PNNL-15420

Pacific Northwest National Laboratory
Operated by Battelle for the U.S. Department of Energy

Clearing Unexploded Ordnance: Bayesian Methodology for Assessing Success

K.K. Anderson

August 2005

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes **any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights**. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service,
U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161
ph: (800) 553-6847
fax: (703) 605-6900
email: orders@ntis.fedworld.gov
online ordering: http://www.ntis.gov/ordering.htm

This document was printed on recycled paper.
(9/2003)

PNNL-15420

Clearing Unexploded Ordnance: Bayesian Methodology for Assessing Success

K.K. Anderson

August 2005

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Clearing Unexploded Ordnance: Bayesian Methodology for Assessing Success

Kevin Anderson

Pacific Northwest National Laboratory
Richland, Washington

August 1, 2005

1 Introduction.

The Department of Defense has many Formerly Used Defense Sites (FUDS) that are slated for transfer for public use. Some sites have unexploded ordnance (UXO) that must be cleared prior to any land transfers. Sites are characterized using geophysical sensing devices and locations are identified where possible UXO may be located.

In practice, based on the analysis of the geophysical surveys, a dig list of N suspect locations is created for a site that is possibly contaminated with UXO. The suspect locations on the dig list are often assigned into K bins ranging from "most likely to contain UXO" to "least likely to be UXO" based on signal discrimination techniques and expert judgment. Usually all dig list locations are sampled to determine if UXO is present before the site is determined to be free of UXO. While this method is 100% certain to insure no UXO remains in the locations identified by the signal discrimination and expert judgment, it is very costly. This paper proposes a statistical Bayesian methodology that may result in digging less than 100% of the suspect locations to reach a pre-defined tolerable risk, where risk is defined in terms of a low probability that any UXO remains in the unsampled dig list locations. Two important features of a Bayesian approach are that it can account for uncertainties in model parameters and that it can handle data that becomes available in stages. The results from each stage of data can be used to direct the subsequent digs.

Prior research at PNNL was based on a probability model involving a known performance matrix and an unknown site-dependent parameter, the true proportion of UXO on the site (Anderson et al. 2003). The performance matrix specifies the probabilities of UXO and non-UXO locations being assigned into each of the K bins, that is, it characterizes the statistical model of the signal discrimination techniques and expert judgment. A Bayesian approach was proposed in the estimation of the probability that there is no UXO left in the

remaining unsampled suspect locations. Digging could be stopped once this probability was determined to be high enough with desired statistical confidence. This paper extends this previous Bayesian research to the case when the performance matrix has uncertainty. The uncertainty is important because the higher it is, the more suspect locations would have to be dug up.

The probability model and the figure of merit are described in Section 2. A discussion of Dirichlet priors, used to model the uncertainty in the performance matrix, is presented in Section 3. The calculation methodology, Markov Chain Monte Carlo, is presented in Section 4. Dynamic dig strategies are presented in Section 5. These strategies are based on additional figures of merit which suggest which suspect locations should be dug up in subsequent stages of sampling in order to reduce the risk of stopping digging. An appendix presents the modifications to the probability model when the signal discrimination techniques and expert judgment produces continuous numeric values instead of assigning suspect locations to discrete bins.

2 The Probability Model.

Assume that there are N suspected UXO locations on the initial dig list for the site of interest. The true state of each suspect location is given by X_i, which is a random variable defined as

$$X_i = \begin{cases} 1 & \text{if UXO is present} \\ 0 & \text{if UXO is not present} \end{cases} \qquad \text{for } i = 1, 2, \ldots, N. \tag{1}$$

Before the geophysical measurements have been taken, each X_i can be considered to be a random sample from a Bernoulli distribution with probability determined by its expectation:

$$\mathrm{E}(X_i) = \Pr(X_i = 1) = p \ , \tag{2}$$

where p is the proportion of identified suspect locations on the site that contain UXO.

