Chilean Journal of Statistics
Vol. 5, No. 1, April 2014, 19–35

Image Reconstruction
Research Paper

3D modelling and Depth Estimation in Archaeological Geophysics

Robert G. Aykroyd[1,*] and Salem M. Al-Gezeri[2]

[1]Department of Statistics, University of Leeds, Leeds, LS2 9JT, UK,
[2]Department of Statistics, University of Benghazi, Benghazi, Libya

(Received: 19 May 2013 · Accepted in final form: 26 July 2013)

Abstract

Modern archaeological investigations are designed to gain maximum information whilst minimising damage to the archaeological site. Hence non-invasive surface studies are a key preliminary stage. One such technique is magnetometry which produces a surface grid of readings, each giving information about the whole 3D subsurface. The inversion of this convolution is an inverse problem requiring substantial regularization. Current approaches incorporate prior information describing smoothness in a Bayesian setting. Typically the subsurface reconstruction is as a single layer, and requires various physical parameters, such as the depth and extent of the layer, to be known. In general, simultaneous estimation of depth and extent along with the magnetic susceptibility distribution is not possible without including further prior information – in particular about depth and extent. Here it is proposed that information about the stratigraphy is obtained from a separate technique where vertical cores of material are extracted from the site. This borehole data allows estimation of depth, extent and susceptibility at a small number of locations across the site. Resulting information can then be used in prior densities in the full subsurface reconstruction. Further, a multi-layer subsurface model is proposed which allows different depth and extent across the archaeological site. The approach is illustrated with part of a dataset from a real archaeological site for which excavation records are also available.

Keywords: Bayesian model · borehole data · empirical Bayes · image reconstruction · inverse problems · magnetometry · MCMC · stratigraphy.

Mathematics Subject Classification: Primary 62P99 · Secondary 86A22.

1. Introduction

Everything around us is magnetic, and just as we can describe objects and materials by their size, colour or chemical composition, so we can describe them by their magnetic properties. Since the mid 1940s, archaeologists have been using geophysical methods for surveys of archaeological sites. In particular, during the last 30 to 40 years magnetic surveying has become an important tool of research in archaeology.

*Corresponding author. Email: r.g.aykroyd@leeds.ac.uk

ISSN: 0718-7912 (print)/ISSN: 0718-7920 (online)

http://www.soche.cl/chjs

LeBorgne (1960) showed that soil magnetization is a product of the Earth's field strength and the magnetic susceptibility of the soil, which arises from iron-bearing minerals. Additionally, heated objects, such as clay, brick and pottery, can exhibit thermo-remanent magnetism, in which the object retains its magnetism in the absence of an external field. The sources of this magnetization are the iron compounds in the soils. The key property is the magnetic susceptibility which is a dimensionless quantity, but which is usually expressed with SI or emu after the value to indicate the system of units of the magnetic field intensities. All soils have magnetic susceptibility to some degree. A typical susceptibility value for rock is 10, with subsoils 10-100 and topsoils 100-1000. Once rocks, clays and soil have been heated the susceptibility can increase to 1000-5000. Figure 1 shows the susceptibilities of different subsoil geologies. It is clear that the sand has higher susceptibility compared to the other materials and that topsoil, pits and ditches tend to have higher susceptibility than the subsoil.

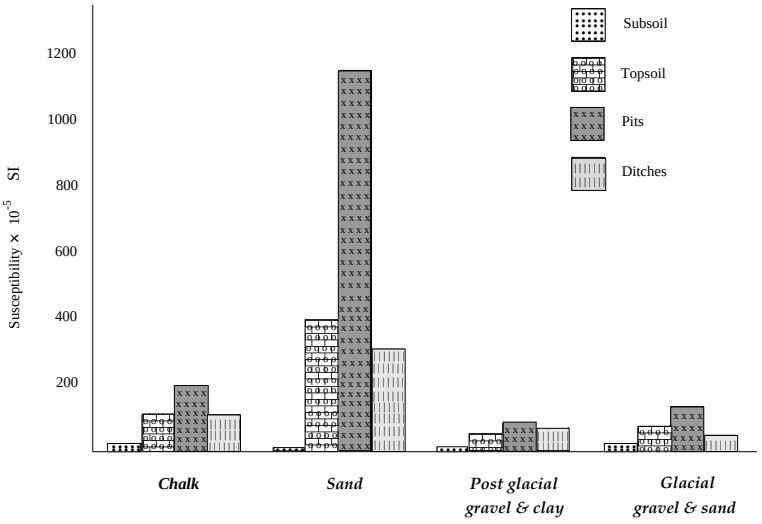


Figure 1. Susceptibility of different types of soil.

Archaeological structures beneath the ground surface cause small local changes, called anomalies, in the Earth's magnetic field due to different magnetic susceptibilities compared to the surrounding ground. The measured data are usually preprocessed and displayed as images then manually interpreted by experts (Scollar 1970). The archaeological interpretation of magnetic anomalies is very difficult for several reasons: first, because it is a $2D$ projection of a $3D$ world. Secondly, the anomalies of nearby structures may be superimposed and finally, there is always a large amount of noise in the measurement caused by the susceptibility variance of the topsoil, by geological structures and by other sources (Eder-Hinterleitner et al. 1995). Although the archaeologist can estimate whether there is an anomaly of an archaeological structure or not and the probable kind of structure, the image of such anomalies itself does not give any detailed description of the buried object, such as its depth and extent.

In most current approaches, a simplified single-layer model is used where the archaeological features are assumed to be at a known depth below the surface, and to have the same vertical extent. Only the magnetic susceptibility is allowed to vary. In real situations, quantities such as depth and extent are unlikely to be known. Any statistical analysis incorrectly fixing such parameters is likely to produce biased estimates of magnetic susceptibility which could again mislead site interpretation. To determine vertical stratigraphy a cylindrical core of material can be obtained using a soil borer, which can penetrate over 1m deep in stone free ground. Often the strata in the core show no varia-

tion in colour or texture, but an analysis of the magnetic susceptibility can differentiate between the separate periods of the now departed landscape. Once collected, the core is passed through a detector coil, allowing readings of the susceptibility to be made along the length of the sample. The use of borehole data for stratigraphy has been studied by Allum (1997) and Allum et al. (1999) and single-layer reconstruction from surface magnetometry data by Aykroyd et al. (2001). This paper concentrates on the most important parameters after the magnetic susceptibility, which are the depth and extent. The aim is to achieve improved estimates for the magnetic susceptibility of the features as well as present a better description of the subsurface.

