Geographic Information Systems I

David Wheatley

Between the lines: the role of GIS-based predictive modelling in the interpretation of extensive survey data

1 Introduction

Extensive field survey (such as fieldwalking and geophysical prospecting) yields information of a fundamentally different nature to intensive investigations. The resulting information lacks the spatial and chronological detail of intensive investigation such as excavations but generally provides greater spatial extent and chronological depth. As a result, field survey provides data with an appropriate spatial and chronological resolution to explore landscape and offsite approaches (e.g., Foley 1981) to archaeological explanation, and to focus attention on the long term effects of the activities of individual human agents within the entire landscape.

However, it is rarely possible within a field survey project to obtain total coverage by fieldwalking. The reasons for this are numerous but the most obvious is that fieldwalking is not generally appropriate unless the area is being actively ploughed, bringing artefacts to the surface. Areas of meadow or woodland within a generally agricultural landscape will therefore not be walked. The resulting data is therefore extensive but discontinuous, requiring the archaeologist to 'fill in the gaps', by estimating how the observed artefact densities might extend outside the surveyed areas.

One of the ideal tools for this kind of analysis is the Geographic Information System (GIS). This provides the archaeologist with the tools to visualise individual artefact types and to explore relationships within the survey data and between the survey data and other landscape indices. The application of GIS to archaeology has now been discussed by a number of authors (see e.g. papers in: Allen *et al*. 1990; Harris 1986; papers in: Lock/Stančič 1995; Wheatley 1993, 1995 for further details), therefore no general introduction to GIS will be provided here. Instead this paper will concentrate on how some of the methods which GIS makes possible might be of benefit in the interpretation of extensive survey data, in this case densities of lithic remains recovered during fieldwalking.

2 The Stonehenge Environs Project

The data used for the investigation is a subset of that collected under the direction of Julian Richards for the Historic Buildings and Monuments Commission for England (HBMC or English Heritage) between 1980 and 1986, generally referred to as the Stonehenge Environs Project (Richards 1990). This data was kindly made available in digital form by Wiltshire County Council Museums Service. The Stonehenge Environs surface collection database covers a series of irregular shaped areas to the east of the river Avon and immediately surrounding the Stonehenge monument itself. The total surveyed area is approximately 7.1 square kilometres, consisting of around 6500 50 m walked transects taken at 25 m intervals. In the course of the project a total of 102,175 pieces of worked flint were collected (Richards 1990: 15). The extent of the surveyed areas is shown in figure 1.

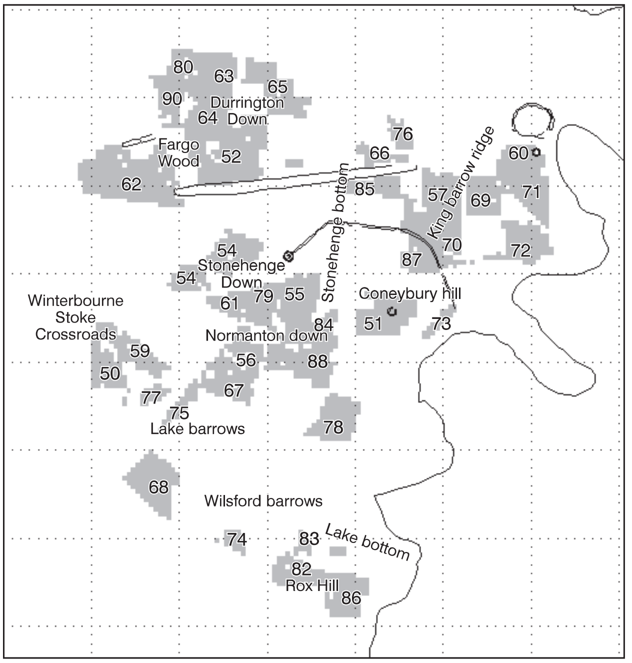

Figure 1. Area represented in the Stonehenge environs database, showing most of the original fieldwalking area numbers (from Richards 1990) and locations mentioned in the text. The grid on this and all subsequent images is the OS National grid, 1 km intervals.

Figure 2. Mapped distributions of total flint densities (left) and the mean-filtered flint densities (right).

Four categories of lithic information (counts of flint flakes, flint cores, burnt flint and retouched pieces) were uniformly present in the database together with total flint density. No details of 'type fossils' (arrowheads, scrapers or other tools) were sought, as the aim of the study was to extract as much information as possible from the majority of the flint data, rather than the minority.

Inspection of the lithic densities revealed a level of random 'noise' in each of the images. This is to be expected in a sample of this type, and represents the chance effects of obtaining unrepresentatively high or low values in the samples for some land units. To reduce the effects of the noise within the images, a second series of data themes were generated by applying a 3×3 mean filter to the images. In order to avoid the distortion which would arise due to edge effects in this case, a basemap of only those cells which had been sampled was used to force the filter to regard cells outside the basemap as 'no data' rather than zero values. Figure 2 shows the density of all flint within the study area, together with the mean filtered version.

3 Previous work

No computer facilities were available for this type of analysis as part of the Stonehenge Environs Project, and the original analysis took the form of plotting the densities of flint categories in categories derived from a frequency histogram of the flint densities (Richards 1990: 16). Using the procedure advocated by Hodder and Orton (1976), the inflexions of this distribution were used to estimate classes for plotting as distribution maps using different sized symbols for the different classes. These distribution maps then formed the basis of the interpretation of the lithic data and the main findings were detailed by Richards (1990) — these are not discussed in detail here for brevity.

Some analysis of the flint density data was undertaken by Maskell (1993), who entered a subset of the data from the paper record to a database, and then undertook some analyses with the IDRISI GIS. Maskell's main conclusions were that the flint density data exhibited spatial autocorrelation (high values in one location made it more likely that high values would occur in neighbouring areas) and that there may be a relationship between aspect and the presence of some categories of flint artefacts.

4 Values at unsampled locations: prediction from correlates

The fieldwalking data can be mapped and analysed in its 'raw' form, but its discontinuous nature makes it difficult to estimate how the lithic densities varied in the unsurveyed areas. One approach to 'guessing' the lithic densities at these unsurveyed locations would be to use spatial interpolation techniques which rely on the spatial structure of the artefact densities themselves, and to assume that the pattern which is observed within the survey data persists outside the surveyed areas. Techniques for interpolating surfaces from point data include polynomial trend surface analysis, linear and non-linear contouring, topographic interpolation with splines, inverse distance weighting and optimal interpolation methods such as 'Kriging'. Interesting though these are, however, each has problems when applied to data of this type, and none of these methods are the subject of this paper. Instead, it is the aim here to examine the extent to which *predictive modelling* techniques can be used to perform the same task, and whether these offer any advantages over spatial interpolation in this context.

In contrast to spatial interpolation, a predictive model tries to estimate values at unknown locations by making use of the correlations of the observed variable with other spatial variables. In situations where the values of these correlates are known for unknown locations, then this may be used to predict the values of the observed variable outside the sampled locations.

Predictive modelling was primarily developed within the context of North American archaeology (see e.g., Carmichael 1990; Kohler/Parker 1986; Kvamme 1983b, 1985a, 1985b, 1988, 1990; Sebastian/Judge 1988; Warren 1990a, 1990b for examples) initially to aid the management of an extensive, and only partially known archaeological resource. In recent years some of the simpler types of predictive models have been imported for cultural resource management use within European archaeology (e.g., Brandt *et al*. 1992; Van Leusen 1993). These have generally been rule-based models, implemented with map-algebra techniques. The map-algebra expressions which describe the predictions have either been entirely deductive in nature or inductive only to the extent that the weightings for the equations were derived from ratios of expected to observed numbers of sites in given classes of the predictors.

