with Farmland Medieval and Farmland Post medieval).

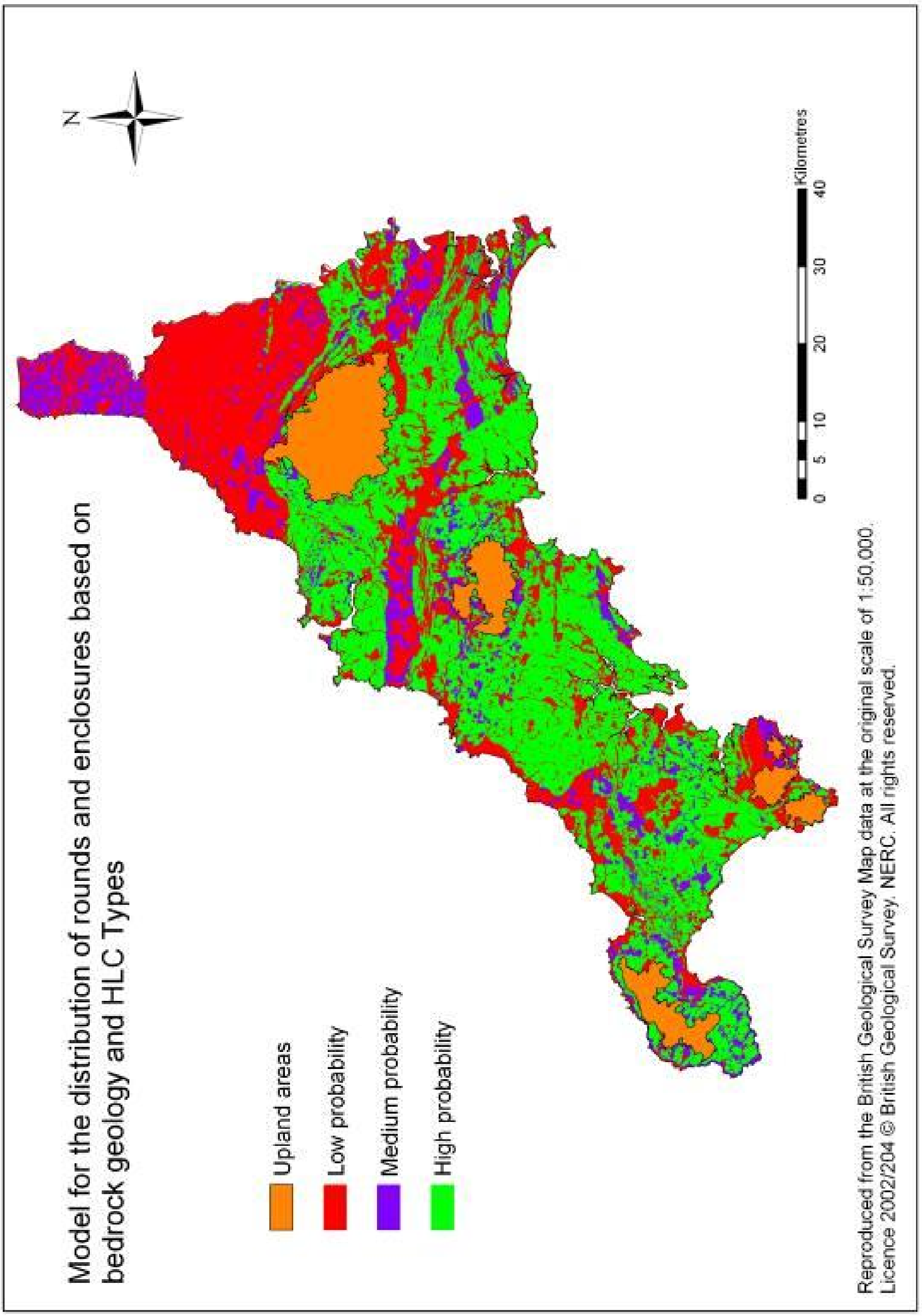

Fig 27. Map showing predictive model for rounds and enclosures using bedrock geology and HLC Types as variables.

Table 54. Model for the distribution of rounds and enclosures correlated with bedrock geology and HLC Types.

Rounds and enclosures/bedrock geology and HLC model. High probability zone						
HLC Type	**Bedrock**	**Area (km2)**	**Rounds**	**PA**	**PS**	**Kj**
Farmland Medieval	Sandstone and [subequal/subordinate] argillaceous rocks, interbedded	146.51	165	0.0459	0.0843	0.0569
Farmland Prehistoric	Granite	65.79	117	0.0206	0.0598	0.1057
Farmland Medieval	Slate and siltstone	280.05	196	0.0878	0.1002	0.1482
Farmland Medieval	Mudstone and sandstone	86.30	106	0.0271	0.0542	0.1869
Farmland Medieval	Granite	112.18	102	0.0352	0.0521	0.2167
Farmland Medieval	Hornfelsed slate and hornfelsed siltstone	69.81	66	0.0219	0.0337	0.2367
Farmland Medieval	Slate and sandstone, interbedded	74.76	67	0.0234	0.0342	0.2560
Farmland Medieval	Mudstone, siltstone and sandstone	33.07	50	0.0104	0.0255	0.2762
Farmland Medieval	Hornfelsed slate	17.39	32	0.0055	0.0164	0.2900
Farmland Medieval	Slaty mudstone with sedimentary rock, metamorphic rock and igneous rock clasts	14.54	29	0.0046	0.0148	0.3028
Farmland Medieval	Mudstone	29.45	33	0.0092	0.0169	0.3142
Farmland Post medieval	Mudstone and sandstone	65.36	41	0.0205	0.0210	0.3212
Farmland C20	Slate and siltstone	33.74	31	0.0106	0.0158	0.3303
Farmland C20	Granite	24.86	27	0.0078	0.0138	0.3394
Farmland Post medieval	Slate and siltstone	76.79	43	0.0241	0.0220	0.3445
Farmland C20	Slate, siltstone and sandstone	40.64	28	0.0127	0.0143	0.3501
Farmland Medieval	Hornfelsed slate and hornfelsed sandstone	36.03	24	0.0113	0.0123	0.3546
Farmland C20	Mudstone and sandstone	15.14	18	0.0047	0.0092	0.3611
Farmland Medieval	Slate, siltstone and sandstone	204.62	84	0.0641	0.0429	0.3551
Farmland C20	Sandstone and [subequal/subordinate] argillaceous rocks, interbedded	16.40	17	0.0051	0.0087	0.3607

Rounds and enclosures/bedrock geology and HLC model. High probability zone						
HLC Type	**Bedrock**	**Area (km2)**	**Rounds**	**PA**	**PS**	**Kj**
Farmland Post medieval	Slate and sandstone, interbedded	12.84	16	0.0040	0.0082	0.3667
Farmland Post medieval	Hornfelsed slate and hornfelsed siltstone	32.35	21	0.0101	0.0107	0.3702
Farmland Medieval	Peridotite and serpentinite	9.65	15	0.0030	0.0077	0.3765
Farmland Prehistoric	Aplitic microgranite	6.31	14	0.0020	0.0072	0.3831
Farmland Post medieval	Sandstone and [subequal/subordinate] argillaceous rocks, interbedded	25.80	18	0.0081	0.0092	0.3867
Farmland Medieval	Microgranite	13.69	15	0.0043	0.0077	0.3919
Totals		**1544.07**	**1375**	**0.484**	**0.7028**	**0.3919**