The two applications to be considered can both be defined as *inverse problems* – see, for example, Ribés and Schmitt (2008) for a general introduction. Further, when written in their usual form both applications are referred to as *linear* inverse problems. Although a full description will be given later, here consider the simple example of linear regression with $n \times 1$ data vector $\mathbf{Y}$, $n \times m$ design matrix X and $m \times 1$ vector of model parameters $\boldsymbol{\beta}$. Assuming independent and identically distributed Gaussian errors would naturally lead to the usual maximum likelihood solution of $\hat{\boldsymbol{\beta}} = (X^T X)^{-1} X^T \mathbf{Y}$. However, it is clear that if the number of parameters is greater than the number of data values, that is $m > n$, then there is no solution. Also, if X is multi-collinear then the solution may change dramatically for only small changes in data. These difficulties characterise an *ill-posed* inverse problem as one in which some combination of the following conditions are true: (i) there is no solution, (ii) there is more than one solution, and (iii) the solution does not depend continuously on the data. In a traditional approach a unique solution and numerical stability are usually achieved by imposing, sometimes arbitrary, additional constraints with point estimation achieved through constrained deterministic optimisation. In this paper, two alternative approaches to the solution of inverse problems are taken. In the first, the problem is re-parameterized to produce a well-posed, but non-linear, inverse problem, whereas in the second carefully chosen prior knowledge is incorporated into the estimation process as part of a Bayesian model and estimation achieved via a stochastic Markov chain Monte Carlo (MCMC) algorithm. A comprehensive review of Bayesian methods for inverse problems, from a mathematical point of view, can be found in Stuart (2010), and examples of applications in industrial process monitoring are given in Watzenig and Fox (2009) and West et al. (2005), and in medical imaging in Li et al. (2011) and West et al. (2004).

The rest of the paper is organised as follows. The next section looks at the physical and statistical modelling for the two inverse problems, and briefly outlines the MCMC algorithm used for estimation. Section 3 describes the data to be analysed. Section 4 considers the first inverse problem, analysis of borehole data, to identify key parameters. In Section 5, this information is then included as part of the model of the surface data to estimate locations of the underground features. Finally, some conclusions are given in Section 6.

2. Physical and Statistical modelling

2.1 Borehole data

A core sample is obtained using a soil borer and although the strata in the core often show no variation in colour or texture, an analysis of the magnetic susceptibility can differentiate between the different layers. Once collected, the core is passed through a magnetic detector coil, allowing readings of the susceptibility to be made along the length of the sample.

Let the output readings be denoted $\mathbf{z} = \{z_i : i = 1, \ldots, n\}$ recorded at positions $\{\mathbf{t}_i : i = 1, \ldots, n\}$ along the core length. Although the core susceptibility, $x(s)$, is a continuous function of position, s, along the core length, it is usual for estimation purposes to consider

the core partitioned into elements. Let these discretized susceptibilities be denoted by $\mathbf{x} = \{x_j : j = 1, \ldots, m\}$ at positions $\{\mathbf{s}_j : j = 1, \ldots, m\}$.

The susceptibility is measured by detecting small changes in the inductance of the coil as the core is passed through it. Since the detector coil is sensitive to the susceptibility across an extended section of the sample, the instantaneous reading indicates not the value at a sharply defined point, but a weighted average of the values over an extended range. Hence the expectation of the observed reading at position i is then given by:

$$\mu_i(\mathbf{x}) = E[z_i] = \sum_{j=1}^{m} x_j h_{ij} \qquad \text{for} \quad i = 1, \ldots, n \tag{1}$$

where the appropriate form of the spread function, h_{ij} is given by (Allum et al. 1999)

$$h_{ij} = h(d = ||\mathbf{s}_j - \mathbf{t}_i||) = \frac{1}{4w}\left[\frac{(d+w)}{\sqrt{(r-a)^2 + (d+w)^2}} - \frac{(d-w)}{\sqrt{(r-a)^2 + (d-w)^2}}\right], \tag{2}$$

where a is the radius of the core, and r and $2w$ are the radius and length of the coil. For the equipment used here, $a = 17.5$mm, $r = 40$mm and $w = 5$mm.

In practice, the observed measurements are subject to error from various sources. Assuming an additive Gaussian error model the conditional distribution of the data given the truth is

$$z_i|\mathbf{x} \sim N(\mu_i(\mathbf{x}), \sigma^2) \qquad \text{for } i = 1, \ldots, n \tag{3}$$

with likelihood

$$\pi(\mathbf{z}|\mathbf{x}) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(z_i - \mu_i(\mathbf{x})\right)^2\right\}. \tag{4}$$

Due to the form of Equation (1), and since typically the number of unknowns, m, is much greater than the number of observed readings n, estimation of $\mathbf{x}$ is an ill-posed linear inverse problem which cannot be reliably solved using the likelihood alone. However, rather than resorting to some form of smoothing prior as part of a Bayesian analysis (as in Allum et al. , 1999) a re-parametrization, in terms of a few key parameters, is proposed. The aim being that the re-parametrized model will contain few parameters and hence leads to a well-posed problem in which no prior information is needed.

Suppose that the archaeology occupies a single layer at unknown depth and that it is of unknown extent and susceptibility. Also, that the susceptibility of the remaining core is unknown. Further suppose that data collection starts before the core enters the recording device, and continues for a short time after it has passed completely through. Figure 2 shows a diagram of a simple core along with model parameters and a susceptibility profile.

Let d_1 be the distance over which measurements are taken before the core enters the data recorder. It is assumed that the susceptibility here is exactly zero. The first part of the core, of length d_2, has susceptibility x_B which represents a background level. The archaeological feature, with susceptibility x_F, is contained in the next part which is of length d_3. Then there is a second background part which is of length d_4 and has susceptibility x_B. Finally, there is a distance d_5 of zero susceptibility before the data recording stops.

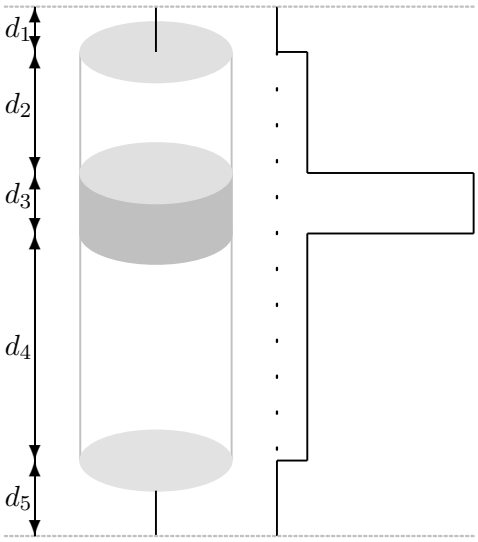


Figure 2. Diagram of the extracted core and corresponding susceptibility profile.

Let the susceptibility along the core be denoted x, and at location t $(0 \leq t \leq l)$

$$x(t) = \begin{cases} 0 & 0 \leq t < d_1 \\ x_B & d_1 \leq t < d_1 + d_2 \\ x_F & d_1 + d_2 \leq t < d_1 + d_2 + d_3 \\ x_B & d_1 + d_2 + d_3 \leq t < d_1 + d_2 + d_3 + d_4 \\ 0 & d_1 + d_2 + d_3 + d_4 \leq t \leq d_1 + d_2 + d_3 + d_5 = l. \end{cases} \tag{5}$$

This re-expresses the parameter set as $\Theta = \{d_1, d_2, d_3, d_4, d_5, x_F, x_B\}$. It would be possible to include prior information regarding the values of these parameters based on previous investigations, or expert opinion, but here uniform prior densities have been used for all parameters hence our aim is to estimate these by maximising the likelihood

$$\pi(\mathbf{z}|\mathbf{\Theta}) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(z_i - \mu_i(\mathbf{\Theta}))^2\right\}. \tag{6}$$

2.2 Surface data

Surface data are collected using a gradiometer which measures the local magnetic field gradient using two sensors about 50cm apart, one above the other. Suppose that n measurements $\mathbf{y} = \{y_i : i = 1, \ldots, n\}$ are recorded at locations $\{\mathbf{s}_i : i = 1, \ldots, n\}$. In general, the locations will form a regular square grid, but the proposed approach is valid for general locations.

