14

Image segmentation techniques for archaeological geochemical data

C. E. Buck*

C. D. Litton†

14.1 Introduction

There are many ways of displaying and analysing spatial data, ranging from relatively simple methods such as dot density plots and greyscales to the mathematically more sophisticated techniques of contouring and image processing. In what follows we shall describe the use of one of the more sophisticated techniques, namely image segmentation. In any situation, the choice of technique must depend upon the nature of the data and the questions posed by the archaeologist. Here we use for illustration the case of the analysis of soil phosphate data. We know that there is considerable variation in the background phosphate level well away from known archaeological sites (see Cavanagh *et al.* 1988): techniques such as dot density may depict this variation (which is of little archaeological importance) as well as describing the on-site variation in which the archaeologist is particularly interested. Moreover, the accuracy of the chemical analysis is important. Since the chemical analysis used for most archaeological soil phosphate surveys is of the quick, cheap but crude type, that is the measurements are relatively inaccurate, this should be borne in mind at the data analysis stage. Using a sophisticated technique such as contouring on data that are subject to relatively large measurement errors is not at all sensible and any results produced may be misleading or unreliable. This point about the high noise level relative to the difference between the on-site and off-site levels associated with phosphate data is often not appreciated by workers more accustomed to techniques such as soil resistivity survey.

Another crucial point to be considered is that the data analysis can proceed by either of two routes. The first attempts to produce a description of observed data usually a visual image. The second postulates a statistical model of the spatial distribution of the phosphate levels and then proceeds, using statistical techniques, to estimate the parameters of the model and so to produce a visual image. The former approach only describes the data obtained whereas the latter provides a means of investigating the underlying process generating the data. Fieller and Flenley (1988) make a similar point in the context of soil particle sizes.

In simple situations where the noise level is low or, equivalently, where there is little variation in the background level, dot density, contouring and greyscales can be very effective in aiding the interpretation. However, the choice of how many levels to display and what these levels should be, remains to be made. Usually, simple sample statistics, such as mean, standard deviation or percentiles of the raw data are used (perhaps after some simple transformation and/or filtering). However, these techniques do suffer from serious deficiencies which severely limit their use in

* Departments of Archaeology & Mathematics, University of Nottingham, Nottingham NG7 2RD

† Department of Mathematics, University of Nottingham, Nottingham NG7 2RD

more noisy and complex situations. For example, contouring is susceptible to noisy or anomalous data points and is not very suitable if fairly abrupt changes in level are to be expected; dot density and greyscales assume that any noise in the data is relatively small and therefore an observation is an accurate representation of (in our case) the true phosphate level.

For these and other reasons we believe that dot density, greyscales, contouring and similar methods are not the most appropriate techniques for the spatial analysis of soil phosphate data arising from field surveys. Moreover, at this stage in the development of phosphate analysis for archaeology, since the chemical processes are little understood, it seems inappropriate to try to determine from the data anything more than areas or zones that have roughly the same phosphate level. If we accept this, then the problem becomes one of image segmentation. In a previous paper (Buck *et al.* 1988) we have described how change-point analysis may be used objectively to divide an area into regions of high or low phosphate concentrations. In the next section we discuss how other forms of image segmentation technique may be used to determine regions corresponding to more than two levels of phosphate.

14.2 General methodology

Consider a continuous two-dimensional region S which is partitioned into an $M \times N$ rectangular array of cells labelled $(i, j), i = 1, 2, \ldots M; j = 1, 2, \ldots, N$. Each cell can take one of L levels, labelled $1, 2, \ldots, L$ where for the moment L is finite. An arbitrary scene of S will be denoted by $\boldsymbol{x} = (x_{11}, x_{12}, \ldots, x_{mn})$, where x_{ij} is the corresponding level of cell (i, j). Let $\boldsymbol{x}^*$ be the true but unknown scene where $\boldsymbol{x}^*$ is a realisation of a random vector $\boldsymbol{x}$. Let y_{ij} be the observed record at cell (i, j) and $\boldsymbol{y}$ be the corresponding vector, which is a realisation of the random vector $\boldsymbol{y} = (Y_{11}, Y_{12}, \ldots, Y_{mn})$. Then given $\boldsymbol{y}$, what can be inferred about the true scene $\boldsymbol{x}^*$? Besag (1986) describes a mathematical and statistical framework, based on Markov random field theory, for tackling this problem in a logical fashion which enables prior information to be properly incorporated into the analysis. In doing so, two assumptions are made.

