Data analysis of a magnetic survey to contrast the most common treatments of data procedures in shallow archaeological surveys[1]

José Antonio Peña
Dept. de Prehistoria y Arqueología, Universidad de Granada, Spain.
Instituto Andaluz de Geofísica, Universidad de Granada, Spain.
E-mail: peruano@ugr.es

Jose Antonio Esquivel
Dept. de Prehistoria y Arqueología, Universidad de Granada, Spain.
Instituto Andaluz de Geofísica, Universidad de Granada, Spain.
E-mail: esquivel@gea.ugr.es

Antonio Ramos
Dept. de Prehistoria y Arqueología, Universidad de Granada, Spain.

María del Mar Osuna

Esther Rull

Introduction

Geophysical surveys applied to archaeology generate a great amount of raw data, defined by two spatial variables (X,Y) and a quantitative variable (Z) measuring characteristic magnitude: magnetic field intensity, eletrical intensity and potential, etc. These raw data sometimes need a preliminary analysis, focused on obtaining a physical magnitude, related to the subject to be studied; we eliminate, if necessary, some external perturbations (e.g., temporal fluctuations in the magnetic field). Further analysis could provide the information necessary to estimate archaeological inferences based on data results.

In this paper we have conducted a geophysical survey, developed by means of the magnetic gradient method in the Nazzarian site at Garbín (Baza, Granada, Spain) which has recently been excavated. The data obtained have been analyzed through a complex set of methods, which include new techniques developed by our team and others; these methods are commonly used in archaeological studies, such as spatial statistics, directional derivatives, filtering techniques, and so on. These methods are focused on obtaining underground shallow structures, which point out the local anomalies. The results of these analyses are contrasted with the structures and the information provided by the excavation of the site.

Data Set

Data acquisition was carried out on a 25 m. square, N-S oriented, including the underlying structure, and superficial remains were removed to avoid external perturbations.

The experimental device was configured of two sensors located in the same vertical line: the bottom sensor was placed 70 cm. distance to soil and the top sensor 210 cm. distance to soil (the distance between sensors was 140 cm.), with N-S and S-N profiles separated by a distance of 1 m. and with the same distance between sucessive sensor placements (Breitner, 1973).

Data was treated as if it had been taken simultaneously, computing a set of values, that constitute an estimation of gradient (the gradient method, *sensu strictu*, demands that the distance between sensors must be an infinitesimal term of depth, where the object producing the anomaly is located) by means of:

$$g_m = \frac{t_m - b_m}{d}$$

g_m being the vertical magnetic gradient, t_m = the top sensor measure, b_m = the bottom sensor measure and d = the sensors' distance. This method eliminates the influence of time variation, emphashizing local anomalies *versus* regional trend.

Excavation Information

The results of the excavation, to be contrasted with the data yielded by prospection, revealed a rectangular tower built on a small foundation ditch. The walls (nearly 90 cm. thick) were built with mud brick, with sand and some stones; the adjacent areas were composed of alluvial clay lying under 50 cm. layer of agricultural soil. This structure was closed in the past, as revealed by consecutive clay and mortar levels, but it was dug by clandestine excavators and was later filled with the same material obtained during these illegal excavations.

Three important, collapsed brick walls were found surrounding

---

[1] Paper related to SEC94-0633 CICYT Project

the tower. The most important of these were located in the S-W oriented area of the tower structure.

Data Analysis

Analysis of the data set yielded by the geophysical surveys was carried out by means of gridding methods and by previous data analysis, to point out the main anomalies, to eliminate white noise (if possible) and to obtain clearer results. The most common methods used were the following:

Gridding Methods

The most commonly used technique of data display is a map of contours, drawing 2D level curves, that plot the geophysical variable on a grid, fitting the raw data through interpolation methods. These methods applied to data set obtained by the magnetic survey in Garbín site, have supplied the following results:

Geostatistical krigging. This method attempts to show the trends that exist in the data, by means of a contour and surface display. Spatial variability is measured by the variogram function, providing the weights for interpolation of the existing points between grid nodes.