In our model, the archaeological site subsurface is divided into L co-planar *layers*, with each layer further subdivided into m 3D rectangular pixels. Previous approaches have used only a single layer. The distance from the surface to the mid-plane of layer l is denoted d^l. Within a layer the centres of the pixel are denoted $\{\mathbf{t}_j^l : j = 1, \ldots, m\}$ and the corresponding magnetic susceptibility values by $\mathbf{x}^l = \{x_j^l : j = 1, \ldots, m\}$. The complete sets of susceptibilities will be denoted $\mathbf{x}$.

The measurement at location s_i is influenced by all components of $\mathbf{x}$, but by an amount dependent on the distance between. The expected magnetometer reading is the superpo-

sition of the influence of all the susceptibility values in all L layers

$$\mu_i(\mathbf{x}) = \sum_{l=1}^{L} \sum_{j=1}^{m} x_j^l h_{ij}^l \tag{7}$$

where the point spread function (see Aykroyd et al. 2001) has negative as well as positive values, is symmetric in the West-East direction, but not in the North-South direction. Its maximum value is not at the origin, but at a distance to the south which depends on the depth of the feature. The noise component of the data is appropriately represented by an additive Gaussian error model (Allum, 1997) with zero mean and variance σ^2. The corresponding likelihood $\pi(\mathbf{y}|\mathbf{x})$, is the joint distribution of the data given a susceptibility distribution $\mathbf{x}$

$$\pi(\mathbf{y}|\mathbf{x}) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i - \mu_i(\mathbf{x})\right)^2\right\}. \tag{8}$$

The form of Equation (7), and since the number of unknowns, $L \times m$, is much greater than the number of observed readings, n, this is another ill-posed linear inverse problem which, again, cannot be reliably solved from the likelihood alone. The traditional approach to solving such problems is to impose additional constraints through *regularization*, which is mathematically equivalent to the approach of penalized likelihood. Here, as in many other inverse problems, the Bayesian framework using prior information to regularize is adopted. This gives an intuitive hierarchical structure to the modelling with results which can easily be interpreted. The posterior density of a susceptibility distribution given the data is the result of combining a prior density with the likelihood using Bayes' theorem:

$$\pi(\mathbf{x}|\mathbf{y}) = \frac{\pi(\mathbf{y}|\mathbf{x})\pi(\mathbf{x})}{\pi(\mathbf{y})}.$$

The unknown susceptibility distribution $\mathbf{x}$ is modelled by a prior density $\pi(\mathbf{x})$. This prior should be designed to give high probabilities to susceptibility distributions which agree with our *aprior* expectations and low probabilities to unacceptable susceptibility distributions. Here we expect neighbouring locations to have similar susceptibilities and hence the prior should encourage smoothness. Although any form of probability density function which suitably describes smoothness can be incorporated into the estimation, it is helpful to have a representation which is easy to describe and interpret. In image analysis, this prior information is commonly quantified in the form of a *Gibbs distribution*

$$\pi(\mathbf{x}) = \frac{1}{z(\boldsymbol{\beta})} e^{-U(\mathbf{x};\boldsymbol{\beta})} \tag{9}$$

where $z(\boldsymbol{\beta})$ is the normalising constant such that $z(\boldsymbol{\beta}) = \int_{\mathbf{x}} e^{-U(\mathbf{x};\boldsymbol{\beta})} d\mathbf{x}$. The constant $\boldsymbol{\beta}$ is a vector of non-negative smoothing parameters reflecting the degree of correlation between neighbouring pixels and determining the level of influence the prior has in the posterior distribution. The origins of this type of distribution is in statistical mechanics, to describe gas thermodynamics for example, and hence the function $U(\mathbf{x};\boldsymbol{\beta})$ is often called the *energy function*, and is designed to assign high probabilities to the expected configurations which have *low energy*. Similarly, the parameter $\boldsymbol{\beta}$ is related to *temperature*, and so increasing $\boldsymbol{\beta}$ increases the energy of the configuration.

The three dimensional prior model and subsequent estimation procedures are a direct generalization of those developed for two dimensional applications (see, for example,

Aykroyd et al. 2001). The prior beliefs about the subsurface of an archaeological site are that neighbouring pixels within a feature have similar susceptibilities and sharp changes exist only at the edge of a feature. This applies also when the same feature extends vertically into neighbouring layers.

Hence in the multi-layer model the energy, $U(\mathbf{x};\boldsymbol{\beta})$, has two components which represent the *within-layer energy* and the *between-layer energy*, that is

$$\text{Total energy} = \text{within-layer energy} + \text{between-layer energy}$$

$$U(\mathbf{x};\boldsymbol{\beta}) = \sum_{l=1}^{L} \beta_W^l \sum_{i\sim j} w_{i,j}\, \phi(x_i^l - x_j^l) + \sum_{l=1}^{L-1} \beta_B^l \sum_{j=1}^{m} \phi(x_j^l - x_j^{l+1}) \tag{10}$$

where $i \sim j$ indicates all pairs of neighbouring pixels, and $w_{i,j}$ gives different weights to different neighbouring pairs. Here a *second-order* neighbourhood is chosen containing the horizontal and vertical nearest neighbours (with $w_{i,j} = 1$) and the four diagonal nearest neighbours (with $w_{i,j} = 1/\sqrt{2}$). This is considered to be the smallest neighbourhood capable of capturing patterns seen in practice. The parameter set $\boldsymbol{\beta} = (\boldsymbol{\beta}_W, \boldsymbol{\beta}_B)$ is accordingly divided into two groups, smoothing within layers, $\boldsymbol{\beta}_W$, and smoothing between layers, $\boldsymbol{\beta}_B$. In this formulation it is expected that the smoothing parameters will vary from one layer to another, and also could be allowed to vary within layer – though only the former is considered here. The *potential* function ϕ measures the difference between susceptibility values. The most commonly chosen quadratic, corresponding to a Gaussian prior, often over-smooths susceptibility distributions and so many other *implicit discontinuity* or *edge-preserving* alternatives have been proposed. One such example is the absolute value, which corresponds to a double-exponential or Laplace prior. Here we use the implicit discontinuity prior recommended in Aykroyd et al. (2001).

2.3 Estimation algorithms

A simple MCMC algorithm is used to produce approximate samples from the posterior distribution by simulating a Markov chain with the required distribution as its equilibrium distribution. The use of such methods for parameter estimation, and more general density exploration is widespread. See Besag et al. (1995) and Chapter 4 of Voss (2013) for general details and Aykroyd et al. (2001) for an archaeological example.