Far more promising for the interpretation of complex archaeological patterns and remains are predictive models which are based on multiple regression techniques in which the relationship between the archaeological dependent variable and the predictors (independent variables) is automatically obtained from multivariate statistical computation. Such an approach has many advantages, particularly that regression procedures provide estimates of the influence each supposed predictor has on the result, and of the total extent of the variability within the result which is accounted for by all the independent variables (Shennan 1988).

5 Multiple regression

Initially, a straightforward multiple regression will be attempted, treating total lithic density as the dependent variable, and a variety of possible correlates as independent variables. Filtered values of lithic density will be used as the dependent variable because it is held that these more closely represent the true form of the population from which the samples were drawn. Densities of retouched, burnt flint and cores provide rather low ranges of values which are therefore more prone to the effects of chance in sampling. The difference between the total flint densities and total flake densities is minimal, and therefore the experiment attempted to predict the total flint densities.

Two types of independent variables were selected as possible correlates: environmental indices and indices describing the relationship to cultural features of the landscape. Environmental variables employed described the topography (elevation, slope, aspect), soil class, geological substrate and proximity to water sources. These seemed likely to have influenced the densities of lithics either through the choices of the human individuals who dropped the flints, or through differential recovery rates.

Three cultural indices were also included in the analysis in the belief that the location of ceremonial monuments may have had considerable influence on the activities of the people responsible for leaving the lithic debitage, and consequently on the form of the scatters. One of the criticisms raised of predictive models (by e.g., Wheatley 1993), has been that of environmental determinism. The most obvious feature of the cultural landscape is the existence of earthen burial monuments, in the form of earlier Neolithic long mounds and later round mounds. There is now some evidence (Lock/Harris forthcoming; Wheatley 1995a) that the locations of these monuments has had some influence on the activities of contemporary and later generations of people. Three variables intended to reflect cultural aspects of the landscape were introduced as possible predictors of lithic density. These were density of round barrows, distance to long barrows and visibility of long barrows.

It is not the intention to devote much space to the details of multiple regression analysis here. However, it should be pointed out that one of the requirements of linear multiple regression analysis is that the dependent and independent variables are approximately normally distributed. If they are not, then appropriate transformations such as logarithmic or square root transforms must be applied to generate new, normally distributed variables (Shennan 1988). Frequency distributions for all the possible dependent variables were examined. Each of the histograms reveals the skewed nature of the flint densities. The transformed distributions for total flint and flake densities, however, are markedly different and although both exhibit a minor skew to the right and a suggestion of bimodality the distributions are apparently quite close to normal (fig. 3).

Frequency distributions of the independent variables were also examined, and appropriate transforms applied to obtain approximately normally distributed variables. Although

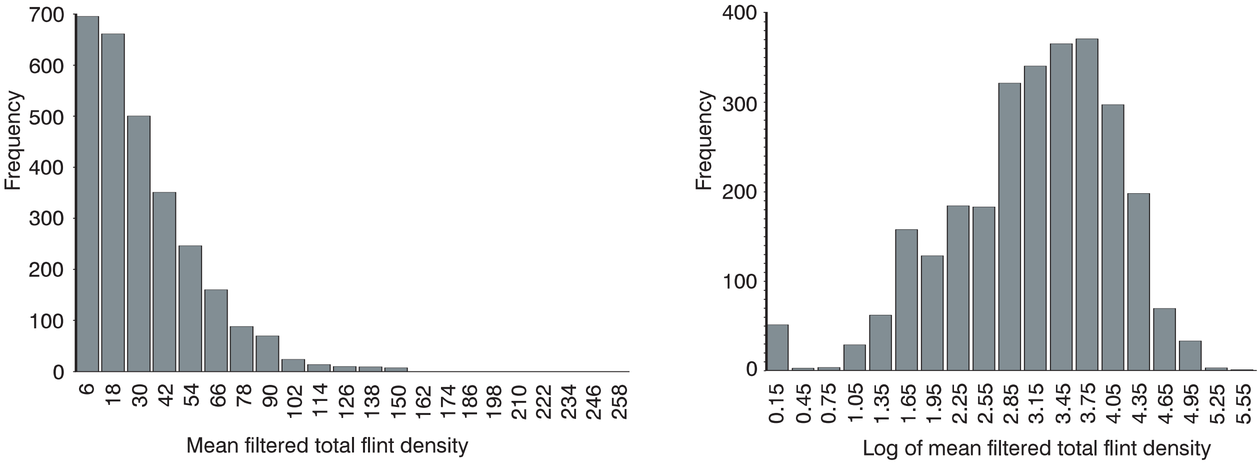

Figure 3. Frequency distributions of mean filtered flint density (left) and the ln mean filtered flint classes (right) showing some skewness and possible bimodality.

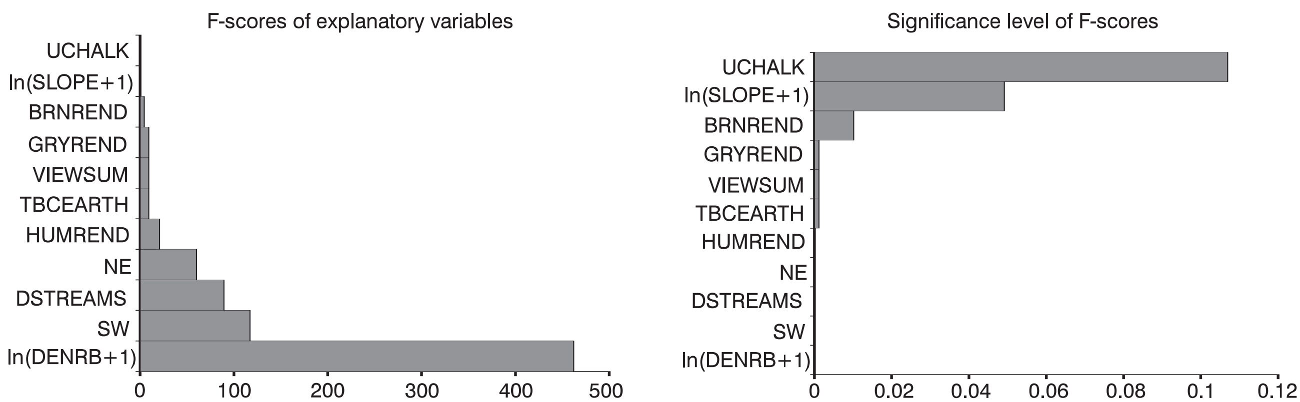

Figure 4. Explanatory power of the variables remaining in the regression expressed as F-scores (left) increasing with explanatory power, and significance of F-scores (right) reducing with explanatory power. Note the dominance of the 'density of round barrows' index.

some care was taken to approximate normality, it should be noted that results were not ideal, particularly in respect of skewness within some of the predictors. Consequently there must remain a suggestion that this may have biased the result a little. However, although not recorded here in detail, regressions omitting the most doubtfully normal variables did not produce significantly different results which supports the notion that the independent variables' skewness was not influencing the general result. Categorical variables (geology and soil class) were dissembled into dummy variables, and introduced into the regression. Further details of individual predictors are not given here for brevity but may be found fully described in Wheatley 1995a.

5.1 Application of the model

A stepwise linear multiple regression was then undertaken, a summary of the results is given in table 1. This procedure examines each variable as it is included in the model to identify to what extent it contributes to the model. Independent variables were only included in the regression if they proved significant at the 0.15 level. The flint densities show significant correlations with 11 of the variables, which in reality represent 7 of the selected variables once the dummy variables for geology and aspect have been counted out.