Analysis was accomplished by adjusting the data set to several theoretical variogram models: linear (Figure 1), exponential, gaussian, wave or hole-effect, and rational quadratic model. The application of these models to our raw data provide the same results as the linear variogram model,without any any variations: the anomalies are fitted to ground bodies but they can be originate by any computational objects.

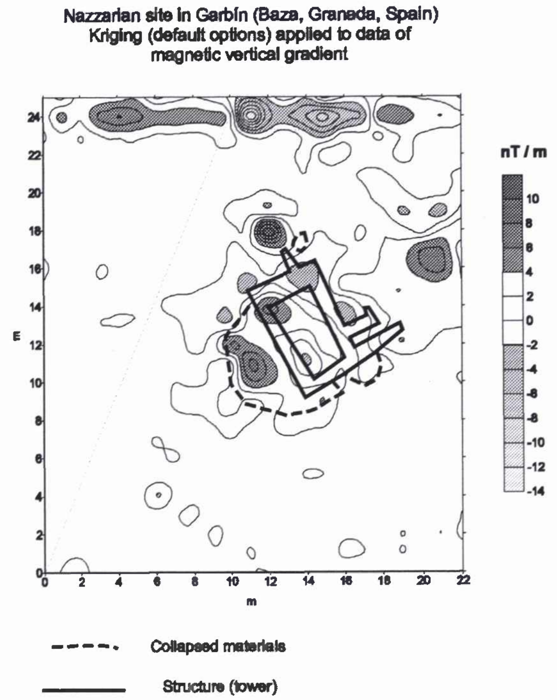

Figure 1

Inverse distance to a power. This is a weighted average interpolator method, that assigns a weight to each point, proportional to the inverse power of distance to the nearest grid node. This technique is controlled by the power parameter p that determines the contrast between the results: the greater the value produced, the greater the contratst, and conversely.

Analysis was accomplished using the powers $p=2$ and $p=8$, showing similar results to those obtained through linear krigging interpolation, but these results were more accute. If $p=8$, the anomalies appear to be more contrasted.

Minimum curvature. This method generates an interpolated surface that passes through all the points with minimum curvature. Its application provided equal results to those obtained using the inverse distance to a power method.

Nearest neighbor. This is a method of fitting data to a previous grid by means of no-interpolation, assigning the value of the nearest point to each node in the previously established grid. The application of this method to our data set gave equal results to those obtained using the inverse distance with a power parameter greater or equal than 2.

Polynomial regression. This method adjusts data to a surface of previously defined type, in order to model the existing trend in the data. It is mainly used to define large scale trends, softening the individual values of the variable. Its application to our data set produced computational objects only, as were expected.

Radial basis functions. This is a method that adjusts the point data to grid nodes by means of interpolations based on different previously selected functions, these functions are those which assign corresponding weights to each point during interpolation.

The application of this method to our data set was accomplished using multiquadric, inverse multiquadric, multilogarithmic, cubic spline and thin plate spline functions, showing results similar to those obtained using the linear krigging or by inverse distance (multiquadric function only).

Shepard's method. This procedure adjusts data, using the least squares approximation to fit the weights, and it provides us with the inverse distance; therefore it is similar to the inverse distance to a power interpolator. The results are very similar to those obtained using the inverse distance to a power, but created computational objects.

Triangulation w / Linear interpolation. This method produces triangles between the points, in such a way that no triangle edges are intersected by other triangles. Triangulation works best when the data are evenly distributed over the grid area. The application of this method to our data set provided poor results: increased the length of large anomalies and created computational objects.

Standard Data Analysis

The interpretation of raw data frequently requires the definition and quantification of characteristics that may not be obvious by looking at a contour or surface map. The most common data analysis methods avalaible for contour maps are directional derivatives, terrain modeling and differential operators. These methods applied, to the Garbín data, provided the following results:

Directional derivatives. This method provides information about the slope of the surface in a previously fixed direction, but not the steepest slope at a given point. The most common methods of data analysis using directional derivatives are the following:

1) *first derivative,* provides the slope of the surface along the direction of the line chosen.

