C01

Proceedings of the 5th International Conference on Inverse Problems in Engineering: Theory and Practice, Cambridge, UK, 11-15th July 2005

A HIGH-LEVEL MODEL AND STATISTICAL ESTIMATION FOR PASSIVE GAMMA RAY TOMOGRAPHY OF NUCLEAR WASTE VAULTS

B. A. CATTLE[1], C. GODDARD[2] and R. M. WEST[3]
[1]*Department of Applied Mathematics, University of Leeds, Leeds, LS2 9JT, UK.*
e-mail: brian@maths.leeds.ac.uk
[2]*British Nuclear Group, Berkeley Centre, Berkeley, Gloucs., GL13 9PB, UK.*
e-mail: chris.goddard@britishnucleargroup.com
[3]*Bio-statistics Unit, University of Leeds, 30 Hyde Terrace, Leeds, LS2 9PL, UK.*
e-mail: r.m.west@leeds.ac.uk

Abstract - This paper demonstrates a parameter identification problem encountered in the assay of intermediate-level nuclear waste vaults. A high-level model of the vault contents is employed in which a likely geometry is assumed so as to reduce the number of unknown quantities in the resulting inverse problem. The inverse problem of identifying model parameters is solved with limited, but quantifiable precision.

1. INTRODUCTION

Transmission tomography, in which radiation is passed through a target object as in an X-ray photograph, is of long-standing interest in the nuclear community [8, 11]. Transmission tomography suffers from the drawback that it can be used only for the examination of small-scale containers and vessels and in particular, for each projection, a source must be placed on one side of the object in order to detect radiation that has passed through the object. Thus access to a considerable portion of the boundary is a pre-requisite for transmission tomography and so is not a practical option for many industrial targets such as storage vaults. Recently, *emission* tomography, in which radiation emitted by the target object is measured, has also been studied for the assay of large-scale nuclear process vessels and waste vaults [2]. In this paper, a model problem involving the assay of an intermediate-level waste vault is presented. The task of identifying physical variables from gamma ray count data can lead to a problem which is ill-posed.

A 'high-level' model is employed to model the vault and its contents in which the likely geometry is assumed, in contrast to low-level approaches where the domain, or at least the target itself, is discretized into pixels (2D picture elements) or voxels (3D volume elements). The advantage of the high-level approach is that only a small number of key process parameters need to be identified in the inverse identification problem. This is very important since the number of measurements recorded in industrial applications is typically small due to limited access to vaults owing to shielding materials for example. Consequently, with a pixellized approach, the scarcity of measurement data would lead to an irretrievably ill-posed inverse problem, whereas the high-level approach with a parsimonious parameterization, might be less ill-posed and yield parameter estimates with better precision.

In the direct problem, the number of counts, $\Psi(\boldsymbol{r})$, received at a gamma ray detector located at a position $\boldsymbol{r}$ the volume V', is calculated. An integral equation of *mono-energetic* gamma ray transport for the photon flux, $\psi(\boldsymbol{r})$, is employed, see [3, 18]. The required equation is

$$\psi(\boldsymbol{r}_D) = \int_{V'} S(\boldsymbol{r}') \frac{\mathrm{e}^{-\mu|\boldsymbol{r}_D - \boldsymbol{r}'|}}{4\pi|\boldsymbol{r}_D - \boldsymbol{r}'|^2} \mathrm{d}V', \tag{1}$$

where $\boldsymbol{r}_D$ is the location of a gamma ray detector in space, $S(\boldsymbol{r}')$ are the gamma ray sources at locations $\boldsymbol{r}'$ in space, μ is the linear attenuation coefficient and $\mathrm{d}V'$ is an element of the volume containing sources that is in the field-of-view of the detector. The number of counts, $\Psi(\boldsymbol{r}_D)$ (dimensionless), is obtained from $\psi(\boldsymbol{r}_D)$ by applying a point-spread function (PSF) to the flux $\psi(\boldsymbol{r}_D)$. The PSF describes the weighting of the detector response to radiation sources at all positions within the field-of-view. The use of the mono-energetic eqn.(1) is justified practically since detectors exist which can discriminate between gamma ray energies, see [9, 10].