Assumption 1. The random variables $Y_{11}, Y_{12}, \ldots, Y_{mn}$ are mutually independent conditional on the scene $\boldsymbol{x}$, that is the conditional density of $\boldsymbol{y}$ given $\boldsymbol{x}$ is

$$p(\boldsymbol{y}|\boldsymbol{x}) = \prod_{\substack{i=1 \\ j=1}}^{\substack{m \\ n}} f(y_{ij}|x_{ij}), \tag{14.1}$$

where f is assumed to be known. In archaeological terms this means that given the underlying phosphate level for two cells, the actual observed phosphate readings are independent of each other.

Assumption 2. The true scene $\boldsymbol{x}^*$ is a realisation of a locally dependent Markov random field with specified distribution $p(\boldsymbol{x})$. In a Bayesian context $p(\boldsymbol{x})$ is viewed as a prior distribution for the true scene $\boldsymbol{x}^*$. In practise, the likely form of $p(\boldsymbol{x})$ for a particular archaeological site may be inferred as being that of other similar sites. Alternatively, $p(\boldsymbol{x})$ may describe some particular features that are suspected of being present at the site.

Besag demonstrates that the estimation of the true scene $\boldsymbol{x}^*$ given the data $\boldsymbol{y}$ poses a huge computational problem. However, by introducing the idea of locally dependent Markov random fields, this problem may be substantially reduced. Let $\boldsymbol{x}_A$ denote

the phosphate levels of the subset A of S and $x_{s\setminus ij}$, the phosphate levels of all cells except cell (i,j), then it is assumed that

$$p(x_{ij}|x_{s\setminus ij}) = p(x_{ij}|x_{\partial ij}), \tag{14.2}$$

where ∂_{ij} is some local neighbourhood of cell (i,j). In other words, the level of cell (i,j) only depends upon other cells in its immediate vicinity. Then, using assumption 1, it is immediately clear as a consequence of Bayes Theorem that

$$p(x_{ij}|y,\hat{x}_{s\setminus ij}) \propto f(y_{ij}|x_{ij})p(x_i|\hat{x}_{\partial_{ij}}), \tag{14.3}$$

where $\hat{x}_{ij}$ is our current estimate of the level at cell (i,j). Initially, $\hat{x}$ may be estimated using maximum likelihood applied to $f(y_{ij}|x_{ij})$. Then the algorithm is applied iteratively until convergence is obtained.

14.3 Detailed methodology

For a particular problem, $f(y_{ij}|x_{ij})$ and $p(x)$ need to be specified. For soil phosphate data, a generally made assumption is that the natural logarithm of the phosphate concentration has a normal distribution (see Cavanagh *et al.* 1988). That is, if Z_{ij} is the random variable representing the phosphate concentration in cell (i,j) then $Y_{ij} = ln(Z_{ij}) \sim N(x_{ij},\sigma^2)$ where x_{ij} is the true level for cell (i,j) and σ^2 is the variance about this true level.

The specification of $p(x)$ is more difficult, but Besag suggests that a suitable form, which takes into account local pairwise but not higher interactions, is

$$p(x) \propto \exp\left\{-\sum\sum_{1\le k<l\le L}\beta_{kl}n_{kl}\right\}, \tag{14.4}$$

where n_{kl} is the number of distinct neighbouring pairs of levels k and l, given a suitable definition of a neighbouring pair.

Given this formulation we have to find the level k that maximises $p(x_{ij}|y,\hat{x}_{s\setminus ij})$, which reduces to finding the level k that minimises

$$\frac{1}{2\sigma^2}(y_{ij}-\mu_k)^2 + \sum\sum_{1\le k<l\le L}\beta_{kl}u_{(k,l)}, \tag{14.5}$$

where μ_k is the mean ln(phosphate) for level k, $u_{ij}(k,l)$ is the number of neighbours of cell (i,j) of level l and β_{kl} is a weighting factor. The first term in the above expression can be thought of as allocating cell (i,j) to the level closest to y_{ij} subject to the second term which is a "smoothness" factor. The choice of the parameters β_{kl} should reflect our perception of how the phosphate is distributed. For example, let us consider a survey area in which there are three levels of phosphate. Does one expect an area of high phosphate (level 3) usually to be surrounded by a region of lower phosphate (level 2) before falling away to the background level (level 1), or is it conceivable that abrupt changes in the level from background to high are likely? The former case could be modelled by taking $\beta_{12} = \beta_{23}$ and $\beta_{13} = 2\beta_{12}$; the latter by taking $\beta_{12} = \beta_{23} = \beta_{13}$.