2) *second derivative,* provides the rate of change of slope along the direction of the line chosen. This procedure makes no sense for our raw data set, because the gradient was assimilable by the first one derivative.

3) *directional curvature*, measures the rate of change of the inclination angle of tangential planes on a profile line, defined by the surface along the directional line (similar to second derivative).

These methods are very useful when the directional line is induced or well known. Their application to the data set of Garbín provided a contour map, that required clearing of the z values around zero to eliminate white noise (see Figures 2 and 3).

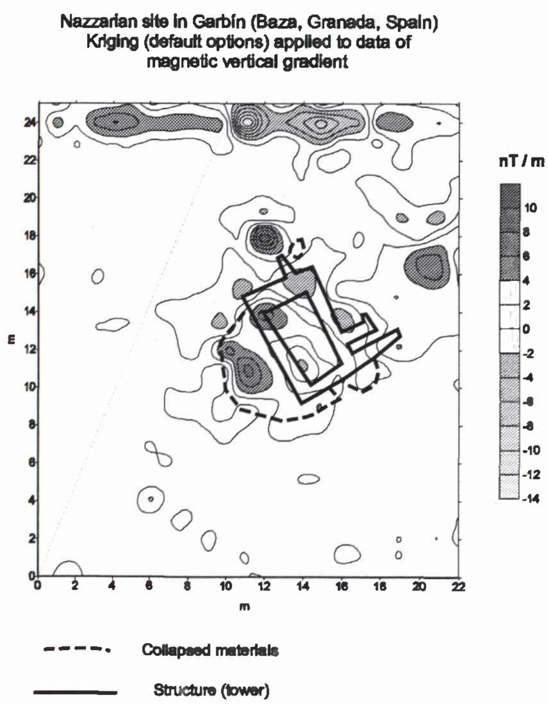

Figure 2

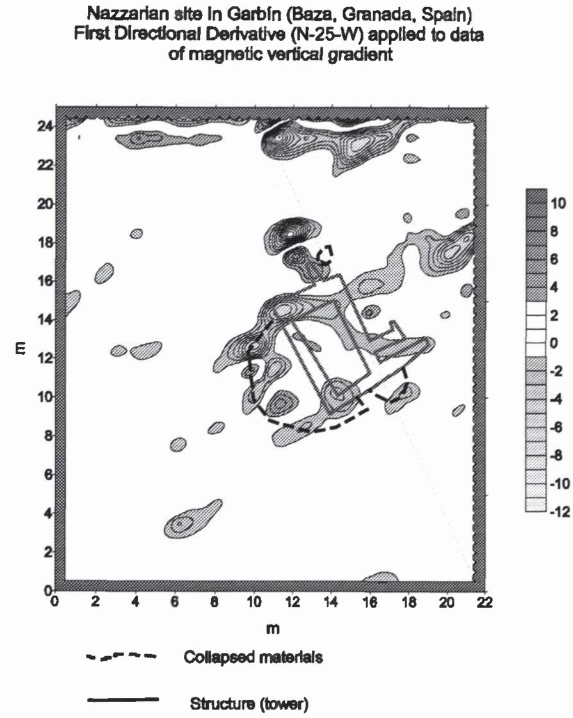

Figure 3

Terrain modeling. This method is based on the direction of the gradient, and it determines the steepest slope and the direction of the gradient (direction of steepest slope) in each grid node. There are five common types of terrain model:

1) *terrain slope* provides the slope, in degrees, at any grid node based on the straight downhill or straight uphill slope at a point, and defines the direction of the gradient at each point on the map (is similar to first derivative).
2) *terrain aspect* provides the direction of the steepest slope at each grid node, using values reported in azimuth.
3) *profile curvature* determines the downhill or uphill rate of change in the slope, in the direction of the gradient at each grid node (similar to second derivative).
4) *plan curvature* points out the rate of change of the angle of the steepest slope, in the horizontal plane (i.e. a measure of the amount of curvature along contour lines).
5) *tangential curvature* is similar to plan curvature, but measures the rate of change of the angle of steepest slope in relation to the vertical plane.