A statistical approach to the inverse problem is adopted which allows the quantification of the reliability of reconstructions.

2. HIGH-LEVEL MODEL AND THE DIRECT PROBLEM

The geometry considered is similar to the 'layered-earth' models employed in geological resistivity imaging, and is described in terms of this particular application. The waste vault may be modeled as a cuboid region

$$\Lambda = \{(x, y, z) : 0 \leq x \leq L, 0 \leq y \leq W, 0 \leq z \leq H\}, \quad \text{where} \quad L > H > W. \tag{2}$$

The z coordinate is such that $z = 0$ at the base of the vault, and increases vertically upwards towards the surface of the waste at $z = H$. Typical dimensions of intermediate-level waste (ILW) vaults would be $L \approx 11$ m, $W \approx 3$ m and $H \approx 8$ m.

Two different processes have been used to fill the vaults leading to the contents being stratified into a two-layer structure, the locations of the layers being described by

$$\Lambda_1 = \{(x, y, z) : 0 \leq x \leq L, 0 \leq y \leq W, h_1 \leq z \leq H\}, \tag{3a}$$

$$\Lambda_2 = \{(x, y, z) : 0 \leq x \leq L, 0 \leq y \leq W, 0 \leq z < h_1\}, \tag{3b}$$

so that $\Lambda = \Lambda_1 \cup \Lambda_2$. The surface layer Λ_1, contains magnesium oxide (magnox) metal waste of emission density S_m Bq m^{-3}, with a proportion of nimonic alloy springs of emission density S_n Bq m^{-3}, with $S_\mathrm{m} \ll S_\mathrm{n}$. Note that the units of volumetric emission density are Becquerels per cubic metre. The linear attenuation coefficients of the magnox and nimonic materials are μ_m m^{-1} and μ_n m^{-1}, respectively, with $\mu_\mathrm{m} \sim O(\mu_\mathrm{n})$. The depth of the top layer is $h = H - h_1$. Note that in general layer Λ_2 is much deeper than Λ_1, see Figure 1. The substrate layer Λ_2, contains the same magnox metal with a smaller proportion of nimonic alloy springs.

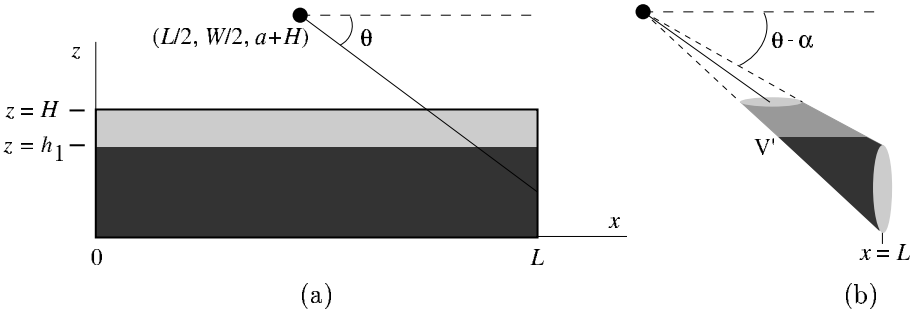


Figure 1: Schematics of the vault: (a) schematic of the vault in the (x, z)-plane; the coordinate y goes into the page and (b) a schematic of the volume V' of material containing sources that is viewed by the detector.