The interpretation of $p(x)$ as representing our prior knowledge or belief about how the phosphate levels are distributed, will be illustrated by continuing these two examples. Before doing so it is convenient to note that equation (14.5) may be rewritten so that the problem becomes that of finding the level k that minimises

$$(y_{ij}-\mu_k)^2 + \sum\sum_{1\le k<l\le L}\gamma_{kl}u_{ij}(k,l), \tag{14.6}$$

where $\gamma_{kl} = 2\sigma\beta_{kl}$.

Let us now consider the eight neighbour situation

$$\begin{matrix} 3 & 3 & 3 \\ 3 & X & 3 \\ 3 & 3 & 3 \end{matrix}$$

where for the reasons explained above our prior belief is high that the unknown level of the central observation (represented by X) is also of level 3. In fact, using (14.4), it is easy to see that

$$p_1 = P(X = 1|\text{ other cells}) = ce^{-8\gamma_{13}},$$
$$p_2 = P(X = 2|\text{ other cells}) = ce^{-8\gamma_{23}}$$

and

$$p_3 = P(X = 3|\text{ other cells}) = c,$$

where

$c^{-1} = 1 + e^{-8\gamma_{23}} + e^{-8\gamma_{13}}$ is the normalising constant.

Taking $\gamma_{23} = 0.125$ and $\gamma_{13} = 0.25$, then these probabilities are 0.09, 0.25 and 0.66 respectively. If this does not give sufficient weight to our prior belief that the centre cell should be of level 3, then we could increase γ_{23} and γ_{13} to 0.25 and 0.5 respectively; in this case p_1, p_2 and p_3 are 0.02, 0.12 and 0.86 respectively. Conversely, decreasing γ_{23} and γ_{13} will decrease p_3.

On the other hand, if we believe that the phosphate in the soil is not dispersed in any way, then it may be interpreted that an observation from level 1 or 2 is quite likely to be adjacent to observations from level 3. In this case, if we take $\gamma_{23} = \gamma_{13} = 0.125$ then

$$p_1' = P(X = 1|\text{ other cells}) = 0.21$$
$$p_2' = P(X = 2|\text{ other cells}) = 0.21$$
$$p_3' = P(X = 3|\text{ other cells}) = 0.58.$$

Increasing γ_{23} to 0.25 produces probabilities 0.11, 0.11 and 0.78 respectively.

By considering other eight-neighbour configurations we can investigate how this formulation reflects our preconceptions about the phosphate levels. For example, the situation

$$\begin{matrix} 3 & 3 & 2 \\ 3 & X & 2 \\ 3 & 2 & 1 \end{matrix}$$

yields

$$q_1 = P(X = 1|\text{ other cells}) = c\exp\{-4\gamma_{13} - 3\gamma_{12}\},$$
$$q_2 = P(X = 2|\text{ other cells}) = c\exp\{-4\gamma_{23} - \gamma_{12}\}$$

and

$$q_3 = P(X = 3|\text{ other cells}) = c\exp\{-3\gamma_{23} - \gamma_{13}\},$$

where c is the normalising constant for this situation.

Taking $\gamma_{12} = \gamma_{23} = 0.125$ and $\gamma_{13} = 0.25$, then the values of q_1, q_2 and q_3 are 0.2, 0.4 and 0.4 respectively. Taking $\gamma_{12} = \gamma_{23} = \gamma_{13} = 0.125$, they are 0.27, 0.35 and 0.38 respectively. Thus, in the latter case, the prior probability that $X = 1$ is slightly higher than in the former case.