The application of these methods to our data set yielded the following results: the last four methods made many computational objects disguising the magnetic anomalies. The terrain slope method provided powerful contrasts between anomalies, but required clearing a great amount of values, while keeping only the very big slopes.

Differential operators. These are a set of operators, that define several options using slope values. The most common are the following:

1) *gradient operator,* measures the magnitude of the gradient at any point on the surface, reported as a number: 0 value for a horizontal surface and approaches infinity as the slope approaches the vertical (Figure 4). Results are improved by removing lower values.

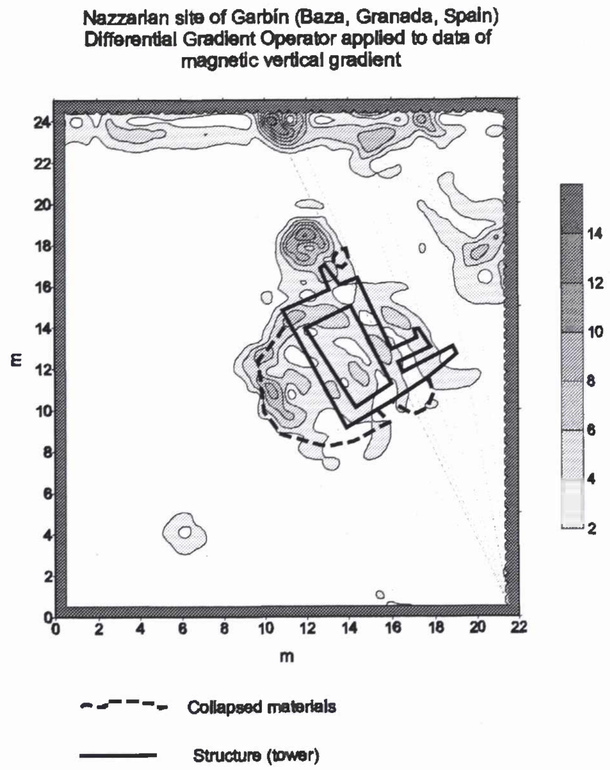

Figure 4

2) *laplacian operator* quantifies the net flow into or out of a local control volume, in physical quantities, which have a local flow rate proportional to the local gradient.

3) *biharmonic operator* is the equivalent of the application of the laplacian operator twice.

Image Data Analysis

Image analysis of data uses methods based on filtering to transform data and to detect underlying patterns, by means of a kernel (moving window usually 3x3 or 4x4 size) focused on obtaining specific transformations. The most usual filtering methods applied to our data provided the following results:

- *low pass filters.* These filters clear high frequencies and produced a great number of computational objects.

- *high pass filters.* These filters clear low frequencies, producing many comptutational objects and white noise effects.

-*zero-sum kernel* uses a kernel with the total sum of raw elements equal to zero. Anomalies fit well to bodies, but do not provided any new information with respect to other methods (Figure 5).

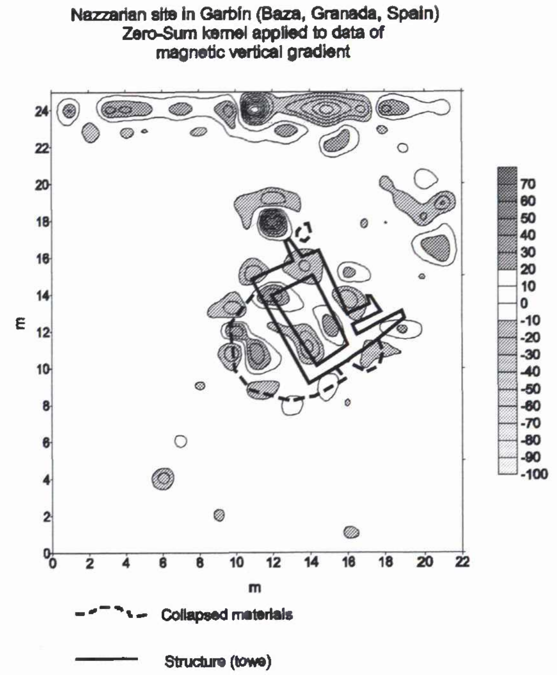

Figure 5

- *vertical boundary* and *horizontal boundary drawing.* These filters attempt to obtain the directional N-S and E-W boundaries. Applied to our data, they produced many computational objects.