Rounds and enclosures/bedrock geology and HLC model. Medium probability zone						
HLC Type	**Bedrock**	**Area (km2)**	**Rounds**	**PA**	**PS**	**Kj**
Farmland Medieval	Sandstone	83.77	34	0.0263	0.0174	0.3885
Farmland Post medieval	Granite	37.99	20	0.0119	0.0102	0.3897
Coastal Rough Ground	Granite	2.40	11	0.0008	0.0056	0.3957
Farmland C20	Hornfelsed slate and hornfelsed siltstone	8.96	11	0.0028	0.0056	0.3999
Farmland Post medieval	Mudstone	3.51	9	0.0011	0.0046	0.4043
Farmland Medieval	Sandstone, siltstone and mudstone	51.36	20	0.0161	0.0102	0.4016
Farmland Prehistoric	Metabasalt	3.30	8	0.0010	0.0041	0.4056
Farmland Medieval	Gabbro	7.24	8	0.0023	0.0041	0.4084
Farmland C20	Mudstone	3.90	7	0.0012	0.0036	0.4115
Farmland Medieval	Basaltic-rock	4.39	7	0.0014	0.0036	0.4145
Upland Rough Ground	Granite	19.26	10	0.0060	0.0051	0.4150

Rounds and enclosures/bedrock geology and HLC model. Medium probability zone						
HLC Type	**Bedrock**	**Area (km2)**	**Rounds**	**PA**	**PS**	**Kj**
Farmland Prehistoric	Hornfelsed slate and hornfelsed siltstone	10.15	8	0.0032	0.0041	0.4170
Farmland Post medieval	Slate	24.25	11	0.0076	0.0056	0.4166
Farmland Medieval	Microgabbro	10.24	8	0.0032	0.0041	0.4185
Farmland Medieval	Basaltic lava	10.54	8	0.0033	0.0041	0.4203
Farmland C20	Slate	21.95	10	0.0069	0.0051	0.4200
Farmland Post medieval	Peridotite and serpentinite	3.13	6	0.0010	0.0031	0.4228
Farmland C20	Slate and sandstone, interbedded	14.62	8	0.0046	0.0041	0.4234
Farmland Medieval	Felsite	5.16	6	0.0016	0.0031	0.4256
Farmland Medieval	Chert	11.64	7	0.0037	0.0036	0.4264
Farmland Post medieval	Hornfelsed slate and hornfelsed sandstone	16.93	8	0.0053	0.0041	0.4263
Farmland C20	Slaty mudstone with sedimentary rock, metamorphic rock and igneous rock clasts	1.63	5	0.0005	0.0026	0.4290
Farmland C20	Microgabbro	1.67	5	0.0005	0.0026	0.4316
Farmland C20	Sandstone, siltstone and mudstone	24.44	9	0.0077	0.0046	0.4298
Plantation and Scrub	Granite	9.80	6	0.0031	0.0031	0.4306
Farmland Post medieval	Sandstone	10.79	6	0.0034	0.0031	0.4311
Farmland Medieval	Metamudstone and metasandstone	5.81	5	0.0018	0.0026	0.4325
Totals		**408.83**	**261**	**0.1283**	**0.1337**	**0.4325**

All other combinations of geology and HLC Types make up the low probability zone.

This model can be seen to represent an improvement on the rounds and enclosures/bedrock geology model (Fig 21) which, by comparison, appears broad brush and crude.

The probability map produced by this model differs from that produced by the rounds and enclosures/HLC model in that large areas in east Cornwall are classed as medium or low probability rather than high probability. There is also a very striking band of low and medium probability land running west–east from Breock Downs eastwards. Another difference is that the area around Camborne and Redruth (though not including the conurbation itself) is classed largely as high probability whereas in the rounds and enclosures/HLC model, much of this area is in the medium probability zone. Another interesting difference is that in this model much of the Lizard peninsula away from the Helford estuary is classed as low probability whereas in other models this area is placed in the high probability zone.

This model also differs significantly from the rounds and enclosures/soils and HLC model. In this model quite extensive areas in east Cornwall are classed as high probability, whilst the soils/HLC model characterises these areas as mainly medium probability with some parts classed as low probability and a scattering of parcels of high probability land. Another difference is that in the soils/HLC model the area immediately to the north of Bodmin Moor is included in the high probability zone; in this model the whole of the area north of Bodmin Moor is placed in the low probability zone. The area around Bude in the far north is also more solidly placed in the low and medium zones of this model than in the soils/HLC model. As mentioned above a large part of the Lizard peninsula is classed in this model as low probability whereas in the soils/HLC model much of the peninsula is included in the high probability zone. The final difference is that the bedrock/HLC model classes more of West Penwith as high probability than the soils/HLC model.

11.2.1 Testing the rounds and enclosures/bedrock and HLC model

The rounds and enclosures/bedrock model was tested using events record data in the same way as the other models discussed so far models (sections 8 and 10).

Testing with all 76 sites

Zone	Area km^2	SA	NS	Predicted PS	PS	Predicted Sites	Sites
High probability	29.22	0.89	26	0.73	0.74	56	56
Medium probability	7.91	0.64	5	0.14	0.08	11	6
Low probability	17.22	0.26	4	0.13	0.18	10	14
Totals	54.35		36			76	76

Table 55. Results of events record testing of the rounds and enclosures/bedrock and HLC model: test based on numbers of sites. NS = notional number of sites predicted.

In this test the high probability zone performs very much as predicted, but the low probability zone captures more enclosures than predicted at the expense of the medium probability zone.

When the test is based on site area rather than number of sites the performance of the medium probability zone is similar (with the PS value only half that expected), but the low probability zone performs much as predicted whilst the high probability zone performs considerably better than predicted (table 56).

Zone	Area km^2	Predicted PS	PS	Predicted site area	Site area
High probability	29.22	0.73	0.78	3.10	3.32
Medium probability	7.91	0.14	0.07	0.60	0.31
Low probability	17.22	0.13	0.14	0.53	0.61
Totals	54.35			4.24	4.24

Table 56. Results of events record testing of the rounds and enclosures/bedrock and HLC model: test based on site area.

Testing with the 33 new sites.

When the test is carried out using only the previously unrecorded enclosures, and based on numbers of sites, the model performs very much as predicted, with the high probability zone capturing 73% of the sites. The low probability zone captures 2% more enclosures than predicted at the expense of the medium zone.

Zone	Area km^2	SA	NS	Predicted PS	PS	Predicted Sites	Sites
High probability	29.22	0.89	26	0.73	0.73	24	24
Medium probability	7.91	0.64	5	0.14	0.12	5	4
Low probability	17.22	0.26	4	0.13	0.15	4	5
Totals	54.35		36			33	33

Table 57. Results of events record testing of the rounds and enclosures/bedrock and HLC model: test based on numbers of new sites. NS = notional number of sites predicted.

When the test is carried out based on site area the high probability zone performs significantly better than predicted, largely at the expense of the low probability zone.

Zone	Area km^2	Predicted PS	PS	Predicted site area	Site area
High probability	29.22	0.73	0.81	1.25	1.39
Medium probability	7.91	0.14	0.13	0.24	0.21
Low probability	17.22	0.13	0.06	0.22	0.11
Totals	54.35			1.71	1.71

Table 58. Results of events record testing of the rounds and enclosures/bedrock and HLC model: test based on new site area.

The apparent discrepancy in the performance of the low probability zone when measured by site area as opposed to numbers of sites (a PS value of 0.06 as opposed to 0.13) is due to the way the area polygons were defined (using present day field boundaries; see section 8.1). The mean size of the polygons defining new enclosures in the low probability zone is 0.7ha, compared with 3.6ha for the medium zone and 3.1ha for the high probability zone.

11.3 The rounds and enclosures/soils, bedrock and HLC model

The HLC model and the bedrock geology and soils model were combined to create a definitive HLC/geology and soils model. If the HLC model was combined with the soils/bedrock model using the spatial union tool in GIS the resulting shapefile would comprise a large number of combinations and a very large number of polygons and the ensuing model building would be extremely time-consuming. For this reason a simplified method was used to create this model. The three probability zones of both existing models were coded as H for the high, M for the medium and L for the low probability zones. A spatial union of the two models was carried out in ArcView based on the probability zone code. The codes were aggregated in the following way.

Bedrock/soils coding	HLC coding	Final codes
L	L	**LL**
L	M	**LM**
L	H	**LH**
M	L	**ML**
M	M	**MM**
M	H	**MH**
H	L	**HL**
H	M	**HM**
H	H	**HH**

Table 59. Code combinations for zones of intersection in the HLC/bedrock and soils model for rounds and enclosures

The model was then created using the Kj parameter. The results are shown in table 60 below.