In the above formulation all adjacent cells are treated equally, but an obvious refinement is to give lower weight to cells that are diagonally adjacent. In this case $\beta_{kl}n_{kl}$ in equation (14.4) is replaced by $\beta_{kl}'n_{kl}' + \beta_{kl}''n_{kl}''$ where n_{kl}' and n_{kl}'' are the respective numbers of first and second-order neighbour pairs. Obvious modifications then have to be made to the conditional probabilities in the above examples. Typically β_{kl}'' is taken as $\beta_{kl}'/\sqrt{2}$ (see Geman & McClure 1985).

14.4 Simulation

In order to demonstrate the effect of the above methodology we consider two simple simulations which idealise common spatial distributions found in real data sets. Both simulations are carried out on a 40×40 grid. The first arises out of work with field survey data where known archaeological sites have been further studied by the analysis of surface soil samples. On such sites, areas of high phosphate are commonly surrounded by a lower, but still enhanced, level before decreasing to the background level. This is commonly interpreted as representing either the archaeological spreading of phosphate-rich material around the outside of the occupied site or some form of post-depositional activity which has led to the movement of the phosphate outwards from its original position. We represent such a situation as a concentration of high phosphate within a circle of radius 10 units with its centre at the centre of the grid. Surrounding this circle we suppose that there is a region of lower phosphate level of width 15 units with the remainder of the grid being at the background phosphate level. We have chosen the high, low and background ln(phosphate) levels to be 2, 1, 0 respectively and the resulting greyscale is shown in Fig. 14.1(a). To the ln(phosphate) in each cell we have added a noise component simulated from a normal distribution with mean zero and standard deviation 2/3. The greyscale representation using the 60, 70, 80 and 90 percentiles is given in Fig. 14.1(b). Assuming that the mean and standard deviation for each level are known, the image corresponding to the maximum likelihood estimate is given in Fig. 14.1(c). We now apply the image segmentation algorithm outlined above with the diagonally adjacent cells downweighted by a factor of $1/\sqrt{2}$. The values of the prior probabilities for the level of X, p_1, p_2 and p_3, are 0.19, 0.31 and 0.50 respectively. The resulting image is given in Fig. 14.1(d). The analysis is repeated but with p_1, p_2 and p_3 being 0.03, 0.17 and 0.80 respectively, and the image is given in Fig. 14.1(e). The final run of the algorithm with $p_1 = 0.002$, $p_2 = 0.048$ and $p_3 = 0.95$ gives the image in Fig. 14.1(f). Figs. 14.1(d), 14.1(e) and 14.1(f) clearly demonstrate how changing the prior information about the scene affects the algorithm's ability to identify the true underlying scene. Even with the highest prior probability for the scene being smooth ($p_3 = 0.95$) there are isolated cells allocated to the wrong level. With $p_3 = 0.5$ there are considerable patches of the background level wrongly designated to the intermediate level and some of these persist when $p_3 = 0.8$ is used. Obviously taking higher values of p_3 will make the image even smoother.

We have conducted a second simulation with two elliptical areas of phosphate above the background level. In contrast to the first simulation, the two areas are adjacent but not concentric, with one ellipse containing a region of high phosphate and the second a region of lower phosphate (but above the background level). Although this situation could arise from many activities, common archaeological interpretations include a site with two phases of occupation or alternatively a site with its associated midden. Fig. 14.2(a) and (b) show the regions before and after the addition of random noise from a normal distribution with zero mean and standard deviation 2/3. The maximum likelihood estimates of the regions are shown in Fig. 14.2(c). The results from three applications of the image segmentation algorithm (with parameters as in the previous example) are shown in Fig. 14.2(d), 14.2(e) and 14.2(f). Again the effect of the algorithm can be readily seen. Despite using prior information which strongly suggests that a region of high phosphate should be not be adjacent to the background level, the algorithm with $p_3 = 0.95$ captures the main features of of the underlying scene. However there is a tendency for a small region at the intermediate level to be erroneously found between the high and background zones.