The partial correlation coefficients (r^2) and F scores for each variable (fig. 4) show that the density of round barrows index has the greatest explanatory power within the

Table 1. Summary of stepwise multiple regression model for total flint density.

Variable	Parameter Estimate	Standard Error	Type II Sum of Squares	F	Prob>F
INTERCEP	4.48422092	0.13117657	755.88447063	1168.59	0.0001
LOGSLOPE	-0.07334661	0.03547191	2.76557247	4.28	0.0388
DSTREAMS	-0.03190818	0.00284291	81.48413338	125.97	0.0001
VIEWSUM	0.01457195	0.00420610	7.76370920	12.00	0.0005
LOGDENRB	-0.48342903	0.03451878	126.86697985	196.13	0.0001
GRYREND	0.41090168	0.08041348	16.88931830	26.11	0.0001
BRNREND	0.16700804	0.03942703	11.60594984	17.94	0.0001
TBCEARTH	0.44753602	0.07039656	26.14250292	40.42	0.0001
UCHALK	0.14062157	0.08714878	1.68412841	2.60	0.1067
NE	0.33501178	0.04170020	41.74820486	64.54	0.0001
SW	-0.36472949	0.03930261	55.70483453	86.12	0.0001
Bounds on condition number:		1.852364,	134.823		

All variables left in the model are significant at the 0.1500 level.
No other variable met the 0.1500 significance level for entry into the model.

Summary of Stepwise Procedure for Dependent Variable LOGMTOT.

Step	Variable Entered	Variable Removed	Number In	Partial R**2	Model R**2	C(p)	F	Prob>F
1	LOGDENRB		1	0.1524	0.1524	348.0545	464.8721	0.0001
2	SW		2	0.0373	0.1897	221.2244	118.7974	0.0001
3	DSTREAMS		3	0.0277	0.2174	127.3268	91.5275	0.0001
4	NE		4	0.0182	0.2356	66.4410	61.4242	0.0001
5	HUMREND		5	0.0067	0.2423	45.3233	22.7708	0.0001
6	TBCEARTH		6	0.0033	0.2456	35.7845	11.4115	0.0007
7	VIEWSUM		7	0.0032	0.2488	26.6531	11.0514	0.0009
8	GRYREND		8	0.0032	0.2521	17.4324	11.1841	0.0008
9	BRNREND		9	0.0019	0.2540	12.9419	6.4831	0.0109
10	LOGSLOPE		10	0.0011	0.2551	11.0785	3.8633	0.0495
11		HUMREND	9	0.0005	0.2546	10.7186	1.6400	0.2004
12	UCHALK		10	0.0008	0.2554	10.1158	2.6036	0.1067

model, alone accounting for some 15% of the variation in lithic density while the aspect variables, distance to stream courses, soil indices and visibility of long barrows also carry limited explanatory value. Surprisingly perhaps, slope accounts for very little of the variation in lithic density (around 1%) and elevation even less.

Probably the most significant statistic within the result however, is the model correlation coefficient r^2 of 0.2554, or around 25% for all 7 explanatory variables. This indicates that the independent variables can account for only around one quarter of the variation within the lithic density, leaving three-quarters of the variation unexplained by the model. This is a very low value for r^2 from a regression analysis of this type, but a number of experimental modifications to the analysis through exclusions of independent variables and cases failed to obtain a value higher than around 27% and as none of these minor experimental alterations improved the methodological rigour of the analysis they are not discussed here in detail.

Accepting the low correlation coefficient for the time being, the model may now be used to express the predicted lithic density at all locations in the landscape, and then be translated into a prediction map (fig. 5 left) through a map-algebra operation. This provides mapped predictions of lithic densities for all locations in the study area.

5.2 Discussion of the model

The overall character of the prediction shows that the model generally predicts rather low values for flint throughout the sampled area, with the prediction never

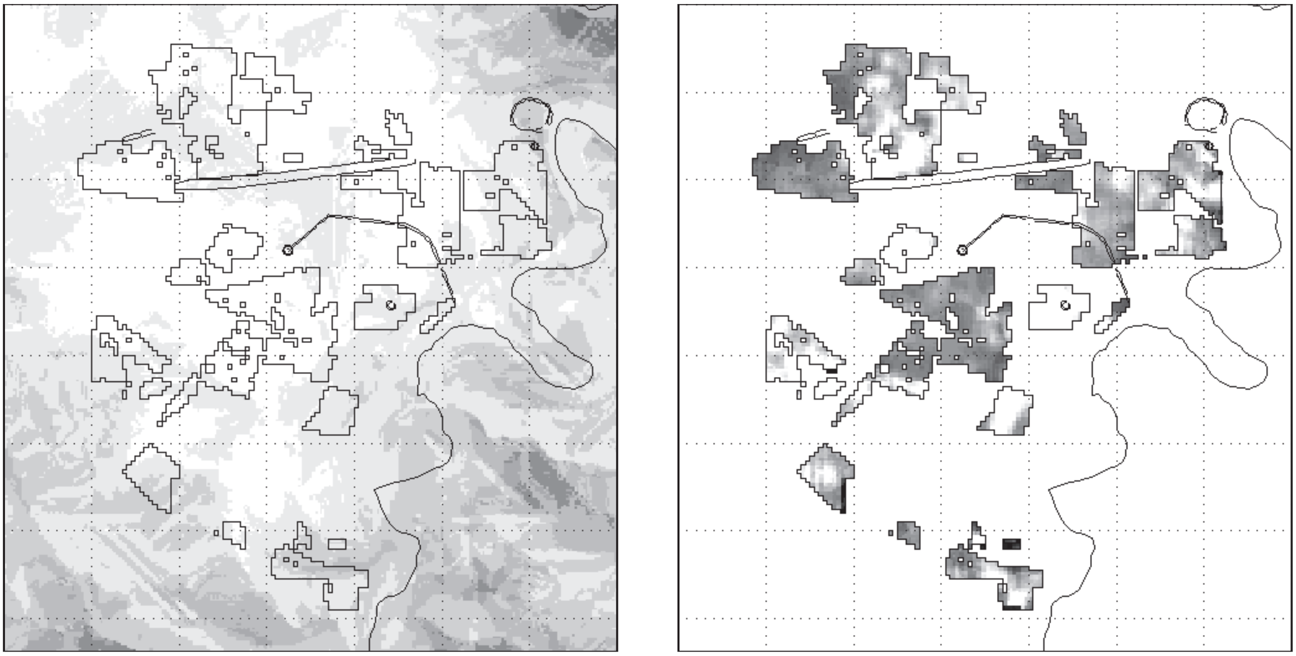

Figure 5. Mapped result of the multiple regression analysis (left) and the residuals from the regression (right).

reaching more than around 80-90 within the convex hull of the data points. The model seems to provide a fairly close approximation of the observed densities for the region around Normanton Down, and shows slightly higher predictions (although never approaching the true densities) for the surrounding areas. The areas of high density north of the Cursus, around Coneybury Henge and north of Stonehenge Down also show up as marginally higher values although, again, these do not approach the real densities for these areas.