Zone	Coding	Sites	PA	PS	Cum Kj
High	**HH**	916	0.24	0.47	0.3239
	MH	336	0.21	0.17	0.3441
Total		**1252**	**0.45**	**0.64**	**0.3441**
Medium	**LH**	299	0.20	0.15	0.3279
	HM	142	0.06	0.07	0.3589
Total		**441**	**0.26**	**0.22**	**0.3589**
Low	**MM**	76	0.05	0.04	0.3488
	HL	54	0.05	0.03	0.3296
	LM	68	0.07	0.03	0.2756
	LL	43	0.07	0.02	0.1747
	ML	23	0.04	0.01	0.0000
Total		**264**	**0.28**	**0.13**	**0.0000**

Table 60. The HLC/bedrock and soils model for rounds and enclosures

A case can be made for defining the high probability zone as only that area with a code of HH. This would provide a precise model, with the high probability zone covering just 24% of the project area and a relatively high Kvamme's gain (0.4788). However only 47% of sites would be captured in this zone so the best all-round model is that listed in table 9 above, in which the high probability zone includes the area coded MH as well as HH.

Overall performance of the model is summarised in table 61 below. Although the Kvamme's gain is reduced to 0.2893, 64% of the sites are captured in the high probability zone. The medium probability zone performs well (with PA and PS values being almost equal), although it is quite large, and the low probability zone (13% of sites in 28% of the project area) is both accurate and precise.

Probability zone	PA	PS	Kvamme's gain	PS/PA
High	0.46	0.64	0.2893	1.41
Medium	0.26	0.23	-0.1605	0.86
Low	0.28	0.13	-1.1039	0.48

An important consideration is that the relative contribution to the combined model made by each of the original models can be compared by analysing the code combinations within each probability zone. The three highest ranked categories (listed in table 60) are made up of the high probability zone (weighted value H) from the HLC model (HH, MH and LH). Where the high probability zone from the geology/soils model is combined with the low probability zone from the HLC model the combined category (HL) is only ranked sixth. This trend can be seen throughout the rankings; MH is ranked higher than HM, LM is ranked higher than ML, and so forth. This suggests that the HLC model is a more powerful indicator of probability than the geology/soils model.

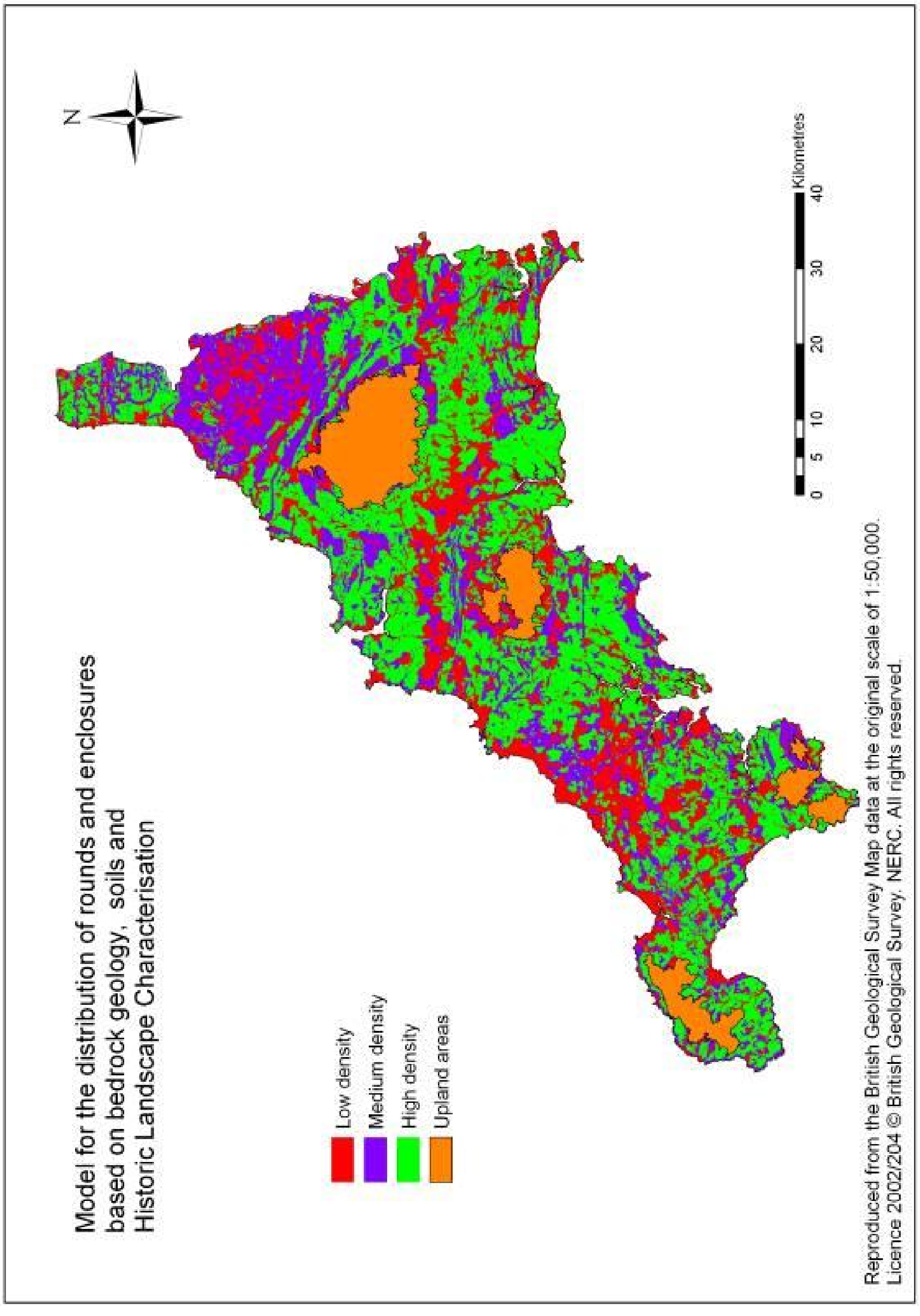

Fig 28. Map showing predictive model for rounds and enclosures using bedrock geology, soils and HLC Types as variables.

The probability map derived from this model offers a compromise between the HLC and bedrock/soils models. On the one hand the area of high probability covering 66% of the project area in the HLC model is reduced by 20% in extent; on the other the abrupt boundary between east Cornwall (largely medium and low probability) and central/west Cornwall (largely high probability zone) which is such a distinctive feature of the bedrock/soils model is to a large degree softened in this combined model.

11.3.1 Testing the rounds and enclosures/bedrock, soils and HLC model

The rounds and enclosures/bedrock, soils and HLC model was tested using events record data in the same way as the other models discussed so far (sections 8 and 10).

Testing with all 76 sites

Zone	Area km^2	SA	NS	Predicted PS	PS	Predicted Sites	Sites
High probability	22.6374	0.86	19.47	0.61	0.55	46	42
Medium probability	13.2111	0.53	7.00	0.22	0.22	17	17
Low probability	18.4676	0.29	5.36	0.17	0.22	13	17
Totals	54.3161		31.83			76	76

Table 61. Results of events record testing of the rounds and enclosures/bedrock, soils and HLC model: test based on numbers of sites. NS = notional number of sites predicted.

In this test the high probability zone does not perform as well as predicted because the low probability zone captures more enclosures than expected. The medium probability zone performs exactly as predicted.

When the test is based on site area rather than number of sites the medium probability zone and the low probability zone perform slightly better than predicted at the expense of the high probability zone. Overall, however, the high probability zone performs better than when the test is based on numbers of sites (table 62).

Zone	Area km^2	Predicted PS	PS	Predicted site area	Site area
High probability	22.6374	0.61	0.58	2.53	2.39
Medium probability	13.2111	0.22	0.24	0.91	0.99
Low probability	18.4676	0.17	0.18	0.70	0.76
Totals	54.3161			4.14	4.14

Table 62. Results of events record testing of the rounds and enclosures/bedrock, soils and HLC model: test based on site area.