In the two uses of MCMC estimation in this paper the parameters to be considered are of quite different natures. For the borehole model the parameters are a mixture of distances and susceptibilities, $\Theta = \{d_1, d_2, d_3, d_4, d_5, x_F, x_B\}$, whereas in the subsurface reconstruction they are all susceptibilities, $\Theta = \{x_j : j = 1, \ldots, m\}$. Despite this the algorithms share a common structure. Let the set of model parameters be labelled $\Theta = \{\theta_1, \ldots, \theta_p\}$, The general approach is to consider single parameter changes; let the parameter being considered be θ_i. A proposed new value θ_i' is drawn from a proposal distribution, $q_i(\theta_i'|\theta_i)$. Let the set of parameters containing the proposed value be $\Theta' = \{\theta_1, ..., \theta_{i-1}, \theta_i', \theta_{i+1}, ..., \theta_p\}$. The proposal is accepted, and the parameter value updated accordingly with probability

$$\alpha(\Theta', \Theta) = \min\left\{1, \frac{p(\Theta'|Y) q_i(\theta_i'|\theta_i)}{p(\Theta|Y) q_i(\theta_i|\theta_i')}\right\} \tag{11}$$

otherwise it is rejected and the previous value retained. Here most of the parameters are non-negative, hence a negative proposal is immediately rejected and no change is made – maintaining detailed balance. A non-negative proposal is accepted with the probability

given in Equation (11). Note that if the proposal distributions are symmetric, $q_i(\theta_i'|\theta_i) = q_i(\theta_i|\theta_i')$, then their ratio in Equation (11) cancels, but that there is no need for the proposal distributions to be the same for the different parameters.

If the algorithm is designed carefully, then as the iterations progress the current parameter set does not depend on the starting values, and can be treated as a correlated sample from the posterior distribution. Key issues then become how to judge when this initial transient behaviour has ended, and the chain is in equilibrium, and how many iterations to perform to have a sufficiently large sample for reliable estimation. Once the sample has been generated from the posterior distribution, the sample mean is used to estimate the posterior mean, and sample percentiles to estimate confidence bounds. It is also possible to inspect low dimensional marginal distributions, and functions of the simulated values.

3. Data Example

Figure 3 shows the magnetometer readings across *the Park, Guiting Power*, which is a late iron-age farmstead. As well as the full site data, an extract which will be analysed here, is also shown. The approximate horizontal location, size and shape of many features can be seen reasonably well. In particular a diagonal linear ditch towards the top, a rectangular boundary ditch surrounding various collections of circular pits and post-holes.

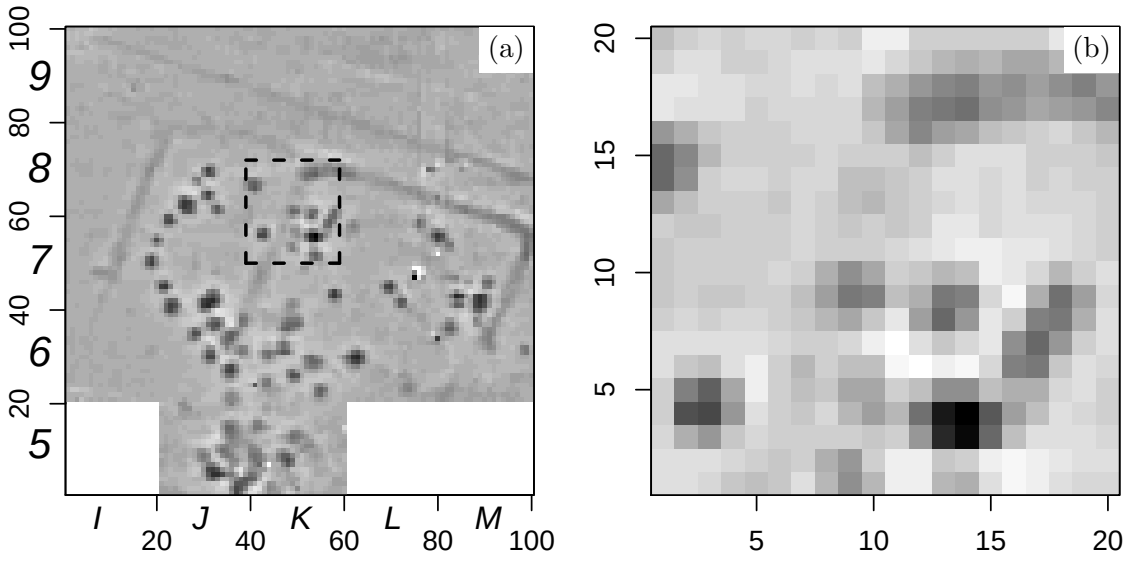
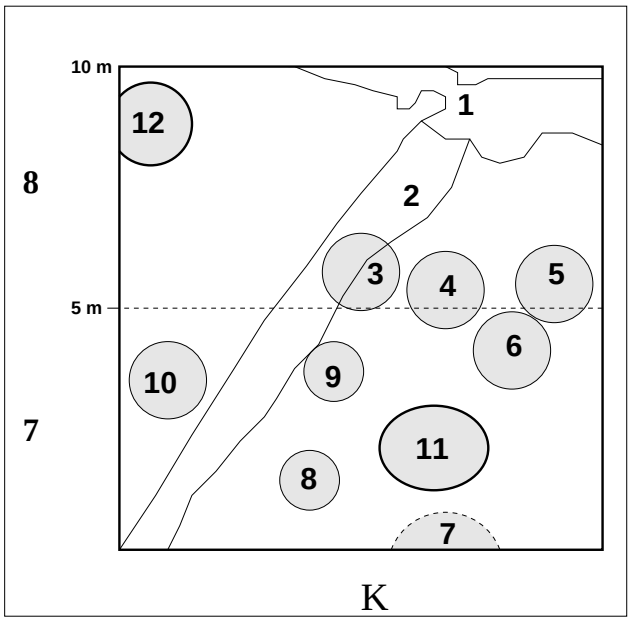


Figure 3. (a) Magnetometry survey of the Park, Guiting Power showing grid labels and distances, and (b) grid selected for analysis.

It is usually assumed that the features are buried at the same depth from the modern site surface, but that they have different susceptibilities and extents. The assumption of constant depth is not unreasonable as ancient structures are often levelled to the prevailing ground level and any pits or ditches in-filled to the same level. At later stages the whole site is uniformly covered with topsoil leading to the common modern surface level. Figure 4 shows a map along with values of the susceptibilities and extents of the selected grid features measured when this part of the site was excavated.

The strength of the surface magnetic reading depends on the depth of the features. If the other parameters are fixed, the magnetometer records high values for shallow features and low values for deep features.



Feature	No.	Vertical Extent (m)	Average Susceptibility (SI × 10^{-3})
Ditches	1	0.7	1.1
	2	0.7	1.5
Pits	3	0.8	1.8
	4		
	5		
	6		
	7		
	8		
	9		
	10		
	11	**1.6**	1.1
	12	**1.5**	1.0

Figure 4. Feature map, susceptibilities and extents.

A key part of our proposed approach is the combination of surface magnetic data with borehole data, but unfortunately no such datasets are currently available – in fact it is expected that this paper will motivate the future collection of combined datasets. Hence we shall use the excavation records to generate phantom cores and corresponding borehole data. Suppose five boreholes are taken at feature locations 1, 2, 3, 11 and 12 as a representative set covering all feature types. The boreholes penetrate all features and go a further 50cm beneath the deepest. This leads to core lengths of 2.4m.