Let Y_i represent the binning response from the geophysical measurement system at the i-th suspect location. The response Y_i is a discrete random variable with K possible outcomes ordered from 1 to K, where 1 means "most likely" to contain UXO and K means "least likely" to contain UXO. The conditional probability distribution of the response Y_i given the true state X_i is determined by the geophysical measurement system performance matrix. That is, there is a $(K \times 2)$ matrix A with (k, l)-element A_{kl}, such that

$$A_{kl} = \Pr(Y_i = k \mid X_i = l) \ \text{ for } k = 1, 2, \ldots, K \text{ and } l = 0, 1 \ . \tag{3}$$

The performance matrix A is presented in Table 1. Each column of the performance matrix is a multinomial probability distribution (and each column sums to one) which characterizes the statistical model of the signal discrimination techniques and expert judgment used to bin each suspect location.

Table 1: Performance Matrix A

Bin	Probability that a location that is truly **not** UXO is assigned to this bin	Probability that a location that is truly UXO is assigned to this bin
1 - highest likelihood of UXO	A_{10}	A_{11}
2 - high likelihood of UXO	A_{20}	A_{21}
$\vdots$	$\vdots$	$\vdots$
K - lowest likelihood of UXO	A_{K0}	A_{K1}

The data from the site of interest (unobserved until dug), (X_i, Y_i), $i = 1, 2, \ldots, N$, can be tabulated into a $(K \times 2)$ contingency table. As counts in a two-way contingency table, each suspect location randomly falls into a cell of the table according to a two-dimensional probability matrix, given in Table 2. As data become available, the likelihood is based on these probabilities.

Table 2: Two-Dimensional Probability Matrix

Bin	Not UXO	UXO	Totals
1	$(1-p)A_{10}$	pA_{11}	$(1-p)A_{10} + pA_{11}$
2	$(1-p)A_{20}$	pA_{21}	$(1-p)A_{20} + pA_{21}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
K	$(1-p)A_{K0}$	pA_{K1}	$(1-p)A_{K0} + pA_{K1}$
Totals	$(1-p)$	p	1

The first data available to the analyst are the binning results, $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_N)$, which are summarized by the bin counts $(N_1, N_2, \ldots, N_K)$, where N_i is the number of suspect locations assigned to bin i. The likelihood function of this data is

$$L(A, p\,; \mathbf{Y}) \propto \prod_{i=1}^{K} \left((1-p)A_{i0} + pA_{i1}\right)^{N_i} \ . \tag{4}$$

As suspect locations are dug, the likelihood changes. Let D be the set of n indices of the suspect locations that have been dug up and let $\mathbf{X}_D = \{X_j\,;\, j \in D\}$ represent the dig results. Then the likelihood is

$$\begin{aligned} L(A,p\,;\mathbf{X}_D,\mathbf{Y}) &\propto L(A,p\,;\mathbf{Y}) \prod_{i=1}^{K} \left(\frac{(1-p)A_{i0}}{(1-p)A_{i0}+pA_{i1}} \right)^{n_{i0}} \left(\frac{pA_{i1}}{(1-p)A_{i0}+pA_{i1}} \right)^{n_{i1}} \\ &\propto \prod_{i=1}^{K} \left((1-p)A_{i0}+pA_{i1}\right)^{N_i-n_i} \prod_{i=1}^{K} \left((1-p)A_{i0}\right)^{n_{i0}} \left(pA_{i1}\right)^{n_{i1}} \end{aligned} \tag{5}$$

where n_i is the number of bin i suspect locations that have been dug up and n_{i1} and n_{i0} are the number of these suspect locations that were found to be UXO or not, respectively.

2.1 Probability of No UXO Remaining.

If A and p were known, the probability that all of the remaining suspect locations are not UXO, P_0, is

$$P_0 = \prod_{i=1}^{K} \left(\frac{(1-p)A_{i0}}{(1-p)A_{i0}+pA_{i1}} \right)^{N_i-n_i}, \tag{6}$$

where $N_i - n_i$ is the number of bin i suspect locations that remain undug. Because p and the elements of the matrix A are random variables in our Bayesian approach, P_0 is also a random variable which depends on the binning results, dig results, and prior distributions. However, for simplicity, this dependence is not reflected in the notation. The posterior distributions of A and p given the data at each stage of analysis $\mathbf{X}_D$ and $\mathbf{Y}$ can be propagated into a posterior distribution for P_0 from which probabilistic inferences can be made, such as determination of lower confidence bounds.