Testing with the 33 new sites.

When the test is carried out using only the previously unrecorded enclosures, and based on numbers of sites, the model performs much as predicted, with the high probability zone capturing 58% of the sites. The medium probability zone captures 5% more enclosures than predicted at the expense of the low zone.

Zone	Area km²	SA	NS	Predicted PS	PS	Predicted Sites	Sites
High probability	22.6374	0.89	26	0.61	0.58	20	19
Medium probability	13.2111	0.64	5	0.22	0.27	7	9
Low probability	18.4676	0.26	4	0.17	0.15	6	5
Totals	54.3161		36			33	33

Table 63. Results of events record testing of the rounds and enclosures/bedrock, soils and HLC model: test based on numbers of new sites. NS = notional number of sites predicted.

When the test is carried out based on site area the high probability zone performs significantly better than predicted, largely at the expense of the low probability zone.

Zone	Area km²	Predicted PS	PS	Predicted site area	Site area
High probability	22.6374	0.61	0.69	1.03	1.16
Medium probability	13.2111	0.22	0.23	0.37	0.40
Low probability	18.4676	0.17	0.08	0.29	0.13
Totals	54.3161			1.69	1.69

Table 64. Results of events record testing of the rounds and enclosures/bedrock and HLC model: test based on new site area.

The apparent discrepancy in the performance of the low probability zone when measured by site area as opposed to numbers of sites (a PS value of 0.08 as opposed to 0.15) is due to the way the area polygons were defined (using present day field boundaries; see section 8.1). The mean size of the polygons defining new enclosures in the low probability zone is 1.4ha, compared with 3.2ha for the high probability zone.

11.4 Discussion

A major issue with the rounds and enclosures/HLC model discussed in volume 1 is its lack of precision, reflected in its Kvamme's gain of only 0.1715, due to the large area covered by the HLC Type Farmland Medieval. By combining soils and bedrock geology with HLC Types an increase in model precision was achieved. Models were made for soils/HLC, bedrock geology/HLC and soils and bedrock geology/HLC. All three models are more precise than the rounds/HLC model, producing Kvamme's gain measures of between 0.2893 and 0.3373. The high probability zones of these models are defined much more precisely – the HLC model's high probability zone covers 66% of lowland Cornwall; the high probability zones of the combined models cover between 41 and 48%. None, however, are as accurate – the HLC model high probability zone captured 79% of the enclosures whereas the high probability zones of these models capture between 62 and 70%.

Analysis of the ranking of the combinations of land classes in the soils/bedrock/HLC model suggests that HLC Types are the most influential element in determining the order in which the weightings are ranked. Nonetheless, and despite the reduced accuracy of these models, it is clear that when soil and geology types are used as variables in combination with HLC Types, more satisfactory models can be produced than by using HLC Types alone, due to their enhanced precision.

All three models were verified to a greater or lesser extent when tested using events record data. The most important test was made using previously unrecorded enclosures contained in the events dataset and the high probability zone of the soils/HLC model performed better than predicted when tested. The high probability zones of the bedrock/HLC and bedrock/soils/HLC models performed as predicted (or slightly better when the test was based on site area rather than number of sites captured). In the bedrock/HLC and bedrock/soils/HLC models the low probability zone performed worse than predicted (especially when the test was based on site area) but in the soils/HLC model this zone performed more or less as predicted whilst the medium zone performed poorly.

Combining HLC Types with soils and bedrock also enabled more nuanced models to be produced than by using soils and geology alone. This is clear when comparing the maps produced by these models. The models based on bedrock and soils alone are notably broad brush (Figs 21 and 23), but when combined with HLC Types a much finer granularity is achieved (Figs 26 and 27).

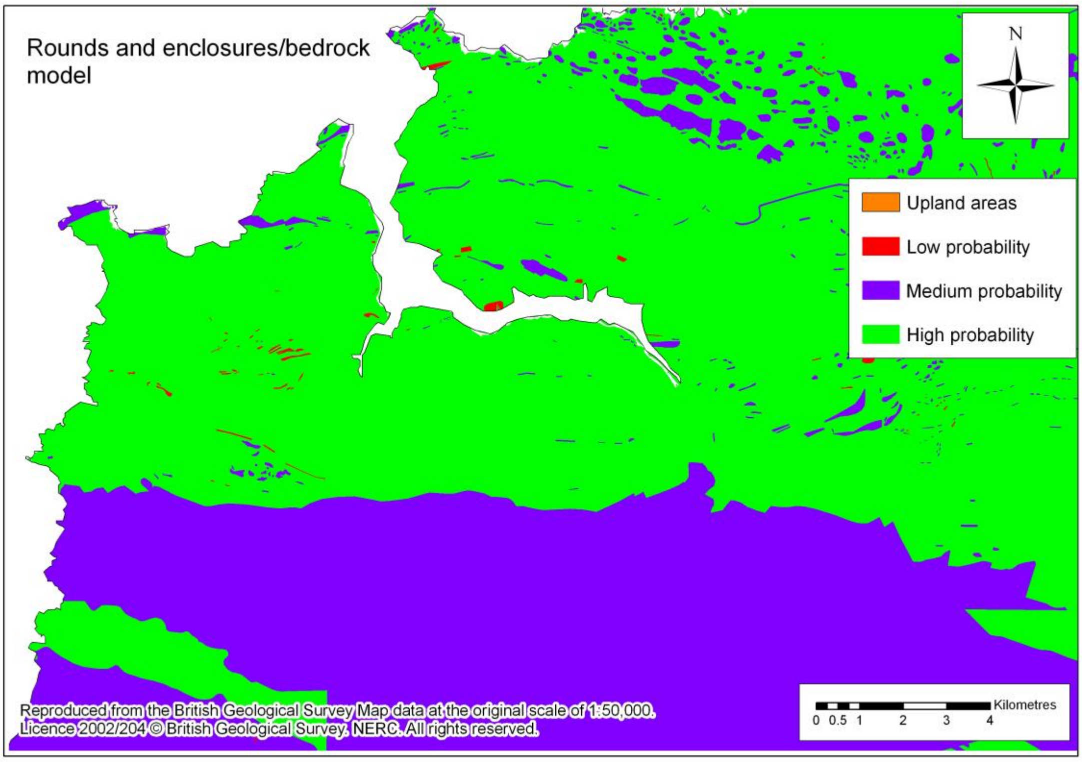

Fig 29. Map showing predictive model for rounds and enclosures using bedrock geology for the Camel Estuary area.

Fig 29, for instance shows a sample of the rounds and enclosures/bedrock model for the area around the Camel Estuary. Much of this area is classed as the zone of high probability, apart from an extensive west–east running band of land classed as medium probability towards the foot of the map. This area corresponds with the high ground of St Breock Downs which overlies Devonian sandstone, siltstone and mudstone of the Meadfoot Group, which is ranked 25th in the model. When HLC Types are combined with bedrock, much of St Breock Downs is reclassified as low probability, as are the river valleys, towns and other tracts of landscape (Fig 30). When soils are also added to the combination, more of St Breock Downs falls into the low probability zone and several tracts of land in the northern part of the estuary are classed as medium probability rather than high (Fig 31).