To generate data we assume that the first 30cm corresponds to modern topsoil deposited after the archaeology was levelled to the prevailing ground level. There is then the archaeological feature of vertical extent dependent on the feature. Finally, there is a depth of background soil taking the core to the chosen $2.4m$. All background susceptibilities have been fixed at 0.2 (SI$\times 10^{-3}$), but the feature susceptibility are taken from the excavation records. The result of this is a perfect step function, which will be called the *truth*, and which is the goal of the estimation process. Natural soil variability and local mixing, however, will introduce departures from this ideal. In an attempt to capture some elements of this complex process, small variance Gaussian noise and small-scale Gaussian blur are combined with the step function. The resulting susceptibility profiles are shown as solid lines in Figure 5 for Core 1 and Core 11. This natural variability and mixing is only included in the data generation model and will not be included in the analysis model.

The borehole cores are placed in the middle of a support tray which is passed through the magnetic coil. The recording starts as soon as the support tray enters the detector and finishes when the tray leaves the detector. Hence the first and last readings correspond to an empty detector. Finally, to produce measurement data these susceptibility profiles are combined with the models defined in Equations (1)-(3).

Figure 5 also shows two example datasets for (a) Core 1 and (b) Core 11 as points. The solid lines show the naturally varying susceptibility profiles and the dashed lines are the true susceptibilities. Table 1 shows the true parameters for all 5 borehole cores for the model illustrated in Figure 5. The parameters d_1 and d_5 correspond to the readings when no core is in the detector, d_2 is the depth of the feature layer, d_3 is the extent of the feature and d_4 is the depth of background below the feature layer. The feature and background susceptibilities are the parameters x_F and x_B. The key parameters are d_2, d_3 and x_F,

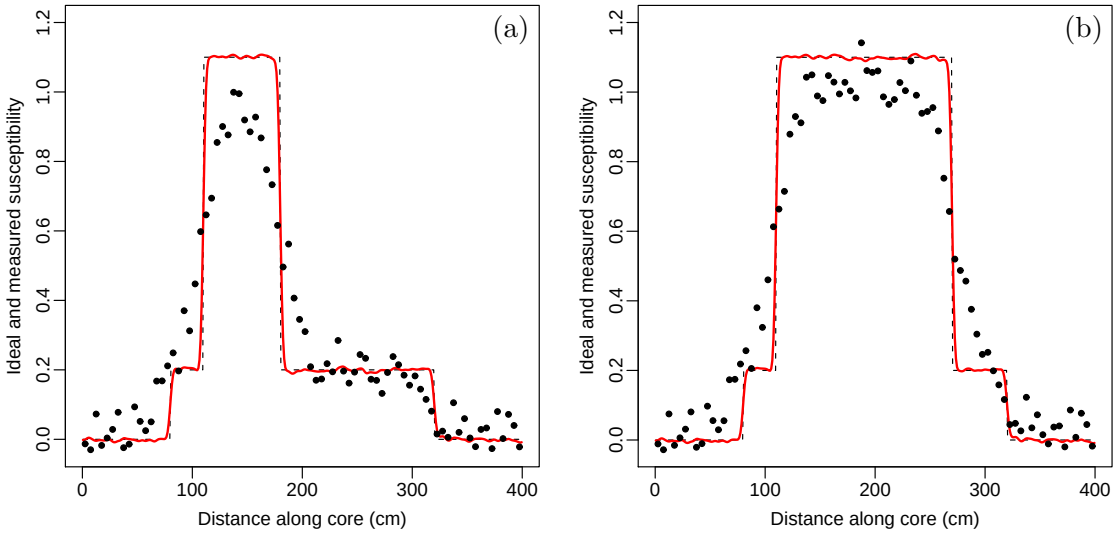


Figure 5. Truth (dashed line) and naturally varying susceptibility profile (solid line) along with data (points) from feature locations (a) Core 1 and (b) Core 11.

whereas d_1 and d_2 can be measured with high reliability, the background susceptibility x_B can be measured by taking a separate sample, and d_4 can be calculated by subtraction given d_2 and d_3 and the length of the core.

Table 1. True values for depth parameters, background and feature susceptibility.

	Distances					Susceptibilities	
Core	d_1 (cm)	d_2 (cm)	d_3 (cm)	d_4 (cm)	d_5 (cm)	x_F (SI$\times 10^{-3}$)	x_B (SI$\times 10^{-3}$)
1	80	30	70	140	80	1.1	0.2
2	80	30	70	140	80	1.5	0.2
3	80	30	80	130	80	1.8	0.2
11	80	30	160	50	80	1.1	0.2
12	80	30	150	60	80	1.0	0.2

4. Stratigraphic Analysis from Borehole Data

As a first stage, the five distance and two susceptibility parameters are estimated using the $1D$ borehole data. These estimates will then be used to describe prior information in the main site reconstruction, which will be the second stage.

Figure 6 shows histograms and kernel density smoothed empirical posterior distributions for the parameters from Core 1, along with a point-wise posterior median of the susceptibility profile with 95% credibility interval. Table 2 summarizes the posterior distributions for all five boreholes using the posterior means and standard deviations. We shall now consider each of the parameters in turn. Although not evident from either Figure 6 nor Table 2, the most important feature of these results is the extreme correlation between some parameter estimates. For example, the correlation between $\hat{d}_1$ and $\hat{d}_2$ is -0.97, between $\hat{d}_4$ and $\hat{d}_5$ is -0.94, and between $\hat{d}_3$ and $\hat{x}_F$ is -0.79. Also, note that the estimates of d_1 have a negative bias and of d_2 a positive bias. For d_4 the estimate is always less than the true value while for d_5 the estimate is greater than the true value. In the case of d_3