In previous research, we assumed that the A matrix was a known input to the problem. Our previous Bayesian analysis only had to contend with the prior distribution of p. In this paper, we extend our Bayesian analysis to account for uncertainty in the A matrix by assuming that two Dirichlet probability distributions are available to characterize the prior distributions of the columns of the A matrix. The uncertainty in the A matrix results in greater uncertainty in P_0. The Dirichlet distribution is presented in the next section.

3 The Dirichlet Distribution.

The Dirichlet distribution is convenient as a prior distribution for this problem because it is a conjugate family for multinomial distributed data (Berger 1985). The $(K-1)$-dimensional

Dirichlet distribution with parameters $(\alpha_1, \alpha_2, \ldots, \alpha_K)$ can be described by the joint density of $(K-1)$ variables

$$f(p_1, p_2, \ldots, p_{K-1}) \propto \prod_{i=1}^{K-1} p_i^{\alpha_i - 1} (1 - \sum_{j=1}^{K-1} p_j)^{\alpha_K - 1} , \tag{7}$$

for $\alpha_i > 0$, $p_i \geq 0$, $\sum_{i=1}^{K-1} p_i \leq 1$, and $f(p_1, p_2, \ldots, p_K) = 0$ otherwise (Johnson and Kotz 1972). The constant of proportionality is not critical in calculations.

The conjugate family property can be seen as follows. Let $\mathbf{M} = (M_1, M_2, \ldots, M_K)$ be multinomial with probabilities $(p_1, p_2, \ldots, p_K)$ and number of trials $n = \sum_{i=1}^{K} M_i$. If n is given and the probabilities have a Dirichlet prior, then the posterior distribution on the probabilities given the multinomial data is

$$\begin{aligned} f(p_1, p_2, \ldots, p_{K-1} \mid \mathbf{M}) &\propto f(M_1, M_2, \ldots, M_K \mid n, p_1, p_2, \ldots, p_{K-1}) f(p_1, p_2, \ldots, p_{K-1}) \\ &\propto \left(\prod_{i=1}^{K-1} p_i^{M_i} (1 - \sum_{j=1}^{K-1} p_j)^{M_K} \right) \left(\prod_{i=1}^{K-1} p_i^{\alpha_i - 1} (1 - \sum_{j=1}^{K-1} p_j)^{\alpha_K - 1} \right) \\ &\propto \prod_{i=1}^{K-1} p_i^{\alpha_i + M_i - 1} (1 - \sum_{j=1}^{K-1} p_j)^{\alpha_K + M_K - 1} , \end{aligned} \tag{8}$$

which is recognized as a Dirichlet distribution with parameters $(\alpha_1 + M_1, \alpha_2 + M_2, \ldots, \alpha_K + M_K)$. Hence, a Dirichlet prior yields a Dirichlet posterior, and further data (i.e., another realization of $\mathbf{M}$) updates the posterior by simply adding results to the Dirichlet parameters. This updating property makes the Dirichlet distribution the natural prior distribution for each column of the performance matrix A.

The Dirichlet distribution features a simple formula for the marginal expected values. If $(p_1, p_2, \ldots, p_{K-1})$ has a prior Dirichlet distribution with parameters $(\alpha_1, \alpha_2, \ldots, \alpha_K)$, then the prior expected value of p_i is

$$E(p_i) = \frac{\alpha_i}{\alpha_0} \text{ for } i = 1, 2, \ldots, K , \tag{9}$$

where $\alpha_0 = \sum_{j=1}^{K} \alpha_j$ and the posterior expected value of p_i after observing $\mathbf{M}$ is

$$E(p_i \mid \mathbf{M}) = \frac{\alpha_i + M_i}{\alpha_0 + n} \text{ for } i = 1, 2, \ldots, K , \tag{10}$$

where $n = \sum_{i=1}^{K} M_i$.