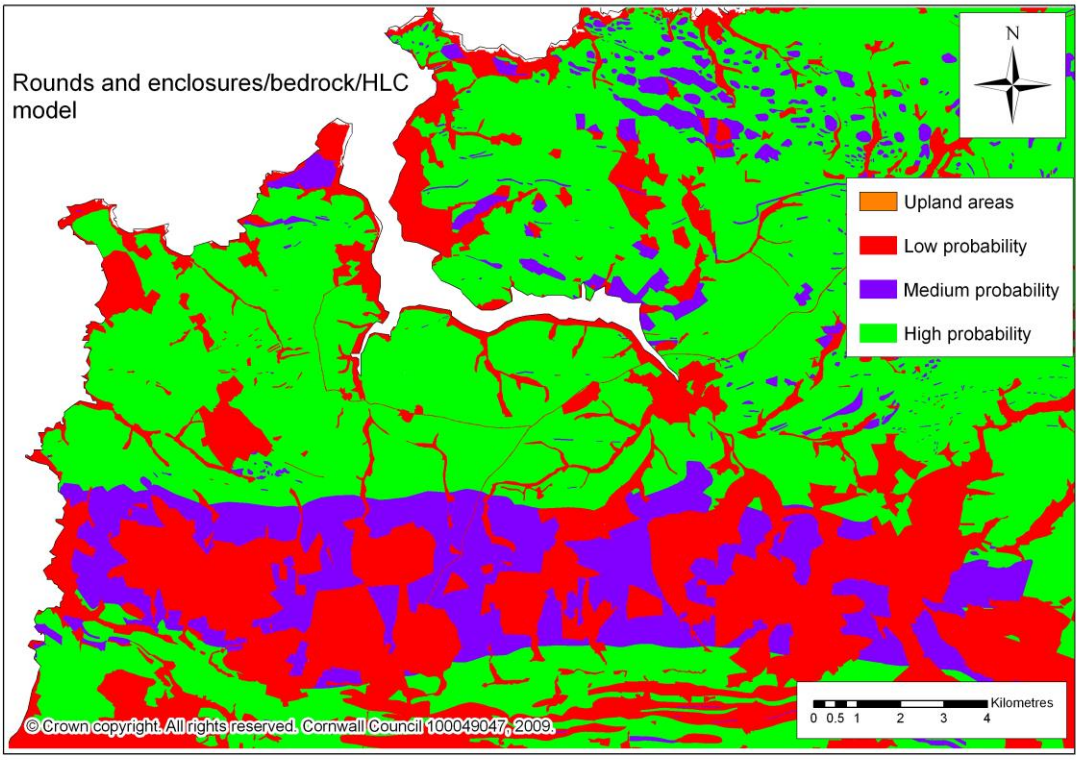

Fig 30. Map showing predictive model for rounds and enclosures using bedrock geology and HLC Types for the Camel Estuary area.

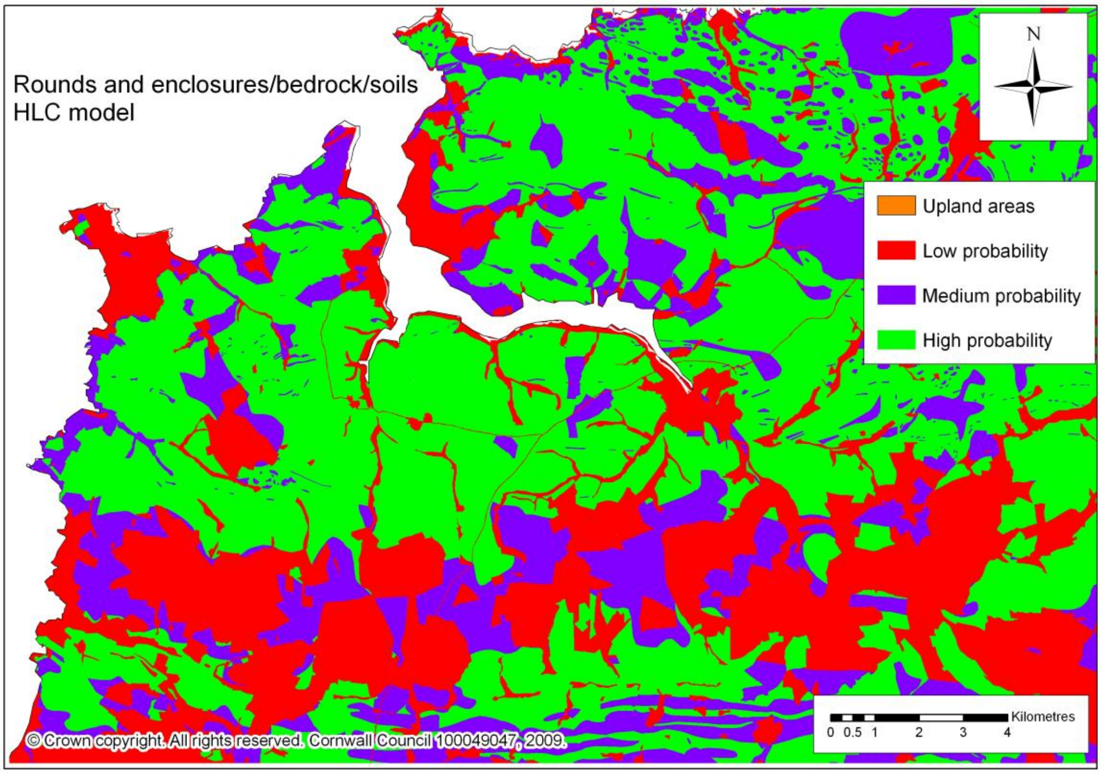

Fig 31. Map showing predictive model for rounds and enclosures using bedrock geology, soils and HLC Types for the Camel Estuary area.

This is less clear when analysing the relative performance of the high probability zones of the models, as summarised in table 65. The right hand side of this table shows the model performance when the land classes are combined with HLC. The soils and HLC model is more precise but less accurate than the model based on soils alone, and the same is true of the bedrock/HLC model compared with the model based on bedrock alone. The opposite is the case when HLC Types are combined with soils and bedrock.

Land class	PA	PS	Kvamme		PA	PS	Kvamme
Soils	49%	71%	0.3023	&HLC	41%	62%	0.3373
Bedrock	55%	74%	0.2534	&HLC	48%	70%	0.3111
Soils and bedrock	35%	57%	0.3860	&HLC	46%	64%	0.2893

Table 65. Performance of the high probability zones of models based on land classes alone compared with land classes combined with HLC Types.

The models based on combinations of land classes with HLC Types better reflect the distribution of known rounds and enclosures than the model based on HLC Types alone because they class much of eastern Cornwall in the low or medium probability zones to a greater or lesser degree. However, there are significant differences between the three models. The soils/HLC model classes southeast Cornwall as a combination of medium and low probability, with scattered tracts of high probability land. The geology/HLC and soils, geology/HLC models both characterise much of southeast Cornwall as high probability, although the area to the east of Bodmin Moor is characterised by extensive tracts of low and medium probability zone. The geology/HLC model classes all of northeast Cornwall (the area between Bodmin Moor and Bude) as low probability with some areas of medium probability around Bude and Morwenstow. The soils/HLC model classes the land immediately north of Bodmin Moor as high probability and the rest of northeast Cornwall as low and medium probability, with limited areas of high probability around Bude. The soils/bedrock/HLC model classes the Culm Measures as medium and low probability, but the area around Bude and Morwenstow as mainly high probability.

There are also differences in other parts of the county. For instance much of the Lizard peninsula is classed as low probability zone in the geology/HLC model: most of this area is classed as high probability in the soils/HLC model and as high and medium probability in the soils/bedrock/HLC model.

Given the differences and apparent contradictions between the models, how best can they be used and which is the most accurate or appropriate? The answer is that all the models have some veracity and they can be used in conjunction with each other. For example, Fig 32 shows a group of fields at Manor Farm, Goldsithney, in the parish of Perranuthnoe, west Cornwall. The fields in the centre of the map are classed in HLC as Farmland Medieval and are therefore in the high probability zone of the HLC model. These fields are bordered to the south by fields classed as Farmland Post medieval (with their typically straight boundaries) which are ranked in the medium probability zone, and to the north by the HLC Type Settlement C20, forming part of the low probability zone. Much of this area overlies the rock type Hornfelsed slate and Hornfelsed siltstone, which is ranked in the high probability zone of the rounds and bedrock model. The area is characterised by three soil types: in the south Trusham, in the north Denbigh 2 and in the east Manod. So this particular group of fields is captured in the high probability zones of the HLC and bedrock models and most of the fields are in the high probability zone of the soils model.

However, Hornfelsed slate and Hornfelsed siltstone overlying Trusham soils are ranked in the low probability zone of the soils and bedrock model. Where they overlie Denbigh 2 soils they are ranked in the high probability zone and where they overlie Manod soils they are ranked in the medium probability zone. Fig 33 shows how the same group of fields are classed in the soils and bedrock model; most of them lie in the low probability zone.