The source densities and linear attenuation coefficients are known for the magnox metal and nimonic springs. However, each layer contains a mixture of *both* of these constituents. Weighted source densities and linear attenuation coefficients are calculated for each layer, taking into account the quantity of each type of material present. The volume fraction of springs present in layers Λ_1 and Λ_2 is denoted by $\sigma_1(z)$ and $\sigma_2(z)$, repsectively and the packing fractions of the materials are denoted by $\rho_1(z)$ and $\rho_2(z)$, respectively. The weighted source and attenuations are given by

$$\left.\begin{aligned} S_1(z) &= \rho_1(\sigma_1 S_\mathrm{n} + (1 - \sigma_1) S_\mathrm{m}), \\ \mu_1(z) &= \rho_1(\sigma_1 \mu_\mathrm{n} + (1 - \sigma_1) \mu_\mathrm{m}), \end{aligned}\right\} \quad \text{in} \quad \Lambda_1, \tag{4}$$

for the surface layer and,

$$\left.\begin{aligned} S_2(z) &= \rho_2(\sigma_2 S_\mathrm{n} + (1 - \sigma_2) S_\mathrm{m}), \\ \mu_2(z) &= \rho_2(\sigma_2 \mu_\mathrm{n} + (1 - \sigma_2) \mu_\mathrm{m}), \end{aligned}\right\} \quad \text{in} \quad \Lambda_2, \tag{5}$$

for the substrate layer, where the argument z of the quantities σ and ρ has been suppressed for simplicity.

The detector is placed at a distance a m above the waste, i.e. at coordinates $(L/2, W/2, a + H)$. The detector makes scans at angles θ measured from $\theta = 0$ parallel to the x axis with θ increasing in a clockwise sense. The measurements of gamma ray counts are recorded according to a measurement protocol which examines a transect along the centre of the vault, i.e. parallel to the x axis at $y = W/2$. A number, N, of measurements are taken at equi-spaced intervals in θ. In this application N will be relatively small, certainly less than 100, say. Figure 1(a) shows the angle θ.

The volume integral in eqn.(1) is an integral over the volume of the waste that contains sources that are visible to the detector. In this case it is a truncated 'cone' whose vertex is at the detector, see Figure 1(b). The integral over this truncated cone is evaluated by employing a local system of cylindrical coordinates (r, ϑ, z) with its origin at $\boldsymbol{r}_D = (L/2, W/2, a + H)$ and the z axis aligned with the direction θ_j, i.e. the angle at which the j-th measurement is recorded. The volume element is therefore $\mathrm{d}V \equiv r\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}z$. The angle α is semi-angle of the (truncated) cone; the so-called *angle of acceptance*. Usually detectors are *collimated* to restrict α; indeed in this work $\alpha = 4\pi/180$, i.e. 4°. The evaluation of the integrals is performed numerically using Gaussian quadrature, [5, 17].

To obtain the number of counts $\Psi(\theta)$ from the flux $\psi(\theta)$ requires the application of a *point-spread* or detector response function, [9]. The application of the point spread function (PSF) is naturally accommodated by the convolution of $\psi(\theta)$ and the PSF, [3]. In order to simulate experimental noise, Gaussian random noise is added independently to each count. Although radiation emission is a Poisson process, in practice the number of observed counts is large enough to allow the large-mean Gaussian approximation to the Poisson distribution to be used, as in [2]. Therefore, the *measurements* are defined to be

$$\boldsymbol{\eta} = \boldsymbol{\Psi}(\theta) + \boldsymbol{\delta}(\xi^2) \tag{6}$$

where $\boldsymbol{\eta} \in \mathbb{R}^{N\times 1}$ are the measured gamma counts including noise, $\boldsymbol{\delta}(\xi^2) \in \mathbb{R}^{N\times 1}$ is a vector whose components are zero-mean Gaussian noise with variance ξ^2 and $\boldsymbol{\Psi}(\theta) \in \mathbb{R}^{N\times 1}$ are the noise-free counts given by the convolution of $\psi(\theta)$ with the point-spread function.

The estimation of a number m of key process quantities, represented by the vector $\boldsymbol{\beta}$, *given* the measurements $\boldsymbol{\eta}$, is the determination of a cause *given* the effect; in essence an inverse problem.

3. ANALYSIS

The measurements (dependent variables) $\boldsymbol{\eta}$ can be thought of as depending on the independent variables $\boldsymbol{\beta}$, called the state variables. In this case the state variables are the gamma ray emission densities of each layer (S_1 and S_2), the linear attenuation coefficients (μ_1 and μ_2) and the depth h of layer Λ_1. The state variables are assembled into the vector $\boldsymbol{\beta} = (h, \mu_1, \mu_2, S_1, S_2)^T$ for which an estimate, $\widehat{\boldsymbol{\beta}}$ is sought; for a detailed account see [1].