3.1 Setting Dirichlet Parameters.

An Empirical Bayes approach to setting the prior distributions on the columns of the performance matrix is natural. For example, if a seeded study of the signal discrimination technique is performed, the columns of the $(K \times 2)$ contingency table of results can be used as the parameters of the Dirichlet priors. Results gathered from the clearing of other similar defense sites might also be used, where similar is in terms of ordnance used and geology.

Alternatively, the parameters of the Dirichlet priors can be set using just expert judgment. For example, if the expert believes his A_{i1} probabilities are the p_i values in column 2 of Table 3, then the expert can set the parameters of the Dirichlet prior as

$$\alpha_i = Cp_i \text{ for } i = 1, 2, \ldots, K \ , \tag{11}$$

for some value of $C > 0$. By Equation 9, the marginal expected values of the Dirichlet distribution are the given p_i values. The choice of C will control the uncertainty, with a larger value giving less uncertainty, as the relative standard deviation is

$$RSD_i = \sqrt{\frac{1 - p_i}{(C+1)p_i}} \text{ for } i = 1, 2, \ldots, K \tag{12}$$

(see Berger 1985). Because there is only one parameter controlling the uncertainty of all K probabilities, the expert cannot set the uncertainty on each probability individually. Table 3 presents two times the relative standard deviations as a percentage for some values of C and for a fixed set of probabilities. For example, if the expert thinks his probabilities are as in column 2 of Table 3 and that his 2-sigma uncertainty on $p_1 = 0.48$ is about $\pm 5\%$ (i.e., $p_1 \in (0.456, 0.504)$), then he should choose $C = 2000$. In general, such a table could be generated for the specific set of p_i the expert has in mind, allowing him to choose an appropriate value for C which models his uncertainties.

Expert judgment can be incorporated with empirical data using a Bayes/Empirical Bayes approach. For example, if the expert models his initial beliefs regarding A_{i1} as a Dirichlet with parameters $(\alpha_1, \alpha_2, \ldots, \alpha_K)$ and empirical results from seeding the site of interest with UXO give bin counts $(X_1, X_2, \ldots, X_K)$, then the prior on the A_{i1} probabilities for the analysis of the site of interest is Dirichlet with parameters $(\alpha_1 + X_1, \alpha_2 + X_2, \ldots, \alpha_K + X_K)$.

After the priors on the two columns of the A matrix are set, the prior on the parameter p must be set. The initial PNNL research suggested the use of a Beta distribution (Anderson et al. 2003). It just so happens that a Beta distribution is a Dirichlet distribution with

Table 3: Example of Dirichlet 2-Sigma Relative Uncertainties

Bin	Probability p_i	Parameter Multiplier C 100	500	1000	2000	5000
1	0.480	20.8%	9.3%	6.6%	4.7%	2.9%
2	0.250	34.6%	15.5%	11.0%	7.7%	4.9%
3	0.150	47.6%	21.3%	15.1%	10.6%	6.7%
4	0.109	57.2%	25.6%	18.1%	12.8%	8.1%
5	0.010	199.0%	89.0%	62.9%	44.5%	28.1%
6	0.001	632.1%	282.7%	199.9%	141.4%	89.4%

$K = 2$. The two parameters can be set using information about dud rates for the ordnance used at the site of interest, results gathered from the clearing of other similar defense sites, or results from initial sub-site clearing. Choice of α_1 and α_2 is made easy using the expressions for the mean and variance of the Beta distribution:

$$\mu = \frac{\alpha_1}{\alpha_1 + \alpha_2} \quad \text{and} \quad \sigma^2 = \frac{\alpha_1 \alpha_2}{(\alpha_1 + \alpha_2)^2 (\alpha_1 + \alpha_2 + 1)} \ . \tag{13}$$

These expressions are inverted as

$$\alpha_1 = \mu \left(\frac{(1-\mu)\mu}{\sigma^2} - 1 \right) \quad \text{and} \quad \alpha_2 = (1-\mu) \left(\frac{(1-\mu)\mu}{\sigma^2} - 1 \right) \ . \tag{14}$$

So if prior information on p is represented as a mean μ and a variance σ^2, then Equation 14 provides the values of α_1 and α_2 for the Beta prior distribution.