- *laplacian boundary drawing.* This filter provides contours by means of the laplacian operator, and applied to our data produced a great number of white noise effects.

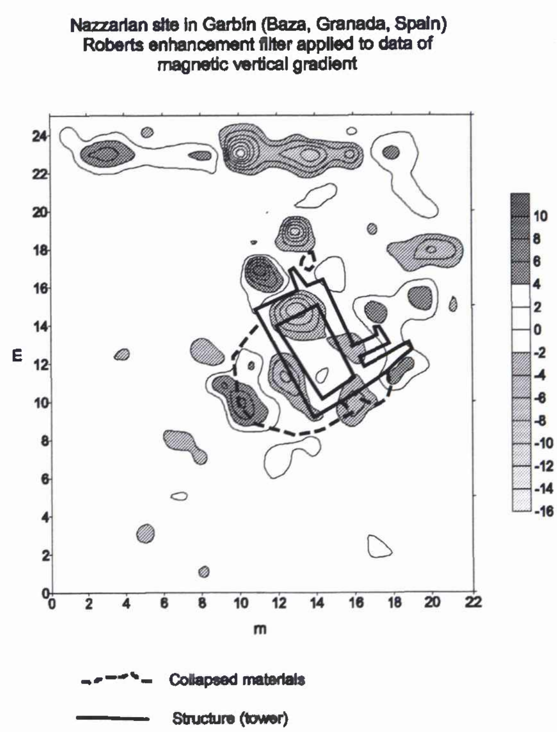

Figure 6

- *Sobel's zero-sum.* This is a family of specific filters, zero-sum type, focused on enhancing the raw data. All of the different Sobel's filters applied to our data, did not produce any new information nor provided any computational objects.

- *Robert's zero-sum.* This is a family of specific filters, zero-sum type, focused on enhancing the raw data using a very simple matrix. One Robert's filter applied to our data fits well the anomalies to bodies, but did not provide new information (Figure 6).

Statistical Analysis With Spatial Components

This statistical method uses a specific algorithm focused on obtaining a partition of *z* variable in *N* intervals, such that the differences between the obtained intervals (z_i, z_{i+1}), i=1...*N* are maximized. First, the procedure divides the area into rectangles to obtain an X,Y spatial grid. Then, a distribution of class , to z values, is selected by means of an algorithm designed by the authors in order to detect significant differences between consecutive z values. This algorithm is not spatial but takes into account the total relation of the data, according to the following design and providing spatial results related to X,Y coordinates (Esquivel, Benjumea & Peña, 1997):

1) z values are put in increasing value order, and the z_k-z_{k-1} differences are calculated. Then, a threshold value (which may be assigned beforehand) and a maximum value of differences are selected, and the variable $a_0=z_1$ is assigned as a first interval limits for data distribution (later, these limits will be provided by the a_i values).

2) if (z_k-z_{k-1})/maximum>threshold, a new interval, a_k-a_{k-1} is accepted, assigning its lower limit as the upper limit of the previous interval. In this way a new maximum value is obtained.

3) find i, such as z_k , is assigned to $[a_i-a_{i-1}]$ interval. If the condition has been fulfilled, this interval will be a new one; otherwise, it will be the last one.

4) steps 2-3 are repeated until all data are classified in intervals.

The algorithm finds the contrasts existing in the data, and obtains the maximum contrasts between anomalies and the environment. The application of the alagorithm to our raw data provided the same results as the lineal krigging method, including previous data transformation by means of log transformation, mean 0 and variance 1 transformation, and the Freeman-Tukey transformation (Cressie, 1991).