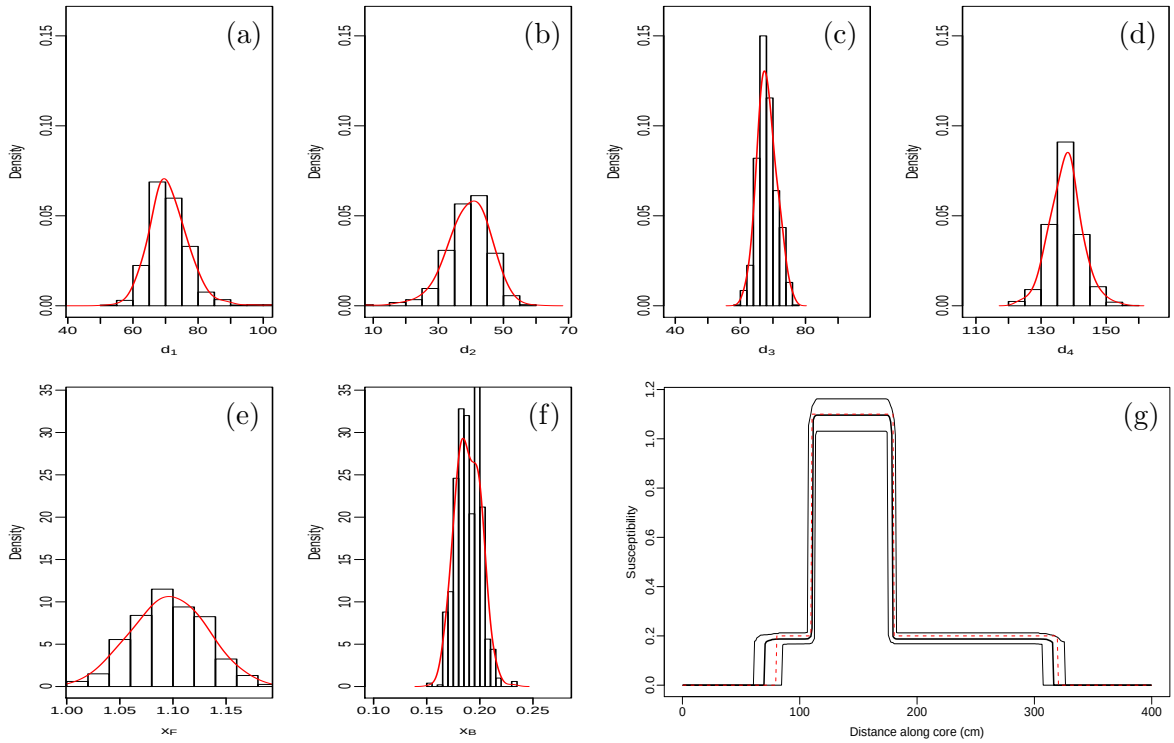


Figure 6. Core 1: Posterior distributions of (a) initial blank, d_1, (b) depth, d_2, (c) extent, d_3, (d) final blank, d_4, (e) feature susceptibility, x_F, (f) background susceptibility, x_B, and (g) susceptibility profile along core.

Table 2. Posterior means and, in parentheses, standard deviations for all distances and susceptibilities.

	Distances					Susceptibilities	
Core	d_1 (cm)	d_2 (cm)	d_3 (cm)	d_4 (cm)	d_5 (cm)	x_F (SI$\times 10^{-3}$)	x_B (SI$\times 10^{-3}$)
1	71.2	39.1	68.2	137.6	84.0	1.10	0.19
	(6.1)	(7.0)	(2.9)	(4.9)	(4.7)	(0.03)	(0.01)
2	71.2	38.8	69.3	137.2	83.5	1.49	0.19
	(5.7)	(6.0)	(1.8)	(5.5)	(5.6)	(0.03)	(0.01)
3	72.5	37.7	78.9	125.7	85.1	1.79	0.20
	(5.8)	(6.1)	(1.5)	(5.1)	(5.2)	(0.03)	(0.01)
11	75.7	35.1	158.8	41.6	88.9	1.08	0.24
	(9.4)	(8.6)	(4.2)	(7.3)	(8.7)	(0.01)	(0.09)
12	75.2	35.0	149.7	51.7	88.4	0.97	0.22
	(6.1)	(6.3)	(3.2)	(5.6)	(6.0)	(0.01)	(0.04)

and x_F, recall these are the width and susceptibility of the feature, the width is estimated slightly too small and the susceptibility also slightly too small – that is there is a tendency to underestimate the archaeological importance of the feature. Hence it is clear that there is insufficient information to separately estimate all parameters. Given that it is possible to record d_1 and d_5 through simple and accurate physical measurement, however, it is reasonable to assume that they are known and the estimation process used for a reduced set of parameters.

Table 3 shows the posterior estimates for the reduced set of parameters. All cases show similar estimates of the feature depth, d_2, of between 29.0cm and 29.7cm, which are now very close to the true values and with dramatically reduced variability. Although these

Table 3. Posterior means and, in parentheses, standard deviations for the reduced set of distances and susceptibilities.

	Distances			Susceptibilities	
Core	d_2 (cm)	d_3 (cm)	d_4 (cm)	x_F (SI$\times 10^{-3}$)	x_B (SI$\times 10^{-3}$)
1	29.7	68.7	141.5	1.10	0.19
	(1.8)	(3.3)	(1.9)	(0.03)	(0.01)
2	29.7	69.4	140.9	1.49	0.19
	(1.1)	(1.9)	(1.2)	(0.03)	(0.01)
3	29.6	80.1	130.3	1.78	0.19
	(0.8)	(1.4)	(0.9)	(0.02)	(0.01)
11	29.4	160.4	50.2	1.08	0.20
	(1.4)	(2.4)	(1.3)	(0.01)	(0.02)
12	29.0	151.2	59.8	0.97	0.20
	(1.7)	(2.8)	(1.6)	(0.01)	(0.02)

appear slightly biased, as the true value is 30cm, given the sampling interval of 5cm and the standard deviations of about 1cm it is not considered important. For d_3, the feature extent, there is a slight reduction in bias and variability. For d_4, as for d_2, there is a substantial improvement due to the removal of d_5 from the estimation. There is virtually no change in the estimates and variability of the susceptibility parameters.

Although this gives a new method for analysing borehole core data – which will be useful in its own right – recall, that the reason for this analysis is to suggest prior distributions for depth, extent and susceptibility for use in the analysis of the surface data. Hence, from the above it is reasonable to assume a common depth for the whole site of about 30cm. An alternative approach would be to define a Gaussian depth prior with mean of 30cm and standard deviation 2cm.

For extent the posterior distributions are rather different, in particular in terms of the posterior mean values. Boreholes 1, 2 and 3 have estimates of about 70cm, while for boreholes 11 and 12 the posterior means are 160cm and 151cm respectively. Clearly, although this gives very definite information about extent at the borehole locations, it is difficult to infer the extent of features further away without some other information. Hence we choose only to use this information at the borehole locations. The background susceptibility posterior means are all close to each other, and hence a suitable fixed value might be 0.2, or a prior distribution with this mean and small standard deviation could be used. Finally, the feature susceptibility ranges from about 1.0×10^{-3} to 1.8×10^{-3}, but with each estimate having a very small standard deviation of about 0.02×10^{-3}. Clearly, we would not want to fix the susceptibility of the features as a single value, but it might be reasonable to describe their susceptibility by a mixture distribution reflecting the susceptibilities of likely features or a vague uni-modal prior could be used which covers the range of values seen in the various boreholes. For example with a mean of 1.4×10^{-3} and standard deviation of 0.2×10^{-3}.

5. Fusion of Depth and Surface Data

5.1 General

Although Bayesian approaches to subsurface reconstruction have been successfully used, showing important improvements in both spatial and gray-scale resolution, a more realistic subsurface model is needed. This is because in previous models various physical parameters,

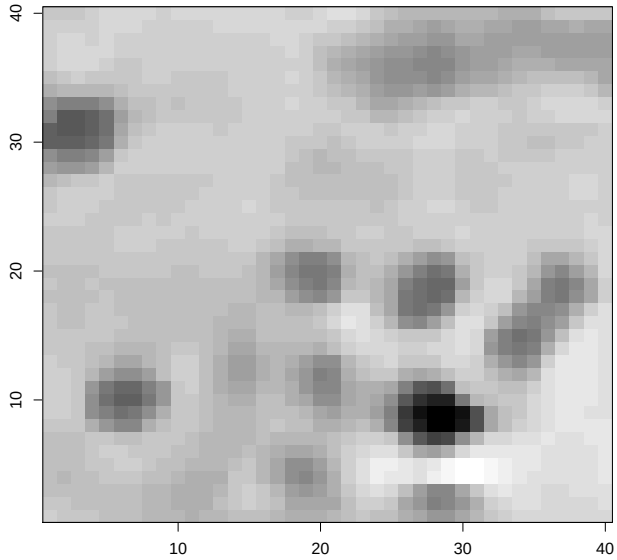


Figure 7. Standard single layer reconstruction.

such as depth and extent, are currently assumed known, when in fact these quantities should be key outputs of an analysis. In this section, the multi-layer model will be used with stratigraphic information supplied from the borehole analysis. The rationale for the technique is that having knowledge of some of the parameters allows available data to estimate remaining parameters more reliably. This should allow not only the detection of a typical anomaly, as in standard magnetic surveys, but should also provide "diagnostic" information about the main physical properties such as size, depth or an idea about the extent of the buried features. The development of such new approaches brings the promise of much improvement to subsurface estimation.