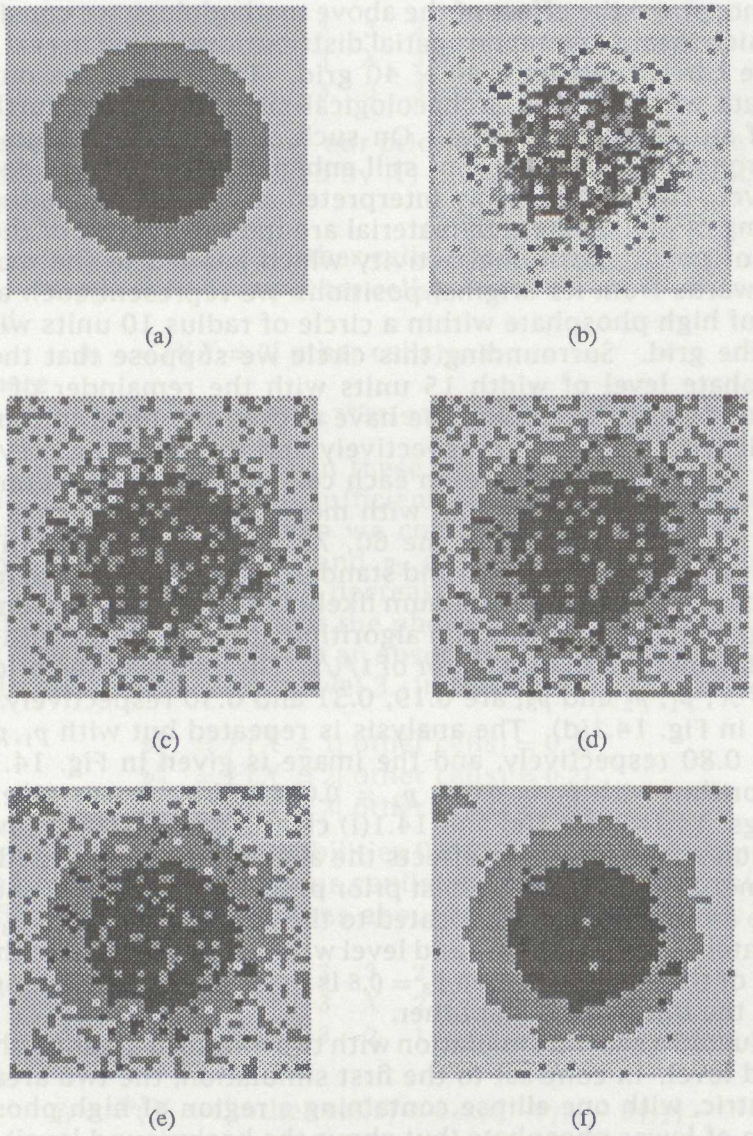

Figure 14.1: Simulation with a circular area of high phosphate surrounded by an area of lower phosphate.

(a) The original scene.

(b) Greyscale representation of the original scene plus noise with percentiles at 60, 70, 80 and 90%.

(c) Maximum likelihood estimated image.

(d) Image segmentation with $p_1 = 0.19$, $p_2 = 0.31$ and $p_3 = 0.50$.

(e) Image segmentation with $p_1 = 0.03$, $p_2 = 0.17$ and $p_3 = 0.80$.

(f) Image segmentation with $p_1 = 0.002$, $p_2 = 0.048$ and $p_3 = 0.95$.