In the model problem presented here, the state variables employed in the direct problem are calculated from the values $S_\mathrm{m} = 1$ Bq m^{-3}, $S_\mathrm{n} = 450$ Bq m^{-3}, $\mu_\mathrm{m} = 1$ m^{-1} and $\mu_\mathrm{n} = 1.5$ m^{-1} via eqns (4) and (5), giving $\boldsymbol{\beta} = (0.50, 0.67, 0.68, 29.84, 6.79)^T$. With this information, the measurement curve for the angle θ plotted against the measurement $\boldsymbol{\eta}$ appears as in Figure 2, with 10 % random noise as an example value.

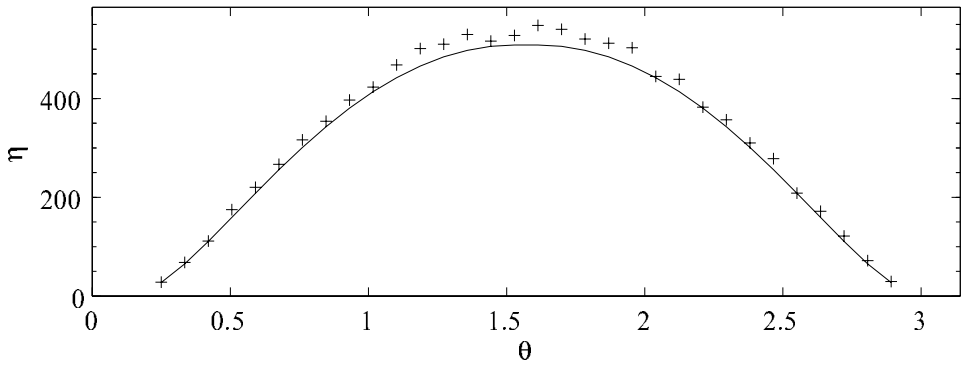


Figure 2: Number of counts $\boldsymbol{\Psi}$ (solid line) and $N = 32$ noise-corrupted measurements $\boldsymbol{\eta}$ (plus signs), plotted against the angle θ.

The sensitivity of the measurements $\boldsymbol{\eta}$ to perturbations of the state variables, i.e. the components of $\boldsymbol{\beta}$, may be examined by calculating the sensitivity relations and assembling these into a matrix $\mathbf{Z} \in \mathbb{R}^{N\times m}$, gives

$$\mathbf{Z}(\eta_j, \beta_k) = \mathbf{Z}_{jk} = \frac{\partial \eta_j}{\partial \beta_k}, \quad \text{for} \quad j = 1, \ldots, N \quad \text{and} \quad k = 1, \ldots, m; \tag{7}$$

where N is the number of measurements taken and m is the number of quantities to be estimated, here $m = 5$. The sensitivity relations are computed by a central-difference approximation. The matrix $\mathbf{Z}$ can

be viewed as the Jacobian of a transformation from the state variables, i.e. the components of $\boldsymbol{\beta}$, to the measurements $\boldsymbol{\eta}$.

The singular value decomposition of the matrix $\mathbf{Z}$ is defined by,

$$\mathbf{Z} = \mathbf{USV}^T = \sum_{i=1}^{5} \boldsymbol{u}_i s_i \boldsymbol{v}_i^T. \tag{8}$$

Analysis of the singular values, s_i, and right singular vectors, $\boldsymbol{v}_i^T$, of the $(N \times m)$ matrix $\mathbf{Z}$, can give an indication of which quantities may be difficult to estimate with good confidence, for full details see [6]. The singular value decomposition of the (32×5) matrix $\mathbf{Z}$ yields singular values, s_i, shown in the log-linear plot below; Figure 3(a). Figure 3(b) shows the matrix $\mathbf{Z}$ interpolated onto a fine grid for visualization, whilst (c) and (d) show examples of the right singular vectors $\boldsymbol{v}_i^T$.