One of the most obvious features of the model are the high values which are predicted for many areas away from the sampled region, for example in the northeast and southwest corners of the image. It may be that the variable most influential in the regression (density of long barrows) is the main factor. It is, for example, a measure which may be particularly prone to edge effects: the model predicts high flint densities at low round barrow densities, and the edge areas of the map may underestimate that index. It is also possible, however, that the model is actually rather 'over-trained' in regard to this variable. The variation of lithics with round barrow density may, at the scale of the sampled area, be as described in the regression equation, but at a larger scale be entirely the reverse. In other words, within the general cluster of monuments on Salisbury Plain the density of lithics is indeed higher in the gaps between the monuments, but nevertheless the flint densities at the edge of the study area might also be expected to decline rather than increase with density of monuments. The high predictions away from the core of the study area seem best explained in this way, therefore, and, put simply, it is that the model has no experience of lithic densities at the edge and extrapolates unreasonably high densities as a result.

The main conclusion which may therefore be deduced from the experiment is that while some of the 'shape' of the distribution is modelled within the core of the study area, this is clearly a poor model in the sense that it fails to adequately account for a large proportion of the variation within the dependent variable and in that it cannot be used outside the scope of the area for which it was designed.

The residuals of the regression were obtained within the GIS by subtracting the result of the application of the regression equation from the original observations in the form of the mean filtered density map. The result is shown in figure 5 (right), coded to show areas for which the model underpredicts and areas for which the model overpredicts. This supports the interpretation presented above, that the model fits adequately the areas which exhibit a low-level and low variation of lithic density. Thus the Normanton Down areas are a good fit, as is the peripheral area southwest of Fargo Wood (62). Practically all the very high areas of flint appear as high residuals. It seems, therefore, that the model explains the regular and therefore predictable variation within the lithic values which might be termed 'background' variation. The areas which show high residuals must then be interpreted as 'unpredictable' areas which deviate dramatically from the trend.

It is worth making the point that the lack of success of this approach should not be taken to devalue the experiment itself. The failure of the model to adequately predict absolute values for flint densities with the available independent variables is, of itself, an interesting finding.

6 Logistic multiple regression models

One reason for the lack of success of this model may be the choice of the dependent variable: total flint density at any given point in the landscape may not be closely related to the chosen independent variables because it is archaeologically the result of a compound of different activities involving manufacture, use and discard of flint artefacts.

However, most of the other flint variables do not show sufficient range of variation to be used as dependent variables in linear multiple regression analysis. Logistic regression, unlike linear multiple regression, can be used to predict presence/absence of particular classes rather than interval level values. Using logistic approaches, it is therefore possible to turn to the flint classes with low variability, and overall low values, such as core and retouched pieces density, and to define some characteristic of these variables which may be worth predicting.

6.1 Core: retouch ratios

It is possible to postulate, as a broad generalisation that those areas with high levels of cores in comparison with retouched pieces generally represent areas in which *manufacturing-related* activity dominated. Conversely, that areas with high levels of retouched artefacts represent areas in which *discard* activity was more common than procurement or manufacture. These might be termed 'discard areas' without prejudicing any interpretation of such activity (the discard could be structured symbolically or domestic and functional for example). It follows that if the proportion of cores to retouched pieces is spatially patterned, this might be evidence for persistent use of the landscape for different activities. If there is no evidence of differentiation between the distributions of cores and retouched pieces then it may be that flint was manufactured, used and discarded generally in the same places.

This ratio of cores to retouched pieces is biased towards the distribution of cores (the maximum density of cores is 39 pieces per 50 m^2, while the maximum count for retouched pieces is just 13), but an unbiased variable can be obtained by dividing the counts for each variable by these maxima. The result varies around 0 for areas with similar 'normalised' core and flake densities, is positive below 1 for areas with higher core values than retouch, and negative below -1 for areas with higher values for retouch than core. Positive values can roughly be equated with 'manufacturing areas' and negative ones with 'discard areas'. In practice the index varies between about -0.5 for 'discard areas' and 0.5 for 'manufacture areas'.

Mapping this variable (fig. 6 left) and mean-filtering it as above (fig. 6 right) clearly confirms an area of high core/retouch ratio north of the cursus, and also several clusters of low value. It is obviously tempting to interpret the latter as 'domestic' sites, but it should be remembered that what is actually revealed are areas in which, over an extended period of time, more discard activity took place than manufacturing. They may be domestic sites of some type, but equally they may represent deliberate discard of artefacts at these places as part of ritual or ceremonial activities. The high values around the Avenue are particularly suggestive in this respect.

6.2 Use as dependent variables

While the distribution of total flint densities may be the result of many different activities (each with different correlates), the core/retouch variable can be used for the differentiation of these activities into separate variables. Consequently it may be possible to use this variable for predicting which activities might be expected in which parts of the landscape.

To this end two normalised core/retouch ratio thresholds were set which allowed the production of binary maps indicating those areas which seem to have been used for manufacturing activity (fig. 7 left) and those which seem to have been used for discard (fig. 7 right). The aim of setting the thresholds was to obtain two variables which provided a clear distinction between these areas. After a little experimentation, values of 0.9 or less were classed as discard areas, while areas of 1.1 or above were classed as manufacture areas.

6.3 Application of the models

The same independent variables were adopted for the logistic regressions as for the linear multiple regression experiment, the only modification to the procedure for the linear multivariate model was that all variables were used untransformed because the logistic procedure does not require normally distributed variables (Rose/Altschul 1988). The output from the LOGISTIC procedure provides a summary of the results of the regression and gives the parameter estimates for those independent variables remaining in the model (tables 2, 3).

These intercept and parameter estimates were then used to generate an estimate of the *probability* of the event represented by the dependent variable for all locations

Figure 6. ‘Normalised’ core:retouch ratio index (left) and the same index mean-filtered as for total flint (right).

Figure 7. Manufacturing areas (left) defined as areas where the ‘normalised’ core:retouch ratio is below 0.9, and discard areas (right) where the ratio is above 1.1.

Table 2. Summary of logistic multiple regression model for 'manufacture areas'.

Step	Variable Entered	Variable Removed	Number In	Score Chi-Square	Wald Chi-Square	Pr > Chi-Square
1	DLBARS		1	78.6853	.	0.0001
2	DRIVER		2	88.4108	.	0.0001
3	ELEV		3	86.4427	.	0.0001
4	CHALK		4	28.1790	.	0.0001
5	SLOPEF		5	42.2459	.	0.0001
6	SE		6	33.7102	.	0.0001
7	NE		7	98.5428	.	0.0001
8	BREND		8	14.5768	.	0.0001
9	DSTREAMS		9	20.9877	.	0.0001
10	VIEWSUM		10	9.1127	.	0.0025
11	DENSRBAR		11	4.6568	.	0.0309

Analysis of Maximum Likelihood Estimates.

Variable	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate
INTERCPT	3.6050	2.1878	2.7152	0.0994	.
ELEV	0.1080	0.0320	11.3757	0.0007	0.574173
SLOPEF	0.5415	0.0710	58.1830	0.0001	0.749372
NE	-2.8268	0.3715	57.8992	0.0001	-0.638711
SE	-2.5466	0.3264	60.8808	0.0001	-0.692974
CHALK	-6.9748	1.2567	30.8042	0.0001	-0.835570
BREND	-1.3809	0.2650	27.1575	0.0001	-0.380738
DSTREAMS	0.1549	0.0289	28.6652	0.0001	0.526867
DRIVER	-0.0547	0.0110	24.7250	0.0001	-0.748168
DLBARS	-0.1525	0.0189	65.4072	0.0001	-1.015312
VIEWSUM	0.2034	0.0577	12.4014	0.0004	0.476920
DENSRBAR	0.0666	0.0296	5.0416	0.0247	0.425465

Association of Predicted Probabilities and Observed Responses

Concordant = 93.2%	Somers' D = 0.874
Discordant = 5.8%	Gamma = 0.883
Tied = 1.0%	Tau-a = 0.085
(390630 pairs)	c = 0.937

within the study area. This is achieved by the use of the cumulative logistic distribution function (Kvamme 1988: 371). Both models were therefore returned to the GIS, and the parameter estimates were used to solve the logistic equation from the landform and cultural overlays for all locations. Unfortunately, the difference in sample sizes between nonsites and sites produces a prediction heavily biased to the prediction of the larger sample: in this case nonsites. This is often undesirable, but it is possible to correct for the sample size bias after running the model by adjusting the intercept parameter by the natural log of the ratio of the sample sizes (Kvamme 1983b) and this adjustment was made to both of these models.