The HLC/bedrock and soils model (Fig 34) offers a compromise, with the fields now ranked in the medium or high probability zones.

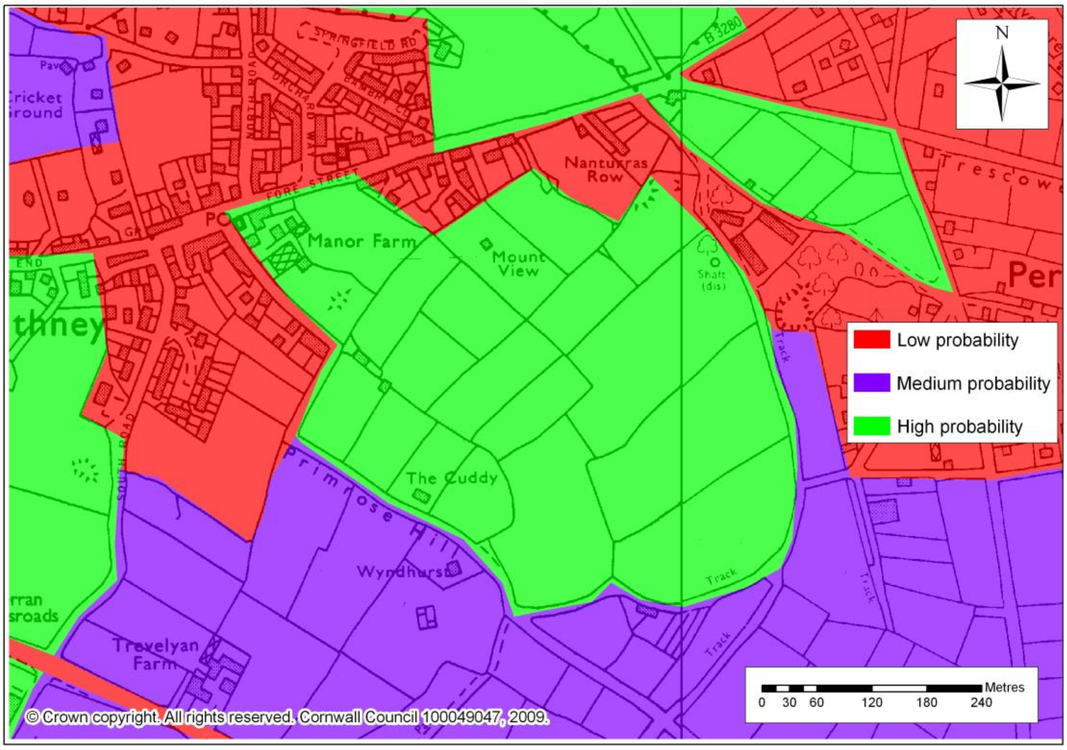

Fig 32. Map showing the rounds and enclosures/HLC model at Manor Farm, Goldsithney

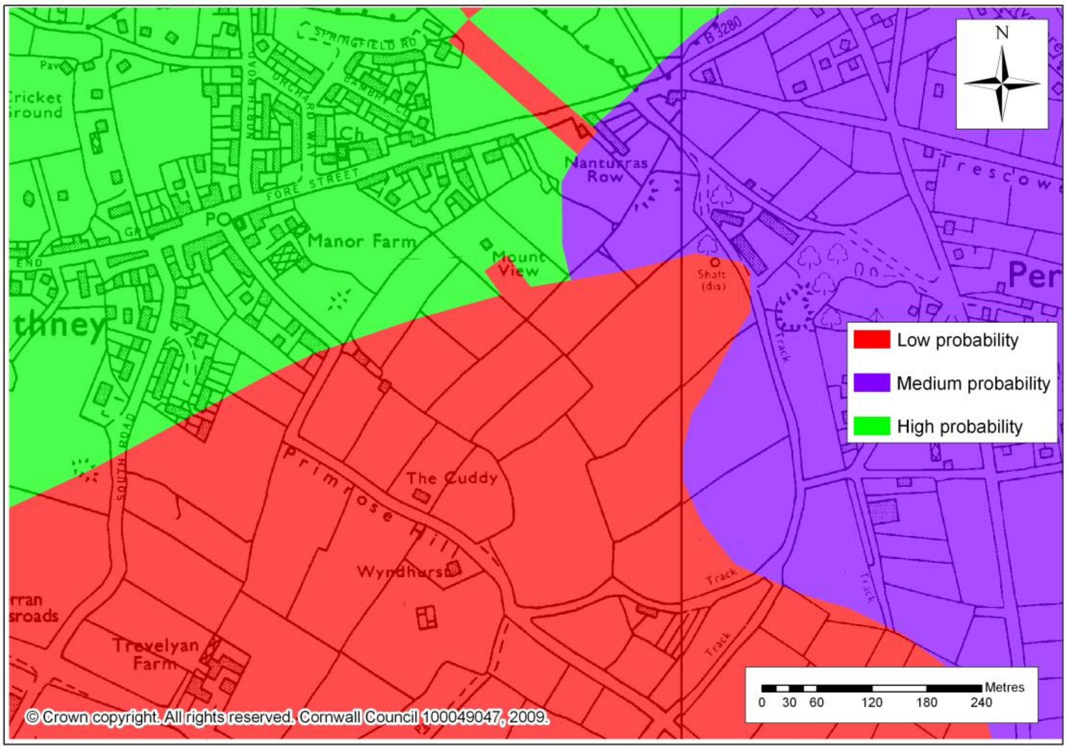

Fig 33. Map showing the rounds and enclosures/bedrock and soils model at Manor Farm, Goldsithney

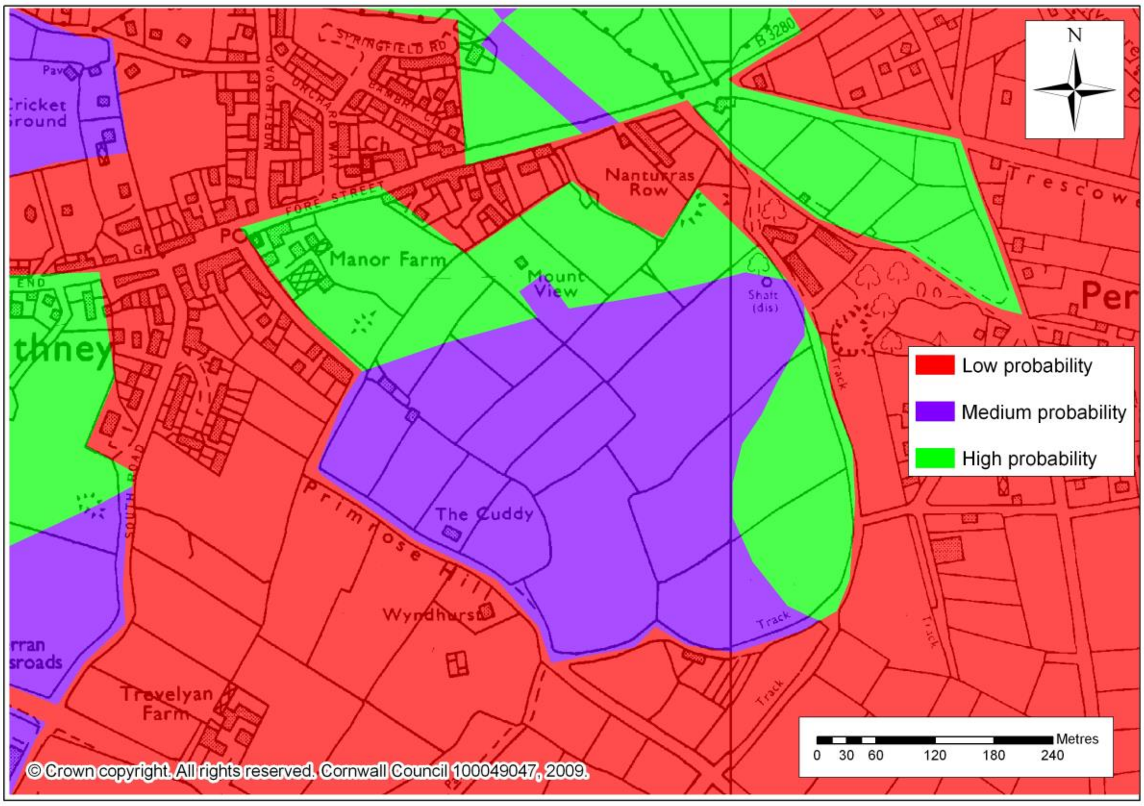

Fig 34. Map showing the rounds and enclosures/HLC, bedrock and soils model at Manor Farm, Goldsithney

To arrive at some measurement of the probability of there being below ground remains of enclosures in this group of fields a simple scoring system could be used. Table 66 shows the probability zone in which these fields (or most of them) fall for each of the seven models, allocating a value of 3 for the high probability zone, 2 for the medium and 1 for the low zone. The fields in question achieve an overall score of 17 out of a possible 21 maximum, which indicates a reasonably high probability.