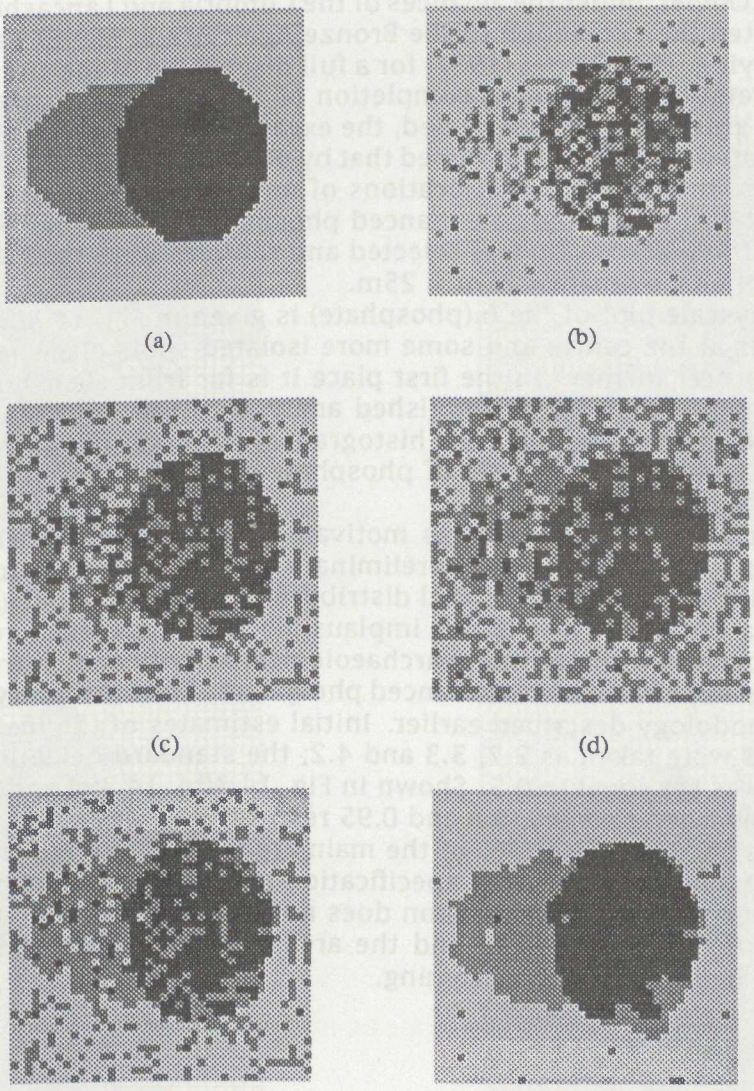

Figure 14.2: Simulation with two overlapping elliptical areas, one of high phosphate and one of lower phosphate.

(a) The original scene.

(b) Greyscale representation of the original scene plus noise with percentiles at 60, 70, 80 and 90%.

(c) Maximum likelihood estimated image.

(d) Image segmentation with $p_1 = 0.19$, $p_2 = 0.31$ and $p_3 = 0.50$.

(e) Image segmentation with $p_1 = 0.03$, $p_2 = 0.17$ and $p_3 = 0.80$.

(f) Image segmentation with $p_1 = 0.002$, $p_2 = 0.048$ and $p_3 = 0.95$.

14.5 Example

In 1982 Adrian Olivier, under the auspices of the Cumbria and Lancashire Archaeological Unit, directed the excavation of the Bronze Age cairn at Manor Farm in advance of gravel quarrying. (See Olivier (1987) for a full discussion of the site, its excavation and its interpretation.) Following completion of the excavation, when inhumation and cremation graves had been revealed, the excavation team collected soil samples from the excavated surface. It was hoped that by phosphate analysis of these samples it might be possible to reveal the locations of any graves not recoverable through excavation, but still remaining as enhanced phosphate in the soil. With this aim in mind a survey interval of 0.5m was selected and samples were taken over the entire cairn, an area of approximately 25m $\times$ 25m.

An initial greyscale plot of the ln(phosphate) is given in Fig. 14.3(a) and indicates enhanced levels at the centre and some more isolated spots elsewhere, but is very difficult to interpret further. In the first place it is far from clear how many levels should be distinguished. In the published analysis of the raw data (Olivier 1987) four levels were chosen. However, the histogram of the natural logarithm of the data suggests that there are three levels of phosphate concentration and this forms the basis of our analysis.

Although the phosphate survey was motivated by a desire to locate burials not detected during excavation, we see a preliminary step towards this aim to be gaining an understanding of the general spatial distribution of the phosphate over the whole site. In such a burial cairn it is not implausible that there exist regions of high phosphate associated with maximum archaeological activity and, surrounding these, regions of lower but nevertheless enhanced phosphate. Thus the data were processed using the methodology described earlier. Initial estimates of the means for each of the three levels were taken as 2.7, 3.3 and 4.2; the standard deviation of the noise for each level was set equal to 0.5. Shown in Fig. 14.3(b), 14.3(c) and 14.3(d) are the resulting images when p_3 is 0.5, 0.8 and 0.95 respectively.

These images show the locations of the main areas of phosphate enhancement on the site. In fact, changing the prior specification to increase the chance that a high-level region abutts a background region does not alter the general picture. Further discussion of the data processing and the archaeological interpretation of Manor Farm are given in Buck et al. forthcoming.