6.4 Manufacturing model

The *manufacturing* model (mapped as fig. 8 left) is complex result, incorporating all of the independent variables to generate the response. The probability of a site being a manufacturing site is increased by lower slope, lower elevations, easterly aspects, presence of chalk rather than clay with flint or valley gravel, proximity to shelter, low visibility of long barrows, low density of round barrows, greater distance from the Avon and presence of brown rendsina soil.

Geographically, the manufacturing model generally shows high values where they would be expected close to the region of the sampled areas: north of the Cursus (area 52) and around the southern rim of Normanton Down

Table 3. Summary of logistic multiple regression model for 'discard areas'.

Summary of Stepwise Procedure

Step	Variable Entered	Variable Removed	Number In	Score Chi-Square	Wald Chi-Square	Pr > Chi-Square
1	BREND		1	94.3617	.	0.0001
2	DLBARS		2	38.0198	.	0.0001
3	VIEWSUM		3	29.1272	.	0.0001
4	NE		4	23.2454	.	0.0001
5	DRIVER		5	19.0265	.	0.0001
6	DENSRBAR		6	20.4722	.	0.0001
7	ELEV		7	10.0973	.	0.0015
8		VIEWSUM	6	.	0.0711	0.7898
9	TBCEARTH		7	9.6831	.	0.0019

Analysis of Maximum Likelihood Estimates.

Variable	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate
INTERCPT	6.1250	1.3314	21.1635	0.0001	.
ELEV	-0.0521	0.0123	17.9639	0.0001	-0.276782
NE	-1.0030	0.2136	22.0439	0.0001	-0.226632
BREND	-2.7628	0.3861	51.1922	0.0001	-0.761737
TBCEARTH	2.7358	1.0944	6.2491	0.0124	0.420462
DRIVER	0.0383	0.00625	37.5140	0.0001	0.524288
DLBARS	0.1045	0.0108	93.1905	0.0001	0.695530
DENSRBAR	0.1170	0.0187	39.2001	0.0001	0.747777

Association of Predicted Probabilities and Observed Responses

Concordant = 87.5%	Somers' D = 0.757
Discordant = 11.8%	Gamma = 0.762
Tied = 0.7	Tau-a = 0.063
(336660 pairs)	c = 0.878

(south of area 67). Low values on Normanton Down (61, 79, 55), around Coneybury (51), south of Durrington Walls (60, 71, 69) and at the eastern end of the Cursus (76, 66 and 85) are also consistent with the data.

However the model predicts manufacturing areas in an unlikely proportion of the area southeast of the river. Examination of the equation, and of the independent variables suggests that the high values in the southeast are primarily influenced by the distance to long barrows and the density of round barrows indices. Both show extreme values in this area, and both have the type of dual relationship discussed above in relation to the linear multiple regression model. The area of high value in the northwest of the study area may be related to distance from the river Avon, although this seems to be also partly an effect of the distance to long barrows index.

For the same reasons as for the linear regression therefore, it seems likely that predictions away from the surveyed areas are unreliable and the model must be regarded as internal to the cluster of monuments rather than portable.

6.5 Discard model

The probability of a location being a discard site is increased by the presence of brown rendsina soils and absence of calcareous earths, higher altitudes, northeastern aspect, proximity to the Avon, low density of round barrows and low distance to long barrows. Both density of round barrows and distance to long barrows decrease the probability, which is curious as the two are inversely related. This is in contrast with the manufacturing area model, although the effects of these two variables may be

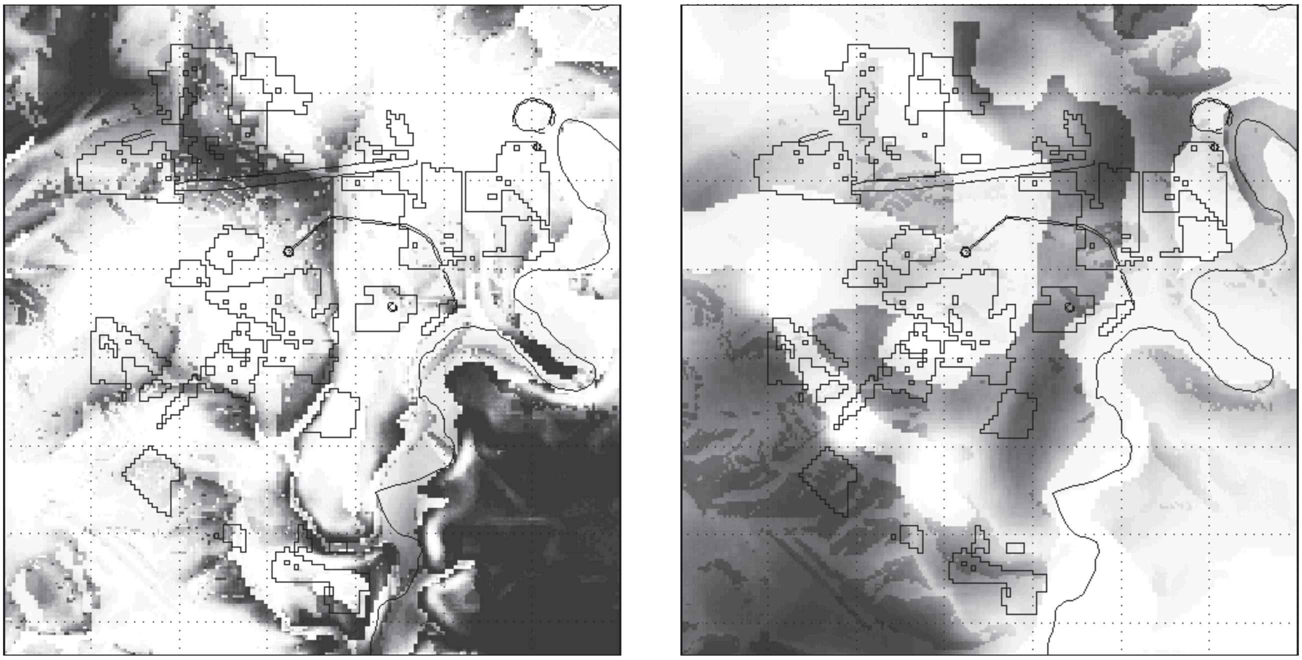

Figure 8. Probability of any location being a manufacturing area (left) or a discard area (right) according to the intercept-adjusted logistic regression equation.

marginal compared to soil. The marginal influence within the model of the cultural variables is enforced by the lack of long barrow visibility which shows insufficient correlation to appear in the model.

In spatial distribution (fig. 8 right), the 'discard' model also shows some encouraging features: high probabilities of discard sites around the elbow of the avenue (87), south of Normanton Down (67) and south of Winterbourne Stoke Crossroads (50) each seem to fit the data. Low probabilities north of the Cursus (52) contrast with the high values of the manufacturing model as should be expected as this forms the major axis of variation between the two areas selected as dependent variables.