4 Markov Chain Monte Carlo.

The goal of the Bayesian analysis is to produce the joint posterior distribution of the parameters A and p conditional on the available data. It is then possible to make probabilistic inferences on the parameters or functions of the parameters, such as P_0 in Equation 6. The Bayesian analysis is based on the fact that the joint posterior density function is proportional to the product of the prior density functions and the likelihood function. Section 3 proposed the use of Dirichlet distributions as priors for the two columns of A and for p. Likelihood functions were given in Section 2, Equations 4 and 5, depending on what data is available. The clearing of the site can be performed in stages, with a Bayesian analysis at each stage assessing how well cleared the site is and guiding the subsequent stage of clearing. More discussion on this is in Section 5.

Unfortunately, the likelihood functions in Equations 4 and 5 are too complicated to admit a closed form solution for the posterior distribution. Thus, producing the posterior distribution at each stage in the site clearing process is a numerical analysis problem. Markov Chain Monte Carlo algorithms (MCMC) are powerful tools for solving the problem (Gilks et al. 1996). MCMC produces pseudo-randomly generated data from the desired posterior distribution. These simulated data are then used to make the desired probabilistic inferences on the parameters or functions of the parameters.

We propose the use a form of MCMC known as an "independence sampler Metropolis-Hastings" algorithm (Tierney 1994). While it is one of the easiest MCMC algorithms to implement, the numerical details will be omitted from this paper. Our approach is to generate the random data from Dirichlet distributions using maximum posterior analyses to set the parameters of the distributions. Our implementation is in the R language and environment for statistical computing and graphics. R is available as Free Software under the terms of the Free Software Foundation's GNU General Public License (see www.r-project.org).

5 Dynamic Dig Strategies.

Our proposed dig strategy is to first sample some suspect locations from each of the K bins and then to let the dig results guide the subsequent sampling. The preliminary sampling will provide a check on the prior beliefs regarding the performance of the discrimination procedure at the site. Of course, all suspect locations judged most likely to be UXO will have to be dug up, so the strategies discussed below will pertain to the suspect locations with lower likelihoods of being UXO.

Interim results from the sampling and Bayesian analysis can be used to guide the subsequent sampling. It is useful to examine the conditional probabilities

$$P_{0i} = \frac{(1-p)A_{i0}}{(1-p)A_{i0} + pA_{i1}} \quad \text{for} \quad i = 1, 2, \ldots, K \ , \tag{15}$$

which represent the likelihood that a remaining Bin i suspect location does not contain UXO. Because these conditional probabilities enter into the calculation of P_0 in Equation 6, they need to be large, i.e., close to 1, in order for P_0 to be large. We can estimate the posterior expected values of these conditional probabilities using the simulated MCMC A and p values.

Other useful quantities are the expected number of UXO remaining in each bin. These are calculated by $(N_i - n_i)(1 - P_{0i})$ for $i = 1, 2, \ldots, K$ and the goal is to minimize them. The

expected number of UXO remaining could have great importance in the decision to transfer a site to public use.

One can use the MCMC analysis in a predictive manner to decide how many suspect locations of any Bin need to be dug up. This involves doing calculations using predicted dig outcomes. For example, if the least likely to be UXO bin remains unsampled, we can calculate the effect on P_0 of digging 100 more suspect locations and finding no UXO among them. In this manner of forward calculations, we can determine approximately how many more samples are required before we can stop digging.

5.1 Digging by Spatial Blocks.

Implicit in the discussion thus far has been the ability to dig up desired individual suspect locations. It may be case that the site of interest is spatially partitioned into blocks and that for cost and effort reasons digging will be done by blocks. This means that instead of choosing to dig all suspect locations from a certain bin, the analyst has to choose which blocks to dig. Hence, the analyst has substantially less freedom to determine from which bins the locations will be sampled. However, the dig results can still be summarized in terms of Bin and UXO/Not UXO counts and the likelihood produced as in Equation 5.