14.6 Discussion

The simulated examples given above are very simple and based upon very naive archaeological models, but we hope that they do illustrate how the image segmentation algorithm works and, in particular, how prior information is clearly and unambiguously incorporated into the analysis. Workers used to dot density and greyscales may find this concept difficult to contemplate. However producing images by these other methods leaves much of the interpretation to the eye and to the prejudices of the beholder. At least the Bayesian analysis used in this paper ensures that any external information used is explicitly stated. Of course it is possible to experiment with different prior probabilities. This is perfectly reasonable and will clearly demonstrate how changing the prior information changes the resulting image and so gives some impression of the strength of evidence supporting a particular interpretation.

One strong impression arising from these examples and from our experiences with other simulations and real data sets is that an isolated high phosphate reading (or perhaps a small group) is not very meaningful without any other supporting archaeological evidence. This is undoubtedly due to the high noise level in the

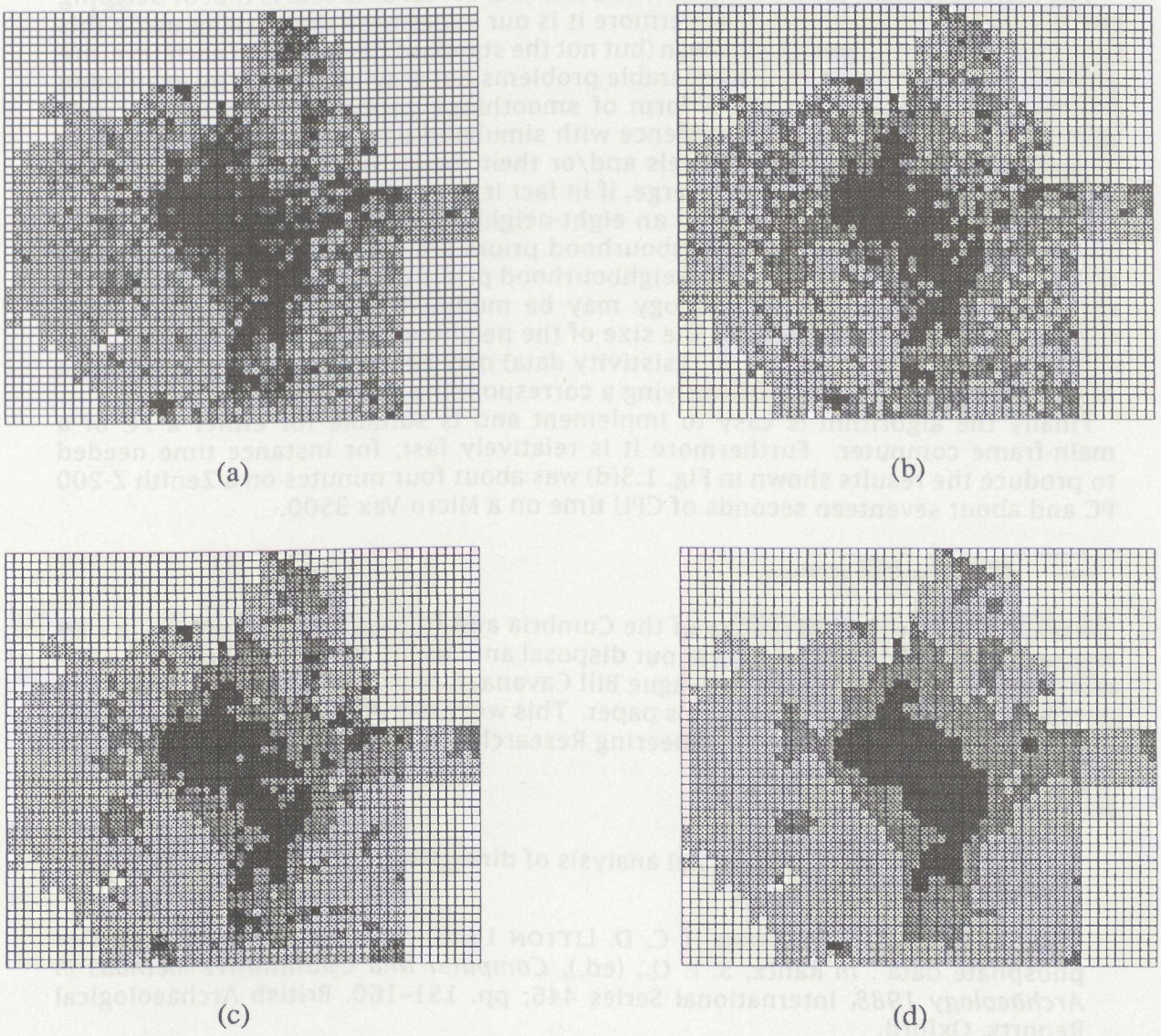

Figure 14.3: Phosphate data from Manor Farm after Olivier 1987

(a) Greyscales representation of the Manor Farm data with percentiles at 60, 70, 80 and 90%.