5.2 Reconstruction without prior information

In this section a simple single layer analysis is considered first, and then the approach will be extended to the multi-layer model. In the former, the main physical parameters are fixed using the borehole data, whereas in the latter the borehole information is used to guide the susceptibility estimation. Note that this is already an improved model compared to the common approach of subjectively fixing these physical parameters.

Suppose that there is an archaeological feature present as a whole in one horizontal layer. In a previous single layer analysis conducted on part of the the same data (Allum 1997), it was found that $\hat{\beta} = 623$, $\hat{\delta} = 175$, $\hat{\sigma} = 0.33$ gave good reconstructions. (Note that δ is a *threshold* parameter in the prior distribution.) As a baseline, Figure 7 shows the reconstruction using these parameter values for a single layer. Comparing this to the feature map in Figure 4, all features are evident, but there is no information regarding their true depth and extent. Also, having fixed depth and extent has the knock-on effect of influencing the susceptibility estimates. In particular notice the dark areas in the top-left and bottom-right, which correspond to features 11 and 12. Given that these actually correspond to deep pits, but that the susceptibility has been estimated assuming a shallow extent, it is likely that this value has been overestimated. Also the white area below feature 12 is likely to be a remnant of the spread function, evident in the data, which has not been corrected.

The straightforward first extension is to repeat the analysis using two layers. So the subsurface is partitioned into two layers of equal extent – without additional information there is no reason to do otherwise. The susceptibility estimates of the two layers, without

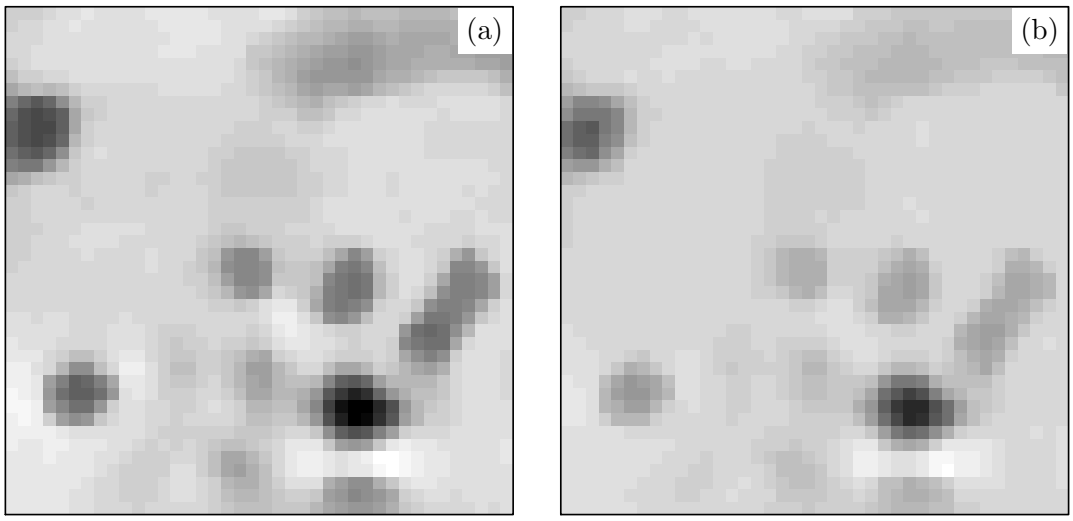


Figure 8. Two-layer reconstruction without prior information: (a) is the upper layer and (b) the lower layer.

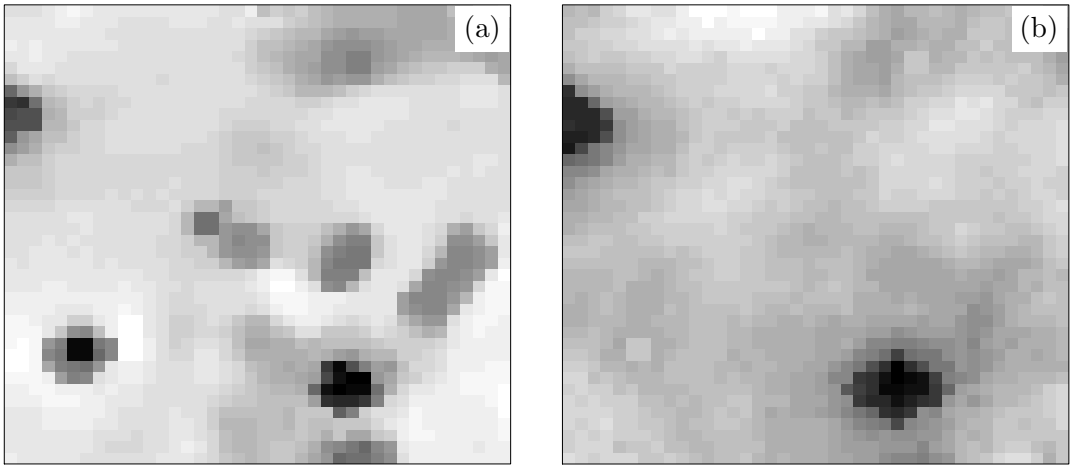


Figure 9. Two layer reconstructions using susceptibility prior information: (a) is the upper layer and (b) the lower layer.

between-layer smoothing, are shown in Figure 8. It can be seen that the estimation in each layer is affected by the other layer with features expected only in the first layer also appearing in the second layer. Given that the between-layer smoothing parameter is set to zero, that is the two layers are independent, this apparent depth correlation is somewhat unexpected.

5.3 Reconstruction with prior information

Although it would be possible to extend the model to many-layers, given the limited range of depths and extents seen across the site, it seems reasonable to continue to define two layers. In contrast, however, to the earlier two-layer reconstruction both the depth, extent and susceptibility information from the borehole data will be used. The top layer is fixed at 30cm below the surface with an extent of 70cm and the second layer with an extent of 90cm. Note that in the previous reconstruction the two layers were assumed of equal extent. The borehole diameter is assumed to be about 35mm which is equivalent to a 2×2

block of pixels. Therefore, at each of these locations, the susceptibility will be fixed, at the background or foreground value as appropriate. Elsewhere the susceptibility is unknown, and is to be estimated. It seems reasonable however to assume that the susceptibility is similar to that encountered in the various boreholes. This is described using a two-part mixture of Gaussian distributions, describing the background and feature susceptibilities with means and variances obtained from the first stage and with a fixed mixing proportion of 10% feature. Although it would be possible to extend the model to allow this mixing proportion to be estimated, in separate experiments the reconstructions have not been found to be sensitive to small changes in this value.