6.6 Optimisation and assessment of performance

Assessment of the degree of confidence which should be placed in the predictions is difficult. In an ideal situation, further samples would be taken throughout the study area and the results compared with the predicted outcome for those locations. This is rarely possible, however, and it is perhaps slightly ironic that in situations where this were possible, there would then be rather less point in constructing a model.

One source of data concerning model performance is the ratio of observed responses to predictions within the data itself, and this is provided with the output from the procedure. In the case of these models, this indicates that the manufacturing area model makes 93% correct predictions against nearly 6% incorrect, while the discard area model is a poorer fit with 87.5% correct against nearly 12% incorrect predictions. However this is widely recognised as an extremely optimistic assessment of the performance of the probability model and Warren (1990b) recommends withholding a random control sample of observations from the prediction and then comparing the predictions with these controls. Carmichael (1990), used the control procedure advocated by Warren and found that a 72% correct prediction rate amongst the sites used for prediction produced only 55% correct prediction amongst the controls.

In this case, however, the samples are grouped tightly together within walked areas so that any randomly selected subset of points would still have fallen within the same surveyed areas as the cases included in the study. Intuition, and experience with the linear regression model suggest that these are likely to be the best performing areas of the model. Consequently little confidence could be held in any assessment of the model based on such a sample of sites, and it was felt that insufficient benefit would probably be obtained to offset the removal of the cases.

One method to assess the performance of the model is to force it to predict the same percentage of the surveyed area

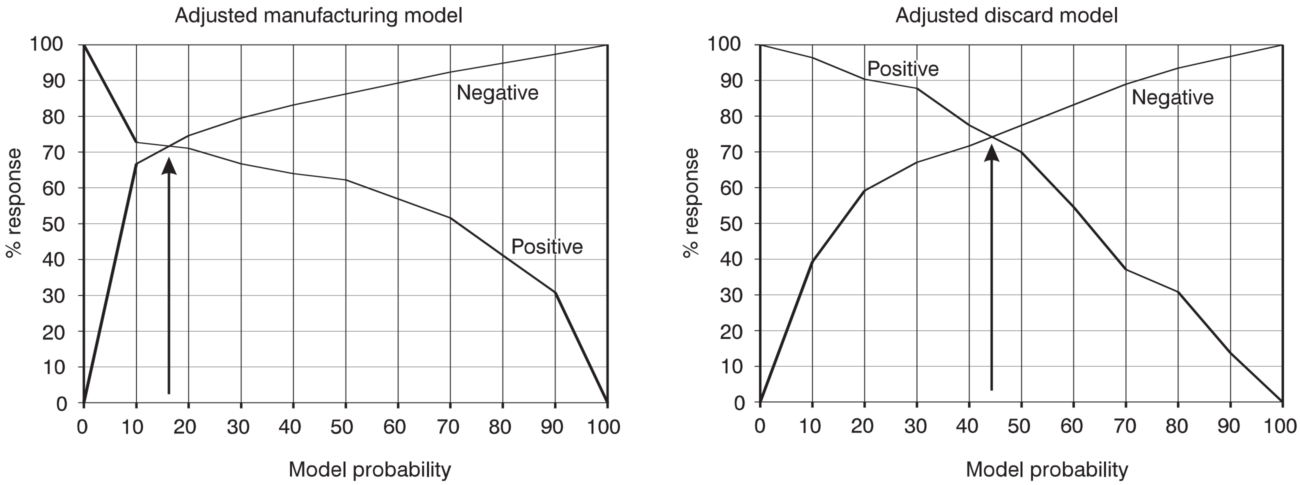

Figure 9. Cut-off points for the intercept-adjusted manufacturing model (left) and for the intercept-adjusted discard model (right) based on the probability which results in an equal proportion of correct positive and negative responses.

as a positive response as occurs within the sample data, and then comparing this prediction with the actual result — a procedure analogous in many ways to the examination of residuals of linear regressions. To 'force' the prediction, a threshold was selected for each probability map which generated the same percentage of positive responses within the worked areas as the original dependent variables. An appropriate threshold was obtained from cumulative frequencies of the probability maps for the surveyed areas. The result of this procedure can then be compared with the original dependent variable, with four possible outcomes. The model may (1) correctly predict no site, (2) correctly predict a site, (3) incorrectly predict a site or (4) fail to predict a site.

It remains to decide on an appropriate threshold for prediction of sites from the adjusted models. It has already been seen that prediction of the 'correct' number of positive responses will not provide a good prediction of presence. Increasing the threshold which is used as a prediction increases the number of incorrect predictions of negative responses but also increases the number of correct predictions of positive responses. The optimum solution must therefore be sought from the models, and occurs when the proportion of correct predictions for positive responses is equal to the proportion of correct predictions for negative responses. This probability value is referred to as the cut-off point for the model, and can be obtained by graphing the observed proportions of correct positive and negative responses against the predicted probability. Observations for these graphs were made by repeated application of the models at appropriate intervals. The point where the positive and negative response curves cross provides both the optimum point of the model and the percentage of correct predictions. From the graphs in figure 9 it can be seen that this cut-off point for the manufacturing model occurs at 15%, where the model correctly predicts 71% of positive responses (manufacturing areas) and 72% of negative responses. For the discard model, the cut-off occurs at 44% where the model predicts 73% of positive responses (discard areas) and 74% of negative responses. The four alternatives within the known data values are mapped as in figure 10 left (manufacture) and right (discard).

6.7 Using the logistic models

The final models provide a method for assessing how likely it is that any location within the landscape would contain a lithic assemblage with either of two particularly interesting characteristics. Clearly the most obvious application of such a model is to the management of the archaeological resource in which they might be used as a method for assessing the relative impacts of alternative courses of action on the (unknown) archaeological resource. However, although using them for this purpose is valid, the models must be understood before they can be best utilised. For one thing, the predictions for manufacturing areas in the far northwest and southeast of the study area are spurious, and should be disregarded. An alternative model may be developed for areas which are marginal or outside the cluster of monuments, but these two are not applicable for the reasons outlined above.

The models can be used to make predictions in a sophisticated way. The optimum predictions which were obtained from the cut-off points of the models may not be