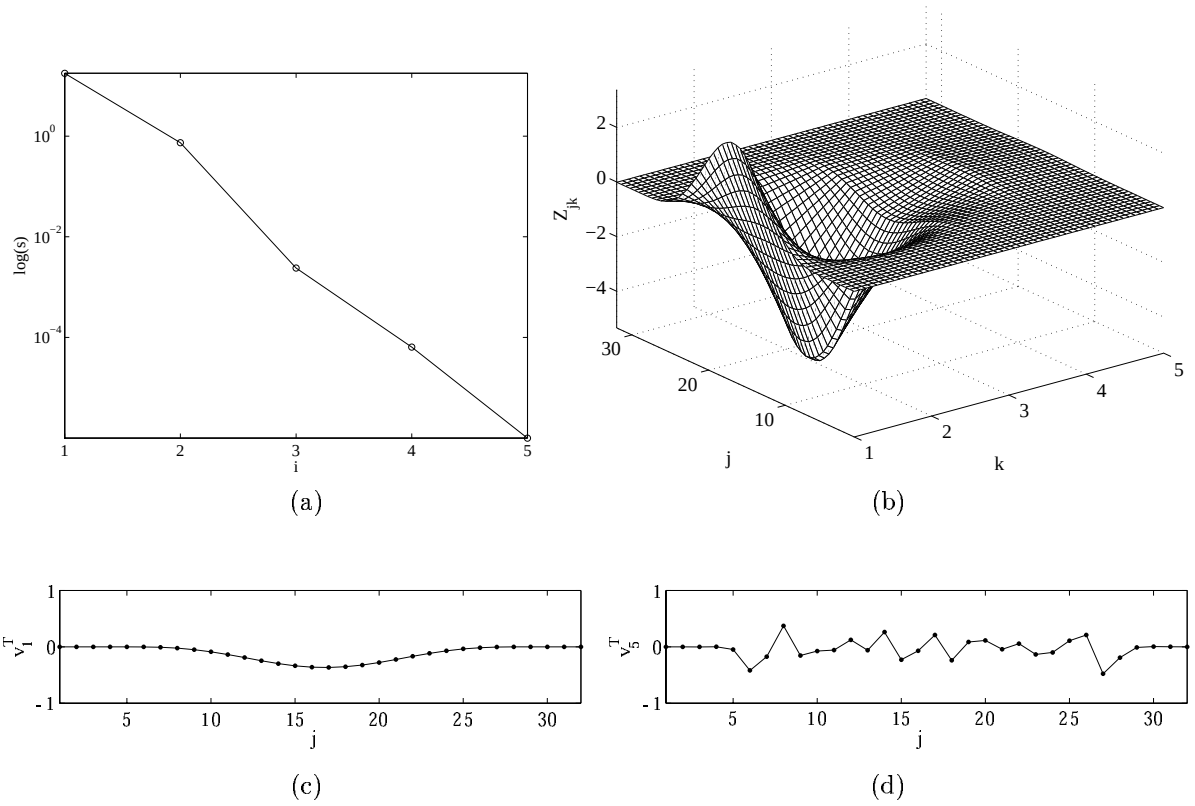


Figure 3: The sensitivity matrix $\mathbf{Z}$; in (a) the singular values and in (b) the matrix interpolated onto a fine grid for visualization. The right singular vectors scaled to the canonical interval $[-1, 1]$, corresponding to (c) the largest singular value, s_1, and (d) the smallest singular value, s_5.

By mathematical convention the singular values are arranged in non-increasing order and therefore their ordering does not necessarily match the order of the components of the vector $\boldsymbol{\beta}$. Consequently, a new index i is introduced for the singular values and vectors as an *aide memoire* of this point.