Using the unweighted median-polish transformation of raw data (Tukey, 1977; Cressie, 1991) we obtained the residuals R(x,y) according to the following expression:

$$z(x_i, y_i) = a + r_i + c_j + R(x_i, y_i)$$

a being = global effect (regional trend), r_i= row effect, c_j = column effect, $R(x_i,y_i)$ = local effect (local trend) and ~ being the estimation of parameters (Cressie, 1991). The results obtained by applying the previous algorithm to residuals were the best of all: anomalies were very well contrasted and fitted to existing structures. The aplication to the residuals, obtained by the median-polish algorithm povided the following results:

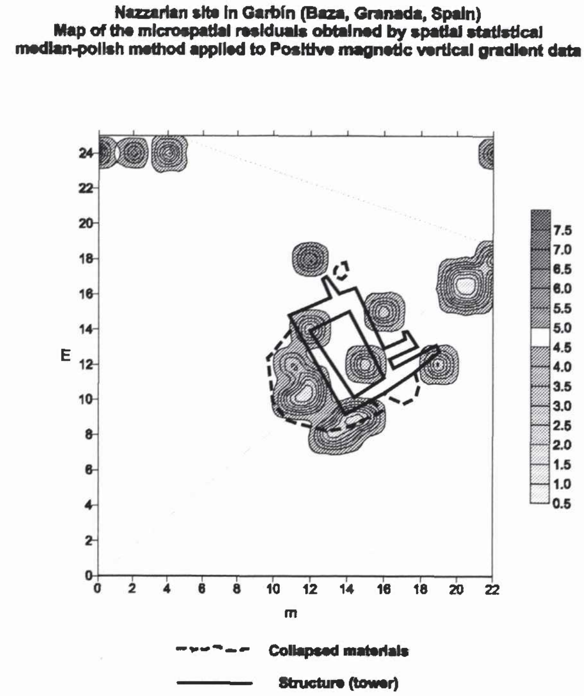

Figure 7

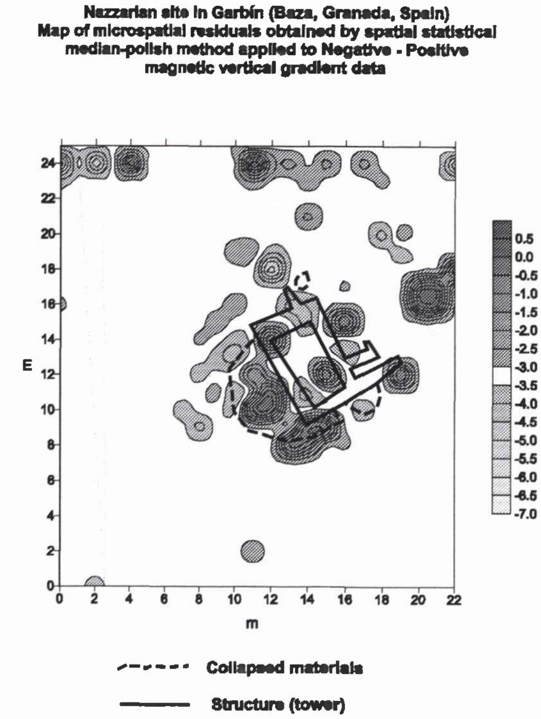

Figure 8

-using data transformation (positive, negative and positive and negative data) the results were similar to those obtained by means of lineal kriging. Application to negative raw data only produced the same results (Figure 7).

- Applied to positive raw data (obtained by spatial statistical median-polish) the microspatial residuals provided the best results: this procedure eliminated white noise and produced the best fit between anomalies and producing bodies. The application to negative-positive residuals showed the same pattern but produced some white noise (Figure 8).

Conclusions

Data provided by geophysical surveys needs to be adjusted to a previously estabalished grid and, in many cases, to be analyzed and transformed by means of different methods of data analysis.