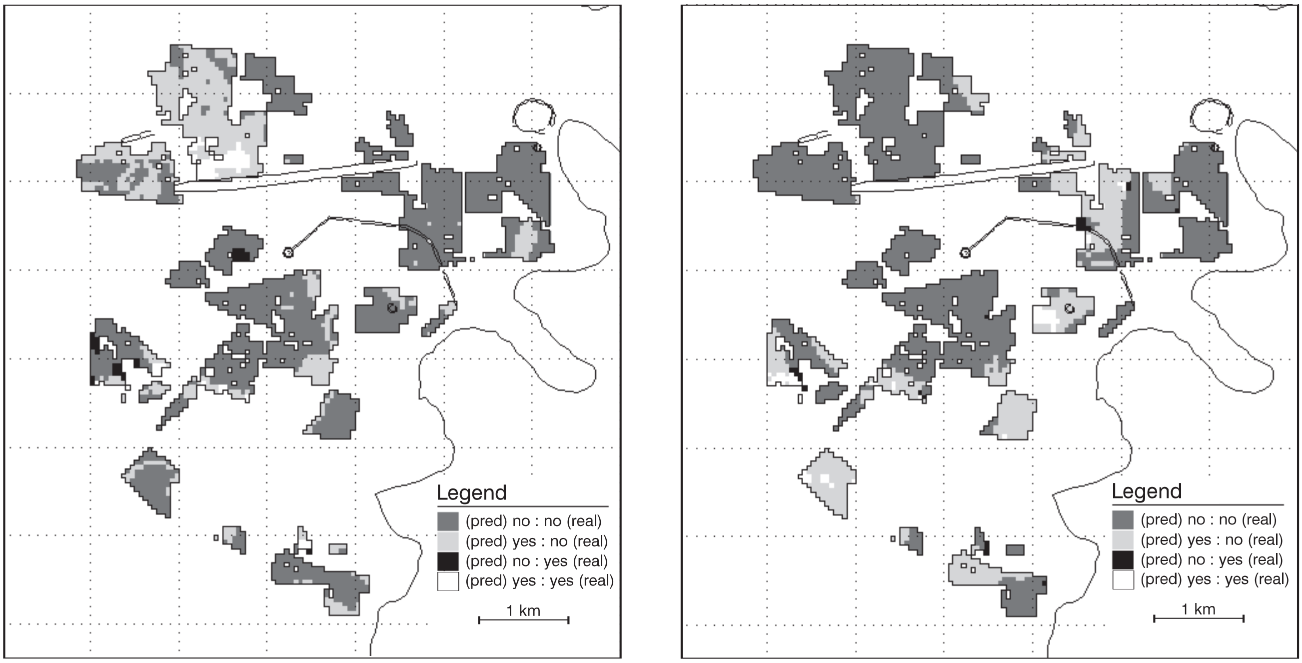

Figure 10. Comparison of predictions made at the selected cut-off points with the observed archaeological values for presence of manufacturing areas (left) and discard areas (right).

the most useful predictions from an archaeological point of view and alternative thresholds might be chosen in order to predict more sites at the expense of nonsites. Given enough resources, of course, an archaeological management strategy would not require a model at all but include provision to survey all areas. In reality, however, management strategies are constrained by resources which in turn restrict the area which may be surveyed. Depending on this, thresholds may be defined which progressively reduce the area which needs to be surveyed, while maximising the likelihood that the interesting areas will be within them. For example, although the manufacturing model predicts roughly 70% of the sites at the cut-off point of 45%, a threshold of 20% may be selected to obtain a prediction which accounts for 90% of the manufacturing areas at the expense of predicting 40% of the non-site areas as manufacturing areas also.

7 Conclusions

It was the aim of this paper to examine whether or not predictive modelling approaches provide a reasonable alternative to spatial interpolation techniques in the analysis of lithic density data. The result of the attempt to use linear multiple regression to predict the lithic densities from correlated variables is, from this point of view, disappointing. That only 25% of the variability within the flint data can be accounted for in this way suggests strongly that the types of deterministic models applied with some success in some North American situations may be inappropriate within complex cultural landscape settings such as this. It is particularly revealing that the majority of the predictive power of the model (low though that is) is accounted for by cultural indices, particularly the density of round barrows, rather than the indices of landform which have been used in other contexts. Although some of the failure of the linear regression to account for the variation in the lithics might be due to the compound nature of the dependent variable, and the explanatory power of the logistic regression models is far more difficult to assess, it is likely to be of the same order as the multiple regression because the same predictors were used for both.

In both cases, the portability of the models is compromised by the use of predictors which seem to have complex relationships with the data. The extreme predictions which occur outside the convex hull of the surveyed areas in both the linear multiple regression model, and both logistic models seem best explained as 'over-training', in the sense that the model is too specifically related to the variation *within* the cluster of monuments to provide a useful estimate of the density of lithics *between* monument clusters. In a sense this is a problem of scale: at the local scale, the relationship between density of monuments and activity may be that the activity takes place away from monuments while at a regional scale this relationship cannot be sustained.

Although untested at present, one way of circumscribing this may be to restrict the variability of the independent variables to that which is found within the data itself through reclassification and in this way restricting the prediction to variation which is within the 'experience' of the regression. Alternatively the independent variables may be left in an unclassified form so as to include the maximum information in the regression, but the prediction may be restricted to the maximum and minimum values of those variables which are observed within the archaeological data.

bibliography

Allen, K.M.S.
S. Green
E.B.W. Zubrow (eds) 1990 *Interpreting Space: GIS and Archaeology*. London, Taylor & Francis.

Altschul, J.H. 1988 Models and the modeling process. In: W.J. Judge/L. Sebastian (eds), *Quantifying the present, predicting the past*, 61-96, Denver: U.S. Department of the Interior Bureau of Land Management.

1990 Red Flag Models: the Use of Modelling in Management Contexts. In: K.M.S. Allen/S.W. Green/E.B.W. Zubrow (eds), *Interpreting Space: GIS and Archaeology, 226-238*, London: Taylor & Francis.

Brandt, R.
B.J. Groenewoudt
K. Kvamme 1992 An experiment in archaeological site location: modeling in the Netherlands using GIS techniques, *World Archaeology* 24 (2), 268-282.

Burrough, P.A. 1986 *Principles of Geographic Information Systems for Land Resources Management*. Oxford: Clarendon Press.

Carmichael, D.L. 1990 GIS Predictive Modelling of Prehistoric Site Distributions in Central Montana. In: K.M.S. Allen/S.W. Green/E.B.W. Zubrow (eds), *Interpreting Space: GIS and Archaeology*, 216-225, London: Taylor & Francis.

Carr, C. 1985 Introductory remarks on regional analysis. In: C. Carr (ed.), *For concordance in archaeological analysis*, 114-127, Kansas City: Westport Publishers.

Clarke, K. 1990 *Analytical and computer cartography*. London: Prentice Hall.

Crumley, C.L.
W.H. Marquardt 1990 Landscape: a unifying concept in regional analysis. In: K.M.S. Allen/S.W. Green/E.B.W. Zubrow (eds), *Interpreting Space: GIS and Archaeology*, 73-79, London: Taylor & Francis.

Ebert, J.I. 1988 Remote sensing in archaeological projection and prediction. In: W.J. Judge/L. Sebastian (eds), *Quantifying the present, predicting the past*, 429-492, Denver: US Department of the Interior Bureau of Land Management.

Ebert, J.I.
T.A. Kohler 1988 The theoretical basis of archaeological predictive modeling and a consideration of appropriate data-collection methods. In: W.J. Judge/L. Sebastian (eds), *Quantifying the present, predicting the past*, 97-171, Denver: US Department of the Interior Bureau of Land Management.

Feder, K.L. 1979 Geographic patterning of tool types as elicited by trend surface analysis. In: S. Upham (ed.), *Computer graphics in archaeology: statistical cartographic applications to spatial analysis in archaeological contexts*, 95-102, Arizona State University Anthropological Research Papers 15.

Foley, R. 1981 A Model of Regional Archaeological Structure, *Proceedings of the Prehistoric Society 47*, 1-17.

Gardiner, J.P. 1984 Lithic Distributions and Neolithic Settlement Patterns in Central Southern England. In: R. Bradley/J. Gardiner (eds), *Neolithic Studies: A Review of Current Work*, 177-187, BAR British Series 133, Oxford: British Archaeological Reports.

Goodchild, M.F. 1987 *Spatial autocorrelation*. Concepts and techniques in modern geography 47. Norwich: Geo.

Harris, T.M. 1986 Geographic Information System Design for Archaeological Site Information Retrieval. In: S. Laflin (ed.), *Computer Applications in Archaeology 1986*, 148-161, Birmingham: University of Birmingham.

Hasenstab, R.J. / B. Resnick 1990 GIS in Historical Predictive Modelling: the Fort Drum Project. In: K.M.S. Allen/ S.W. Green/E.B.W. Zubrow (eds), *Interpreting Space: GIS and Archaeology*, 284-306, London: Taylor & Francis.