In Figure 3(b), the index k *does* correspond with the components of $\boldsymbol{\beta} = (h, \mu_1, \mu_2, S_1, S_2)^T$ and j is the measurement index; $j = 1, \ldots, 32$. Examination of Figure 3(b) shows that only the columns of $\mathbf{Z}$ relating to β_1, β_2 and β_3 exhibit non-zero values. The two source densities, $\beta_4 = S_1$ and $\beta_5 = S_2$ have $\mathbf{Z}_{j4} \approx 0$ and $\mathbf{Z}_{j5} \approx 0$ for all j. This corresponds with the two smallest singular values (s_4 and s_5) in Figure 3(a).

Estimation of the components of $\boldsymbol{\beta}$ corresponding to the smallest singular values will be difficult or impossible due the components of the measurements in those directions, i.e. the corresponding $\boldsymbol{v}_i^T$, being strongly affected by the noise. This can be demonstrated by examining the vectors $\boldsymbol{v}_1^T$ and $\boldsymbol{v}_5^T$. Figure 3(c) shows the vector $\boldsymbol{v}_1^T$ corresponding to the *largest* singular value and which is smooth in appearance. Figure 3(d) shows $\boldsymbol{v}_5^T$ which corresponds to the *smallest* singular value and which oscillates due to the

effects of random noise. State variables corresponding to small singular values are often difficult to estimate with good confidence.

To determine an estimate $\widehat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$, the arguments $\boldsymbol{\beta}$ which minimize an objective function of the form

$$E(\boldsymbol{\beta}) = \frac{1}{\boldsymbol{\eta}} \|f(\boldsymbol{\beta}) - \boldsymbol{\eta}\|_2^2 + \lambda \|\boldsymbol{\beta} - \widetilde{\boldsymbol{\beta}}\|_p^2, \quad \lambda \geq 0, \tag{9}$$

must be performed, where $\widetilde{\boldsymbol{\beta}}$ is a vector encoding the prior information. In eqn.(9) $\boldsymbol{\eta}$ and $\boldsymbol{\beta}$ have their usual meanings, whilst $f : \mathbb{R}^{m\times 1} \rightarrow \mathbb{R}^{N\times 1}$, refers to direct solution of eqn.(1) with arguments $\boldsymbol{\beta}$, i.e. the measurement model. The penalty functional $\lambda\|\boldsymbol{\beta} - \widetilde{\boldsymbol{\beta}}\|_p^2$ increases the value of the objective function if $\widehat{\boldsymbol{\beta}}$ is not 'close' to the prior information thereby regularizing the solution. The p-norm in eqn.(9) is usually taken to be the Euclidean 2-norm, however a sum of moduli 1-norm can also be employed to good effect. The sum of squares of errors in eqn.(9), is weighted by the observed counts to adjust for the variance being proportional to the count.

There are many methods for the optimization of functions of the form of eqn.(9), see for example [4]. However, the analysis performed above which culminated in Figures 3(a) and (b), gives an indication that gradient-based optimization methods will fail due to approximately zero gradients with respect to two of the components of $\boldsymbol{\beta}$. Therefore a gradient-free optimization method should be employed.

4. BAYESIAN APPROACH TO THE INVERSE PROBLEM

Progress can be made if a change of paradigm is considered. Rather than considering the components of $\boldsymbol{\beta}$ as a vector of deterministic values, they are now thought of as random variables, allowing Bayesian statistical methods to be applied.

Each of the N measurements will be regarded as being distributed independently of other measurements. In fact each measurement will have a Poisson distribution (see for example [12]) but for convenience will be approximated with a Gaussian distribution with mean $f_i(\boldsymbol{\beta})$ (for $i = 1, \ldots, N$) and variance $f_i(\boldsymbol{\beta})$, see eqn.(6). This defines the *likelihood*, i.e. the probability of the measurements $\boldsymbol{\eta}$ being recorded *given* the values $\boldsymbol{\beta}$. The Gaussian likelihood is denoted by $\mathcal{L}(\boldsymbol{\beta}|\boldsymbol{\eta})$ and is given by