Model	Zone
HLC	3
Soils	3
Bedrock	3
Soils and bedrock	1
Soils and HLC	3
Bedrock and HLC	2
Soils, bedrock and HLC	2
Score	**17**

Table 66. Probability scoring matrix for the Manor Farm fields, taking into consideration all seven predictive models for rounds and enclosures.

The scoring can be refined by more detailed analysis of the models. For instance there is an apparent contradiction in that the Manor Farm fields achieve a score of 3 in the soils model and the bedrock model but only 1 in the soils and bedrock combined model. As mentioned above the combination of Trusham soils overlying Hornfelsed slate and Hornfelsed siltstone is ranked in the low probability zone of the soils and bedrock model. However, it should be borne in mind that the Kj parameter calculation favours large areas over small. The combination of Trusham and Hornfelsed slate and Hornfelsed siltstone covers only 0.13% of lowland Cornwall (4.2km^2) and captures five enclosures. However these five enclosures constitute 0.26% of all enclosures in lowland Cornwall. In other words the PS value for this soil and bedrock combination is twice its PA value (0.26 as opposed to 0.13). We can conclude from this that there is high potential for the existence of buried prehistoric enclosures or other settlement features in the Manor Farm fields.

By way of comparison, the overall score for the fields to the south of Primrose Hill, Goldsithney (at the foot of the three map extracts) is 14 out of 21, indicating only a medium probability (table 67).

Model	Zone
HLC	2
Soils	3
Bedrock	3
Soils and bedrock	1
Soils and HLC	1
Bedrock and HLC	3
Soils, bedrock and HLC	1
Score	**14**

Table 67. Probability scoring matrix for the fields to the south of Primrose Hill, taking into consideration all seven predictive models for rounds and enclosures.

12 Rounds and enclosures and additional factors: summary and conclusions

The predictive models for cropmarks using bedrock geology and soil types as variables were rejected when tested by events record data. These models therefore appear to indicate that the distribution of known cropmark sites simply reflects the location of those soil and rock types most conducive to cropmark formation. The fact that the majority of rounds and enclosures in lowland Cornwall are recorded as cropmarks therefore raises the possibility that the same applies to their known distribution. To shed light on this question models were made based on the correlation of the distribution of rounds and enclosures with bedrock geology, soils and a combination of the two.

The models for rounds and enclosures using soil types as the variable and bedrock geology as the variable were both accurate (with 70% or more sites captured in the high probability zone) but not particularly precise (the high probability zones covering between 49 and 55% of lowland Cornwall). The model for rounds and enclosures using a combination of bedrock geology and soils as variables was less accurate but more precise.

Whilst there were clear cut off points between the three probability zones of all three models, both the rounds and enclosures/soils and the rounds and enclosures/bedrock and soils model can be regarded as essentially two-zone models because in each the medium and low zones are a similar size and capture a similar proportion of sites.

All three models are broadly similar in that east Cornwall (particularly northeast Cornwall) is placed in the low and medium probability zones and that the greater part of central and western areas are in the high probability zone. There are, however, differences between all three models in terms of the way the three zones are spread across the county.

Unlike the models for cropmarks, the high probability zones of all three of these models were verified when tested using the events record data, although the test for the rounds and enclosures/bedrock geology model was less conclusive than for the other two models. The medium and low probability zones performed less well in the tests, but their performance is less important than that of the high probability zone.

The fact that the models were largely verified by testing suggests that the influence of soils and geology on the location of rounds and enclosures is archaeologically meaningful and does not simply highlight those below ground conditions most favourable for cropmark production. This view is corroborated by analysis of the form of survival of enclosures in each of the probability zones of the three models. The proportion of extant enclosures (i.e. those with above ground earthwork survival) captured in each probability zone of each model is actually very similar to the overall proportion of enclosures captured in each probability zone of each model. In other words the high probability zones of these three models would remain the high probability zones if the cropmark enclosures were removed from the equation, and any bias in the models towards cropmark sites appears to be minimal.

Another significant factor is the varying density of enclosure distribution over soil or rock types that are essentially very similar, but which occur in different parts of the county. The clearest example of this is a comparison of the numbers of enclosures located on Denbigh 1 and Denbigh 2 soils (Fig 25); these two soil types are closely related and overlie similar rock types but where they are found in central and western areas far greater numbers of enclosures are recorded than where they occur in east Cornwall. The suggestion is that, for whatever reason, fewer rounds and enclosures were established in eastern parts of the county than were in the west and central areas.

13 References

13.1 Publications

National Soil Resources Institute 2004. Cranfield University

Verhagen, P. 2007. *Case Studies in Archaeological Predictive Modelling*, Lieden University Press, 120.

Young, A, 2009. *Lowland Cornwall: the Hidden Landscape Project Design.* Internal document for English Heritage

13.2 Web sites

Lowry, R, 2009. http://faculty.vassar.edu/lowry/ch8pt1.html

14 Project archive

The HE project number is **2009028**

The project's documentary, photographic and drawn archive is housed at the offices of Historic Environment, Cornwall Council, Kennall Building, Old County Hall, Station Road, Truro, TR1 3AY. The contents of this archive are as listed below:

1. A project file containing project correspondence and administration.
2. A digital file containing Excel tables, draft documents and notes held in the directory G:\TWE\Projects\Sites_L\ Lowland_Cornwall
3. GIS shapefiles and accompanying metadata are held in the directory: L:\Historic Environment (Data)\HE_Projects\Sites_L\Lowland_Cornwall_2009028\Final report

Appendix 1. Chi-Squared tests

1. Chi-Squared test for cropmarks correlated with Agricultural Land Classification (ALC)

ALC Code	Area	Sites	Expected	Chi-Squared
G3	2045.2161	1283	1137.74871	18.5435833
G4	596.029055	148	331.569504	101.631068
G2	278.679138	251	155.028522	59.411806
NONAGR	69.3863223	12	38.5994412	18.3300652
URBAN	54.9619948	10	30.5752231	13.8458452
G5	33.3669995	9	18.5619801	4.9257387
G3B	29.1210417	27	16.1999642	7.20006363
G3A	27.6814892	13	15.3991447	0.37378018
OTHER	19.7668768	2	10.9962652	7.36002503
Other (Exp < 5)	3.88393534	2	2.16062372	0.01194099
				231.62
5% Sig Chi-Sq value				**8.81**