Figure 9 shows the new two-layer reconstruction using the susceptibility information derived from the borehole data as prior information. The estimates look reasonable, since the second layer estimate is not affected by the features truly in only the first layer. Pits 11 and 12 appear in the image of the second layer as their extents are approximately twice that of the other features.

6. Conclusion

Any archaeological magnetometry data set can be explained by an infinite number of spatial susceptibility distributions. This is because we typically try to estimate many parameters from highly correlated data – increasing the sampling interval adds little new information. Instead, identifying the most plausible models can only be achieved by including additional knowledge as apriori constraints. For example, the number of acceptable models can be reduced markedly if the magnetic signals are generated by features for which the depths and extents are known.

While the surface magnetometry data provides a good indication of the horizontal location of features, it can not resolve the depth distribution on its own. An observed magnetic signal, in the presence of noise, can not uniquely define the shape and depth of the feature producing the signal because an infinite number of combinations of susceptibility and depth can produce essentially the same given magnetic readings. Strong features seen in the magnetic survey data may either represent deep high susceptibility objects, or could equally be produced by shallower low susceptibility objects. If a feature has low susceptibility, or is deeply buried, then it will produce a weaker signal than if it has high susceptibility, or is buried near the surface.

As well as being physically *confounded*, their estimates are also highly correlated with the level of smoothing used in the reconstruction. Increasing the level of smoothing usually leads to the distortion of feature boundaries, especially for weak signals, whereas decreasing smoothing does not help as the signal can be hidden by noise.

Although other site survey methods, such as ground penetrating radar, seismology or electrical tomography, could be used to give complimentary information, all surface based methods suffer from similar confounding of depth and physical properties. The most fruitful source of the necessary information is direct observation, but it is clearly impossible to obtain such information across a full site. It is possible, however, to gather information at a few key locations. The use of extracted borehole cores can provide exactly the information needed, that is estimates of depth, extent and susceptibility values.

Although several papers use data from multiple sensing systems none use it to produce a single underground reconstruction. For example, Chianese et al. (2004) collect magnetic surface data and 3D radar images but make no attempt to combine the two. Similarly, Leopold et al. (2010) record aerial photographs, electrical resistance images, magnetic data and ground penetrating radar images, but only combine these sources by overlaying the data. Argote et al. (2009) perform 3D estimation, with very limited success, but carefully match reconstruction grid size and data dimensions to produce a well-posed problem and

use an approximation for the spread function. Also, as they state, this will only work in cases where feature and background are highly contrasting – which is not usually the case. Traditionally it has not been possible to combine such diverse types of information, but the Bayesian framework provides a natural way to incorporate partial prior knowledge into the estimation process.

This paper proposes an approach for full surface estimation from archaeological magnetometry data which makes efficient use of available data. The approach is based on the concept of exploiting the information gained from the statistical analysis of the borehole data to achieve more precise estimates of the physical parameters of the whole archaeological site. The estimation process has been split into two sequential stages. In the first stage, a parametric model is fitted to the borehole core data and the resulting parameter estimates are used in the prior densities for the second stage estimation. In the second stage, using an MCMC algorithm, all physical model parameters are estimated in a fully Bayesian setting. The approach has produced improved image reconstructions as well as estimation of key physical parameters. The improvements allow a more detailed and useful interpretation of the archaeological site. Further, the general framework can also be used to incorporate other types of data leading to enhanced resolution. This gives the potential to reduce environmental damage on important archaeological sites and, perhaps, eventually allow site interpretation without the need for physical excavation.

Acknowledgements

The authors thank Alistair Marshall of Guiting Manor Amenity Trust, Gloucestershire, UK for allowing us to use the Guiting Power data. Also, we are grateful to the anonymous referees for their helpful comments.

References

Argote, D.L., Tejero, A., Chávez, R.E., López, P.A., Bravo, R., 2009. 3D modelling of magnetic data from an archaeological site in north-western Tlaxcala state, Mexico Journal of Archaeological Science, 36, 1661-1671.

Allum, G.T., 1997. A statistical approach to inverse data problems in archaeological geophysics. PhD Thesis, Department of Statistics, University of Leeds, UK.

Allum, G.T., Aykroyd, R.G., Haigh, J.G.B., 1999. Empirical Bayes estimation for archaeological stratigraphy. Journal of the Royal Statistical Society, Series C, Applied Statistics, 48, 1-14.

Aykroyd, R.G., Haigh, J.G.B., Allum, G.T., 2001. Bayesian methods applied to survey data from archaeological magnetometry. Journal of the American Statistical Association, 96, 64-76.

Besag, J., Green, P. J., Higdon, D., Mengersen, K. 1995. Bayesian computation and stochastic systems, Statistical Science, 10, 1-41.

Chianese, D., D'Emilio, M., Di Salvia, S., Lapenna, V., Ragost, M., Rizzo, E., 2004. Magnetic mapping, ground penetrating radar surveys and magnetic susceptibility measurements for the study of the archaeological site of Serra di Vaglio (southern Italy). Journal of Archaeological Science, 31, 633-643.

Eder-Hinterleitner, A., Neubauer, W., Melichar, P., 1995. Reconstruction of archaeological structures using magnetic prospection, Analecta Praehistorica Leidensia, 28, 131-137.

Le Borgne, E., 1955. Abnormal magnetic susceptibility of the top soil, Annales Geophysique, 11, 399-419.

Leopold, M., Plöckl, T., Forstenaicher, G., Völkel, J., 2010. Integrating pedological and

geophysical methods to enhance the informative value of an archaeological prospection The example of a Roman villa rustica near Regensburg, Germany. Journal of Archaeological Science, 37, 1731-1741.

Li, R., Fahimian, B.P., Xing, L., 2011. A Bayesian approach to real-time 3D tumor localization via monoscopic x-ray imaging during treatment delivery. Medical Physics, 38, 4205–214.

Ribés, A., Schmitt, F., 2008. Linear inverse problems in imaging. IEEE Signal Processing Magazine, July, 84-99.

Scollar, I., 1970. Fourier transformations for the evaluation of magnetic maps, Prospezioni Archeologiche, 5, 9-41.

Stuart, A.M., 2010. Inverse problems: A Bayesian perspective. Acta Numerica, 19, 451-559.

Voss, J., 2013. An Introduction to Statistical Computing: A Simulation-Based Approach. Wiley-Blackwell.

Watzenig, D., Fox, C., 2009. A review of statistical modelling and inference for electrical capacitance tomography. Measurement Science and Technology, 20, 052002 (22 pages).

West, R.M., Aykroyd, R.G., Meng, S., Williams, R.A., 2004. MCMC techniques and spatial temporal modelling for medical EIT. Physiological Measurement, 25, 181-194.

West, R.M., Meng, S., Aykroyd, R.G., Williams, R.A., 2005. Spatial-temporal modelling for electrical impedance imaging of a mixing process. Review of Scientific Instruments, 76, 073703 – 073703-10.