Hodder, I. / C. Orton 1976 *Spatial analysis in archaeology*. Cambridge: Cambridge University Press.

Kincaid, C. 1988 Predictive modeling and its relationship to cultural resource management applications. In: W.J. Judge/L. Sebastian (eds), *Quantifying the present, predicting the past*, 549-569, Denver: US Department of the Interior Bureau of Land Management.

Kohler, T.A. 1988 Predictive locational modeling: history and current practice. In: W.J. Judge/L. Sebastian (eds), *Quantifying the present, predicting the past*, 19-59, Denver: US Department of the Interior Bureau of Land Management.

Kohler, T.A. / S.C. Parker 1986 Predictive models for archaeological resource location. In: M. Schiffer (ed.), *Advances in Archaeological Method and Theory 9*, 397-452, New York: Academic Press.

Kvamme, K.L. 1983a Computer processing techniques for regional modeling of archaeological site locations, *Advances in Computer Archaeology* 6, 26-52.

1983b *A manual for predictive site location models: examples from the Grand Junction District, Colorado*. Colorado: Bureau of Land Management, Grand Junction District.

1985a *Fundamentals and Potential of Geographic Information System Techniques for Archaeological Spatial Research*. Paper presented at 50th meeting of Society of American Archaeology, Denver.

1985b Determining empirical relationships between the natural environment and prehistoric site locations: a hunter-gatherer example. In: C. Carr (ed.), *For concordance in archaeological analysis*, 208-238, Kansas City MO: Westport Publishers.

1988 Using existing archaeological survey data for model building. In: W.J. Judge/L. Sebastian (eds), *Quantifying the present, predicting the past*, 301-428, Denver: US Department of the Interior Bureau of Land Management.

1989 Geographic Information Systems in Regional Archaeological Research and Data Management. In: M.B. Schiffer (ed.), *Archaeological Method and Theory, Volume 1*, 139-203, Tuscon: University of Arizona Press.

1990 The Fundamental Principles and Practice of Predictive Archaeological Modelling. In: A. Voorrips (ed.), Mathematics and Information Science in Archaeology: A Flexible Framework. *Studies in Modern Archaeology* 3, 257-295, Bonn: Holos-Verlag.

Kvamme, K.L. / M.A. Jochim 1989 The Environmental Basis of Mesolithic Settlement. In: C. Bonsall (ed.), *The Mesolithic in Europe*. Papers Presented at the Third International Symposium, Edinburgh 1985, 1-12, Edinburgh: John Donald Publishers.

Leusen, M.P. van 1993 Cartographic modelling in a cell-based GIS. In: J. Andresen/T. Madsen/I. Scollar (eds), *Computing the Past, Computer Applications and Quantitative Methods in Archaeology CAA92*, 105-123, Aarhus: Aarhus University Press.

Lock, G.R.
T.M. Harris forthcoming Danebury revisited: an English Iron Age hillfort in a digital landscape. In: M. Aldenderfer/H. Maschner (eds), *The anthropology of human behaviour through geographic information and analysis*, New York: Oxford University Press.

Lock, G.R.
Z. Stančič 1995 *Archaeology and Geographic Information Systems: a European Perspective*. London: Taylor & Francis.

Maskell, J.E. 1993 *From ploughzone to pixel*. University of Southampton: MSc Thesis.

Richards, J. 1984 The Development of the Neolithic Landscape in the environs of Stonehenge. In: R. Bradley/J. Gardiner (eds), *Neolithic Studies: A Review of Current Work*, 177-187, BAR British Series 133, Oxford: British Archaeological Reports.

1990 *The Stonehenge Environs Project*. Southampton: Hobbs.

Rose, M.R.
J.H. Altschul 1988 An overview of statistical method and theory for quantitative model building. In: W.J. Judge/L. Sebastian (eds), *Quantifying the present, predicting the past*, 173-255, Denver: US Department of the Interior Bureau of Land Management.

Sebastian, L.
W.J. Judge 1988 Predicting the past: correlation, explanation and the use of archaeological models. In: W.J. Judge/L. Sebastian (eds), *Quantifying the present, predicting the past*, 1-18, Denver: US Department of the Interior Bureau of Land Management.

Shennan, S.J. 1980 Meeting the plough damage problem: a sampling approach to area-intensive fieldwork. In: J. Highcliffe/R.T. Schadler Hall (eds), *The past under the plough*, Department of the Environment Occasional Paper 3, 125-133.

1985 *Experiments in the collection and analysis of archaeological survey data: the East Hampshire Survey*. Sheffield: Sheffield University Press.

1988 *Quantifying Archaeology*. Edinburgh: Edinburgh University Press.

Thomas, D.H. 1975 Non-Site Sampling in Archaeology: Up the Creek without a Site. In: J.W. Mueller (ed.), *Sampling in Archaeology*, Arizona: University of Arizona Press.

Thoms, A.V. 1988 A survey of predictive locational models: examples from the late 1970s and early 1980s. In: W.J. Judge/L. Sebastian (eds), *Quantifying the present, predicting the past*, 581-645, Denver: US Department of the Interior Bureau of Land Management.

Wansleeben, M. 1988 Applications of Geographical Information Systems in Archaeological Research. In: S.P.Q. Rahtz (ed.), *Computer and Quantitative Methods in Archaeology 1987*, 435-451, BAR International Series 446, Oxford: British Archaeological Reports.

Warren, R.E. 1990a Predictive Modelling of Archaeological Site Location: a Case Study in the Midwest. In: K.M.S. Allen/S.W. Green/E.B.W. Zubrow (eds), *Interpreting Space: GIS and Archaeology*, 90-111, London: Taylor & Francis.

1990b Predictive Modelling of Archaeological Site Location: a Primer. In: K.M.S. Allen/S.W. Green/E.B.W. Zubrow (eds), *Interpreting Space: GIS and Archaeology*, 201-215, London: Taylor & Francis.

Wheatley, D.W. 1993 Going over old ground: GIS, archaeological theory and the act of perception. In: J. Andresen/T. Madsen/I. Scollar (eds), *Computing the Past, Computer Applications and Quantitative Methods in Archaeology CAA92*, 133-138, Aarhus: Aarhus University Press.

1995a *The application of GIS to archaeology, with case studies from Neolithic Wessex*. PhD Thesis: University of Southampton.

1995b Cumulative Viewshed Analysis: a GIS-based method for investigating intervisibility, and it's archaeological application. In: G. Lock G/Z. Stančič (eds), *Archaeology and Geographic Information Systems: a European Perspective*, 171-185, London: Taylor & Francis.

forth-coming A View to a Hill – the use of line-of-sight maps to understand regional variation in earlier Neolithic Wessex. In: H. Maschner (ed), *GIS and the Advancement of Archaeological Method and Theory*, Carbondale: Southern Illinois University Press.

Williams, I.
W.F. Limp
F.L. Briuer
1990 Using geographic information systems and exploratory data analysis for archaeological site classification and analysis. In: K.M.S. Allen/S.W.,Green/E.B.W. Zubrow (eds), *Interpreting Space: GIS and Archaeology*, 237-239, London: Taylor & Francis.

Zubrow, E.B.W. 1990b Modelling and prediction with geographic information systems: a demographic example from prehistoric and historic New York. In: K.M.S. Allen/S.W. Green/E.B.W. Zubrow (eds), *Interpreting Space: GIS and Archaeology*, 307-318, London: Taylor & Francis.

David Wheatley
Department of Archaeology
University of Southampton
Southampton SO17 1BJ
United Kingdom
e-mail: dww@soton.ac.uk