2. Chi-Squared test for cropmarks correlated with soil types

Soil	Area km2	Cropmarks	Expected	Chi-sq
Denbigh 2	940.2046	700	518.469786	63.5586096
Powys	192.7165	238	106.272276	163.280529
Moretonhampstead	234.2176	159	129.157791	6.89511222
Trusham	126.7805	110	69.9122921	22.9862914
Denbigh 1	686.0884	221	378.33904	65.4322469
Sportsmans	58.7557	48	32.4004532	7.51056959
Neath	144.5191	61	79.6941291	4.38514691
Hallsworth 1	61.6072	41	33.9728946	1.4535179
Yeollandpark	33.2780	23	18.3509393	1.17780159
Manod	334.4618	59	184.436811	85.3104833
Croft Pascoe	16.1958	21	8.93106985	16.3092527
Moor gate	89.1753	21	49.1751462	16.1430911
Hafren	46.8072	15	25.8115297	4.52856439
Hallsworth 2	20.1676	13	11.1212934	0.31736761
Halstow	57.2024	12	31.5438959	12.108963
Sandwich	25.9102	6	14.2880133	4.80760783
Hexworthy	18.7320	2	10.329641	6.71687618
Laployd	23.5789	1	13.0024329	11.0793416
Conway	15.3577	0	8.46890499	8.46890499
Malvern	18.6966	0	10.3101199	10.3101199
Sea	14.6582	0	8.08317021	8.08317021
Other	30.6983	8	16.9283715	4.70900684
				525.57
5% Sig Chi-Sq value				**33.92**

3. Chi-Squared test for cropmarks correlated with bedrock geology

Bedrock	Area	Cropmarks	Expected	Chi-Sq
Slate and siltstone	489092213.4	328	267.161278	13.8543661
Slate, siltstone and sandstone	338230834.8	129	184.754898	16.8255818
Granite	294311320.1	160	160.764343	0.00363402
Mudstone and siltstone	286313858.1	93	156.395817	25.6978074
Sandstone and [subequal/subordinate] argillaceous rocks, interbedded	239587720.2	213	130.872174	51.5386848
Mudstone and sandstone	205346006.3	137	112.168012	5.49735706
Slate	175105234.5	48	95.6493211	23.7373122
Hornfelsed slate and hornfelsed siltstone	160587108	64	87.7189531	6.41353685
Sandstone, siltstone and mudstone	153210334.5	29	83.6894712	35.738525
Sandstone	143376026.6	69	78.317588	1.10853063
Slate and sandstone, interbedded	129899902.8	99	70.9564026	11.0834728
Hornfelsed slate and hornfelsed sandstone	77345127.13	22	42.2489306	9.70484187
Mudstone, siltstone and sandstone	43094770.65	80	23.5400476	135.417153
Mudstone	42165412.93	84	23.0323961	161.383501
Hornfelsed slate	38132202.72	8	20.8292991	7.90189408
Microgranite	35169525.35	8	19.210969	6.54239904
Metamudstone and metasandstone	33617399.15	6	18.363137	8.32358633
Mudstone and sandstone, interbedded	25810969.46	3	14.0989601	8.737305
Chert	24079860.37	9	13.1533606	1.31148267
Slaty mudstone with sedimentary rock, metamorphic rock and igneous rock clasts	21128673.17	24	11.5413069	13.4489999
Hornfelsed slate, hornfelsed siltstone and hornfelsed sandstone	21040578.75		11.4931862	11.4931862
Hornblende schist	20937031.92	34	11.4366249	44.5153967
Peridotite and serpentinite	19022963.17	21	10.3910858	10.831309

Bedrock	**Area**	**Cropmarks**	**Expected**	**Chi-Sq**
Tuff and agglomerate	18464241.55		10.0858902	10.0858902
Microgabbro	15250378.79	21	8.33035279	19.2692872
Basaltic lava	15160217.63	10	8.28110323	0.35678895
Metalimestone and pelite	13545068.53	3	7.39884569	2.61525165
Gabbro	12636941.64	9	6.9027913	0.63717475
Aplitic microgranite	12489452.25	5	6.82222683	0.48671947
Felsite	10778756.37	7	5.8877779	0.2101027
Other	79771238.79	32	30.8920628	0.03973593
				644.81
5% Sig Chi-Sq value				**56.94**

4. Chi-Squared test for rounds and enclosures correlated with bedrock geology

Bedrock	Area	Sites	Expected	Chi-Sq
Slate and siltstone	489.092213	285	303.238929	1.09701794
Slate, siltstone and sandstone	338.230835	123	209.704333	35.8487648
Granite	294.31132	298	182.474076	73.1404672
Mudstone and siltstone	286.313858	65	177.515621	71.316343
Sandstone and [subequal/subordinate] argillaceous rocks, interbedded	239.58772	210	148.545247	25.4244865
Mudstone and sandstone	205.346006	175	127.315262	17.859872
Slate	175.105234	53	108.565874	28.4395665
Hornfelsed slate and hornfelsed siltstone	160.587108	113	99.5645839	1.81299814
Sandstone, siltstone and mudstone	153.210334	43	94.9909578	28.4559684
Sandstone	143.376027	52	88.8936516	15.3120218
Slate and sandstone, interbedded	129.899903	104	80.5384064	6.83458237
Hornfelsed slate and hornfelsed sandstone	77.3451271	37	47.9542567	2.5022959
Mudstone, siltstone and sandstone	43.0947706	56	26.7189126	32.0889585
Mudstone	42.1654129	49	26.1427075	19.9847633
Hornfelsed slate	38.1322027	7	23.6421027	11.7146764
Microgranite	35.1695253	20	21.8052321	0.14945325
Metamudstone and metasandstone	33.6173991	16	20.8429082	1.12526333
Mudstone and sandstone, interbedded	25.8109695	4	16.0028938	9.00271296
Chert	24.0798604	10	14.9295999	1.62770306
Slaty mudstone with sedimentary rock, metamorphic rock and igneous rock clasts	21.1286732	41	13.0998533	59.4219013
Hornfelsed slate, hornfelsed siltstone and hornfelsed sandstone	21.0405787	3	13.0452344	7.73514153
Hornblende schist	20.9370319	33	12.981035	30.8726507
Peridotite and serpentinite	19.0229632	25	11.7943055	14.7859801
Tuff and agglomerate	18.4642415	3	11.4478961	6.23406675
Microgabbro	15.2503788	14	9.45528964	2.18442724
Basaltic lava	15.1602176	10	9.39938939	0.03837835
Metalimestone and pelite	13.5450685	4	8.39799115	2.30320869
Gabbro	12.6369416	16	7.83494922	8.50906016
Aplitic microgranite	12.4894523	17	7.74350527	11.0651045

Bedrock	**Area**	**Sites**	**Expected**	**Chi-Sq**
Felsite	10.7787564	7	6.68286767	0.01504936
Slate and limestone	9.379378	2	5.81524806	2.5030949
Metabasalt	8.15473619	9	5.05596573	3.07664393
Other	53.9197798	53	33.4304572	11.4556317
				543.94
5% Sig Chi-Sq value				**31.34**

5. Chi-Squared test for rounds and enclosures correlated with soil types

Soil	**Area km2**	**Enclosures**	**Expected**	**Chi-sq**
Denbigh 2	940.2046	751	576.830796	52.5889251
Powys	192.7165	183	118.234704	35.4764168
Moretonhampstead	234.2176	268	143.696303	107.528231
Trusham	126.7805	125	77.7818964	28.6641161
Denbigh 1	686.0884	252	420.926379	67.7936167
Sportsmans	58.7557	34	36.0475765	0.11630655
Neath	144.5191	42	88.6648156	24.5599678
Hallsworth 1	61.6072	15	37.797018	13.7498686
Yeollandpark	33.2780	17	20.4165936	0.5717463
Manod	334.4618	113	205.197748	41.4255263
Croft Pascoe	16.1958	23	9.93638641	17.1750567
Moor gate	89.1753	65	54.7104953	1.93516632
Hafren	46.8072	12	28.7169776	9.73143289
Hallsworth 2	20.1676	5	12.3731502	4.39365427
Halstow	57.2024	13	35.0946017	13.9101571
Sandwich	25.9102	12	15.8963286	0.95502409
Hexworthy	18.7320	9	11.4923863	0.54053087
Laployd	23.5789	5	14.4660382	6.19422388
Conway	15.3577	0	9.42219844	9.42219844
Malvern	18.6966	2	11.4706678	7.81938337
Sea	14.6582	1	8.99304383	7.10424089
Other	30.6983	10	18.8338959	4.14347184
				455.80
5% Sig Chi-Sq value				**33.92**