$$\mathcal{L}(\boldsymbol{\beta}|\boldsymbol{\eta}) = \left(2\pi \det(\mathbf{C}_\xi)\right)^{-N/2} \exp\left(-\frac{1}{2}(f(\boldsymbol{\beta}) - \boldsymbol{\eta})^{\mathrm{T}} \mathbf{C}_\xi^{-1} (f(\boldsymbol{\beta}) - \boldsymbol{\eta})\right), \tag{10}$$

where $f(\boldsymbol{\beta})$ represents the direct (measurement) model and $\mathbf{C}_\xi$ is the variance–covariance matrix of the measurements. Since the measurements are assumed to be independently distributed, the covariance matrix takes the form $\mathbf{C}_\xi = \xi^2 \mathbf{I}_N$, where $\mathbf{I}_N$ is the identity matrix in $\mathbb{R}^{N\times N}$. It is assumed that some prior information is known about the components of $\boldsymbol{\beta}$. This could take a number of forms. Here a very simple form only is considered for clarity of exposition. It is assumed that each element of the parameter vector is independently distributed about a mean $\widetilde{\beta}_k$ and variance ζ_k^2, mathematically expressed by

$$\mathcal{P}(\beta_k) = \left(2\pi \det(\mathbf{C}_\zeta)\right)^{-m/2} \exp\left(-\frac{1}{2\zeta_k^2}\|\beta_k - \widetilde{\beta}_k\|_2^2\right), \tag{11}$$

where $\widetilde{\beta}_k$ are *a priori* estimates of the state variable β_k and $\mathbf{C}_\zeta$ is the variance–covariance matrix of the prior distribution. The use of eqns (10) and (11) is becoming more common in tomographic inverse problems, and implicitly underpin all 'regularized least squares' methods. The *posterior density function* describes the distribution of the state variables given the prescribed prior knowledge. The posterior density $\mathcal{Q}(\boldsymbol{\beta}|\boldsymbol{\eta})$, assuming that prior and likelihood distributions are independent, is the product of the likelihood and the prior, that is

$$\mathcal{Q}(\boldsymbol{\beta}|\boldsymbol{\eta}) = \mathcal{L}(\boldsymbol{\beta}|\boldsymbol{\eta})\mathcal{P}(\boldsymbol{\beta}). \tag{12}$$

The interpretation of eqn.(12) is that the solution of the inverse problem is now the posterior probability of the variables $\boldsymbol{\beta}$, conditioned on the observed gamma ray counts $\boldsymbol{\eta}$. By sampling the posterior distribution, estimates of the values of the random variables $\boldsymbol{\beta}$ are determined. Full details of the Bayesian approach may be found in [13].

Assuming for simplicity that the prior knowledge of one parameter is independent of all of the others, the prior variance–covariance matrix is $\mathbf{C}_\zeta = \zeta_k \mathbf{I}_m$. Taking the logarithm of the product of eqns (10) and (11) defines the log-posterior density given by

$$\log \mathcal{Q}(\boldsymbol{\beta}|\boldsymbol{\eta}) = \text{constant} - (f(\boldsymbol{\beta}) - \boldsymbol{\eta})^{\mathrm{T}} (f(\boldsymbol{\beta}) - \boldsymbol{\eta}) - \sum_{k=1}^{5} \frac{2\xi^2}{\zeta_k^2} \left(\beta_k - \widetilde{\beta}_k\right)^2, \tag{13}$$

in which the ratios of the observation variance ξ^2 and the prior density variances ζ_k^2, conveniently defines a regularization parameter for the variable β_k, i.e. $\lambda_k = 2\xi^2/\zeta_k^2$.