(b) Image segmentation with $p_1 = 0.19$, $p_2 = 0.31$ and $p_3 = 0.50$.

(c) Image segmentation with $p_1 = 0.03$, $p_2 = 0.17$ and $p_3 = 0.80$.

(d) Image segmentation with $p_1 = 0.002$, $p_2 = 0.048$ and $p_3 = 0.95$.

data. In the simulations above the standard deviation of the noise relative to the differences between the underlying levels is realistic but perhaps a little on the low side. Several real data sets that we examined appear to have a slightly higher noise level and so the resulting images from the algorithm are not as good.

The major problem confronting us in the analysis of real data sets is that of deciding the number of levels to use. Furthermore it is our experience that having made this decision, good estimates of the mean (but not the standard deviation) of each level are needed. Both of these pose considerable problems but may be tackled by examining histograms (perhaps after some form of smoothing), probability plots and crude greyscale plots of the data. Experience with simulated data suggests that incorrect specification of the number of levels and/or their means is usually typified by the algorithm being very slow to converge, if in fact it does converge.

In these examples we have used an eight-neighbourhood prior. Obvious modifications are to vary this; four-neighbourhood priors are more suitable for detecting rectangular shapes whereas eight-neighbourhood priors are better for curved objects or zones. The general methodology may be modified to search for other more complicated areas by increasing the size of the neighbourhood. For example, linear features (e.g. walls or ditches in resistivity data) may be searched for by selecting a suitable neighbourhood and specifying a corresponding prior distribution.

Finally the algorithm is easy to implement and is suitable for either a PC or a main-frame computer. Furthermore it is relatively fast; for instance time needed to produce the results shown in Fig. 1.3(d) was about four minutes on a Zenith Z-200 PC and about seventeen seconds of CPU time on a Micro-Vax 3500.

14.7 Acknowledgements

We are grateful to Adrian Olivier of the Cumbria and Lancashire Archaeological Unit for putting the Manor Farm data at our disposal and for his helpful suggestions in its interpretation. We thank our colleague Bill Cavanagh for his useful and constructive comments on an earlier draft of this paper. This work has been carried out under the auspices of a UK Science and Engineering Research Council grant (ref: GB/E/30263).

Bibliography

BESAG, J. 1986. "On the statistical analysis of dirty pictures", *Royal Stat. Soc. B*, 48: 259–302.

BUCK, C. E., W. G. CAVANAGH, & C. D. LITTON 1988. "The spatial analysis of site phosphate data". *in* Rahtz, S. P. Q., (ed.), *Computer and Quantitative Methods in Archaeology 1988*, International Series 446, pp. 151–160. British Archaeological Reports, Oxford.

BUCK, C. E., W. G. CAVANAGH, & C. D. LITTON forthcoming. "Tools for the interpretation of soil phosphate data from archaeological surveys". *in Proceedings of the Geoprospection in the Archaeological Landscape Conference, Dorset, January 1989.*

CAVANAGH, W. G., S. HIRST, & C. D. LITTON 1988. "Soil phosphate, site boundaries and change-point analysis", *Journal of Field Archaeology*, 15: 67–83.

FIELLER, N. R. J. & E. C. FLENLEY 1988. "Statistical analysis of particle sizes and sediments". *in* Ruggles, C. L. N. & Rahtz, S. P. Q., (eds.), *Computer and Quantitative Methods in Archaeology 1987*, International Series 393, pp. 79–94. British Archaeological Reports, Oxford.

GEMAN, S. & D. E. MCCLURE 1985. "Bayesian image analysis : an application to single photon emission tomography", *Proc. Amer. Statist. Assoc.*, pp. 12–18.

OLIVIER, A. C. H. 1987. "Excavation of a Bronze Age funerary cairn at Manor Farm, near Borwick, North Lancashire", *Proceedings of the Prehistoric Society*, 53: 129–186.