6 Conclusions.

A probability model was presented which can be used to analyze the type of data available from geophysical surveys and digging of suspect UXO locations at FUDS that are slated for transfer for public use. Dig strategies were discussed that reduce the probability that any UXO remains in the unsampled dig list locations to a tolerable level so that a complete sampling is unnecessary. A Bayesian methodology was proposed to account for the uncertainties involved with discriminating suspect locations into bins ranging from "most likely to contain UXO" to "least likely to be UXO."

References

[1] Anderson, K.K., R.F. O'Brien, and P.G. Heasler (2003). A statistical Bayesian methodology to stop digging geophysical anomalies based on results from incremental digs, white paper, Pacific Northwest National Laboratory, Richland, WA.

[2] Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis*, Second Edition, Springer-Verlag, New York, NY.

[3] Gilks, W.R., S. Richardson, and D.J. Spiegelhalter (1996). *Markov Chain Monte Carlo in Practice*, Chapman & Hall/CRC, London.

[4] Johnson, N.L., and S. Kotz (1972). *Distributions in Statistics. Continuous Multivariate Distributions*, John Wiley & Sons, New York, NY.

[5] Tierney, L. (1994). Markov chains for exploring posterior distributions (with discussion), Ann. Statist. 22: 1701-1762.

7 Appendix.

In this appendix we present the model for continuous metric UXO discrimination. That is, instead of assigning suspect locations to discrete bins, $Y_i = k$, which range from "most likely to contain UXO" to "least likely to be UXO," Real numeric values, $Y_i = y_i$, are produced; the larger the value, the more likely to contain UXO. In this case the performance matrix is replaced by two probability density functions representing the distributions of Y_i given $X_i = 0$ (not UXO) and Y_i given $X_i = 1$ (not UXO):

$$f(y_i\,;\theta_0) \;\text{ and }\; f(y_i\,;\theta_1), \tag{16}$$

respectively, where θ_0 and θ_1 are the parameters of the conditional distributions. Implicit in the above is that the two density functions are from the same distributional family, that only the parameters θ_0 and θ_1 are different. The model could be made more general by allowing two completely different density functions, but using a single distributional family such as Normal or Beta should be adequate. For example, if Normal distributions are used, then $\theta_0 = (\mu_0, \sigma_0^2)$ and $\theta_1 = (\mu_1, \sigma_1^2)$.

The marginal density function of $Y_i = y_i$ (i.e., when X_i is not known) is the mixture density $g(y_i\,;\theta_0, \theta_1, p) = (1-p)f(y_i\,;\theta_0) + pf(y_i\,;\theta_1)$. The likelihood function of Equation 4 is replaced by

$$L(\theta_0, \theta_1, p\,;\mathbf{Y}) = \prod_{i=1}^{N} g(Y_i\,;\theta_0, \theta_1, p)\ , \tag{17}$$

which represents the data before any suspect locations have been dug up. The likelihood

function of Equation 5 is then replaced by

$$L(\theta_0, \theta_1, p\,; \mathbf{X}_D, \mathbf{Y}) = \prod_{i \in R} g(Y_i\,; \theta_0, \theta_1, p) \prod_{i \in D_0} ((1-p) f(Y_i\,; \theta_0)) \prod_{i \in D_1} (p f(Y_i\,; \theta_1)) \tag{18}$$

as suspect locations are dug, where D is the set of indices of the suspect locations that have been dug up, R are those that have not been dug up (the remaining), and D_1 and D_0 are the indices of the suspect locations that were found to be UXO or not, respectively. Prior distributions on θ_0 and θ_1 are the last inputs required for the Bayesian analysis. Consult Berger (1985) for mathematically convenient priors for the chosen distributional family.

The probability of no UXO remaining on the site of interest in Equation 6 is replaced by

$$P_0 = \prod_{i \in R} ((1-p) f(Y_i\,; \theta_0) / g(Y_i\,; \theta_0, \theta_1, p)) \ . \tag{19}$$

The terms in the product represent the posterior probabilities (given $\mathbf{Y}$ and the dig data $\mathbf{X}_D$) that the unsampled suspect location i is not UXO. Typically, these posterior probabilities are smaller for larger values of Y_i, so digging up the suspect locations with the largest values of Y_i will reduce the risk that any UXO remains.