Application to the same data set, contrasted to the later excavation, of the most common methods used in geophysical surveying (as well as some less standard methods) provided the following results:

- linear kriging is a method of fitting data to a previously established grid, which provides good results, fitting well the anomalies to the bodies that produced them. Some clearer results were obtained using the method of inverse distance to a power. The rest of the methods are similar to some of the previous ones and, in many cases, they were found to produce computational objects.

- the methods based on computing directional derivatives produced an important byass in function to the chosen direction, for the realization of the derivation procedure.

- the methods based on the study of terrain modeling models, adjusted the anomalies well to the producing bodies, providing very powerful contrasts among anomalies; but they required the elimination of many z values, while keeping only the very large ones.

- application of filters with different kernels generally produced white noise and computational objects, with the exception of some filters, of zero-sum kernel.

- the method of classification into intervals proposed by the authors (Esquivel & Peña, 1996; Esquivel, Benjumea & Peña, 1997), provided the same results as linear kriging, eliminating almost all of white noise and without creating computational objetcs.

- application of the previous method to residual values, obtained from positive raw data, using the algorithm median-polish of Tukey (Cressie, 1991), provided the best results: a map of perfectly fitted anomalies to producing bodies, the elimination of white noise and the computacional objects, without the intervention of any subjective decision. The contours map or surface image obtained by means of this method shows the most sharp map at all, taking into account the main anomalies only, very well situated and fitted to producing bodies.

References

BREITNER, S. (1973), *Applications Manual for Portable Magnetometers*, Geometrics, Sunnyvale.

CRESSIE, N., (1991), *Statistics for Spatial Data*, John Wiley & Sons, Chichester.

ESQUIVEL, J.A., Benjumea, B. & Peña, J.A. (1997), "Spatial statistical analysis applied to magnetometric archaeolgical data", *Archeologia e Calculatori*, 8, pp. 55-65.

ESQUIVEL, J.A. & Peña, J.A. (1996), "Avance al análisis espacial estadístico de datos geofísicos. Aplicación a la Arqueología", in: Vidal, F., Espinar, M. & Esquivel, J.A. eds, *Homenaje en honor al profesor Fernando de Miguel Martínez*, Servicio de Publicaciones de la Universidad de Granada, Granada, pp. 197-206.

FREEMAN, M.F. & Tukey, J.W., (1950), "Transformations related to the angular and the square root", *Annals of Mathematical Statistics*, 21, pp. 607-611.

HAŠEK, V. & Vencalak, R., (1989), "Methods for Processing Digital Geophysical Data in Archaeology", RAHTZ, S. & RICHARDS, J. (eds.), *Computer Applications and Quantitative Methods in Archaeology, BAR International Series 548*, pp. 179-192.

KEMPERMAN, J.H.B., (1984), "Least absolute value and median polish", in: Tong Y.L. (ed), *Inequalities in Statistics and Probability*, IMS Monograph Series, Vol. 5, Institute of Mathematical Statistics, Hayward, pp. 84-103.

SCOLLAR, I., (1990), *Topics in Remote Sensing 2: Archaeological Prospecting and Remote Sensing*, Cambridge University Press, Cambridge.

SHAFFER, G., (1993), "An Archaeomagnetic Study of a Wattle and Daub Building Collapse", *Journal of Field Archaeology, 20, 1*, Boston University, pp. 59-75.

SNEDECOR, G.W. & Cochran, W.G., (1980), *Statistical Methods, 7hn edition*, Iowa State University Press, Iowa.

TUKEY, J.W., (1977), *Exploratory Data Analysis*, Addison-Wesley, Reading.

UPTON, G.J.G. & Fingleton, B., (1985), *Spatial Data Analysis by Example, Volume 1: Point Pattern and Quantitative Data*, John Wiley & Sons, Chichester.

UPTON, G.J.G. & Fingleton, B., (1989), *Spatial Data Analysis by Example, Volume 2: Categorical and Directional Data*, John Wiley & Sons, Chichester.

WEYMOUTH, J., (1986), "Geophysical Methods of Archaeological Site Surveying", in: Schiffer, M.B., (ed.), *Advances in Archaeological Method and Theory, Vol. 9*, Academic Press, Orlando, pp. 311-395. *