4.1. Estimation

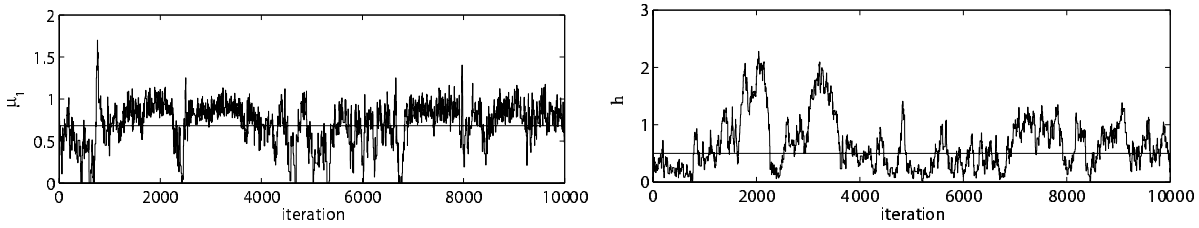
Although the prior and the likelihood are both Gaussian distributions, the likelihood is Gaussian with respect to the measurements $\boldsymbol{\eta}$ whilst the prior is Gaussian with respect to the state variables $\boldsymbol{\beta}$. One approach to estimation is to approximate the posterior with a Gaussian curve and calculate the posterior mean and variance-covariance matrix. This leads to the calculation of regularized least squares estimates which are commonly employed. It will be seen however in Subsection 4.2 that the posterior in this case is not symmetric and therefore the Gaussian approximation to the posterior would be biased and misleading.

Markov chain Monte Carlo techniques have been developed [14, 7] to generate samples from the posterior so that sufficient samples can be obtained in order to describe the posterior distribution of the state variables.

Sampling the posterior is accomplished computationally by the Metropolis–Hastings algorithm which simulates Markov chains in the state space, [14, 7]. The Markov chains, with runs of sufficient length, explore a large range of feasible estimates for $\boldsymbol{\beta}$. At at each step of the algorithm the *regularized energy*, defined by eqn.(13), is computed. At each iteration, a randomly generated perturbation to the model f is proposed, and the consequent change in the energy ΔE is determined. If $\Delta E < 0$, i.e. the perturbation results in a lower energy, the new state is automatically accepted. If $\Delta E \geq 0$, the new state is accepted with a probability $P = \exp(-\Delta E)$. Therefore the Metropolis–Hastings algorithm can accept a perturbation of the current state which results in an increase of energy, i.e. $\Delta E > 0$. This makes the Metropolis–Hastings algorithm a robust, gradient-free method for the optimization of objective functions with multiple local minima, and/or indistinct global minima.

4.2. Results

A simplex method is employed from a random starting state to give an initial estimate, $\widehat{\boldsymbol{\beta}}_0$, which is employed as the starting state of the Metropolis–Hastings algorithm. The Markov chain produced by the Metropolis–Hastings algorithm does not converge but with a sufficiently long run, the mean of the chain is calculated and this mean will converge to the mean of the posterior distribution. The most commonly employed summary of the posterior distribution is the posterior mean, denoted by $\widehat{\boldsymbol{\beta}}_M$. The values $\boldsymbol{\beta}$ which give rise to the smallest value of the energy as per eqn.(9), defines the maximum a posteriori (MAP) estimate, denoted by $\widehat{\boldsymbol{\beta}}_{MAP}$.



(a) (b)

Figure 4: Examples of the Markov chains for (a) the attenuation coefficient μ_1 and (b) the depth, h, of layer Λ_1; 10000 chain samples are employed.

Figure 4 shows two examples of the Markov chains for (a) μ_1 and (b) h. The chains can be seen to vary in the vicinity of 'target' values indicated by the horizontal line. Table 1 summarizes the salient features of the marginal posterior distributions of the estimated values $\widehat{\boldsymbol{\beta}}$.

Note that Table 1 shows that the state variables μ_1 and h have the least degree of regularization: that is the prior information has less influence for these variables. It is indeed these two state variables that it is expected that the measurements will provide the most information, since the layer Λ_1 to which they relate is closest to the detector.

The most important information in Table 1 are the credibility regions. Essentially these give the degree of reliability of the estimates. For example, there is a 95 % chance that the value of S_1 is in the range $(30.61, 31.55)$. The size of the credibility interval may be reduced if a reduction in the degree of reliability of the estimates is accepted. For example, if a 90 % credibility region is adopted, then there is a 90 % chance that S_1 lies in the smaller interval $(30.65, 31.48)$. From an industrial standpoint, the provision of

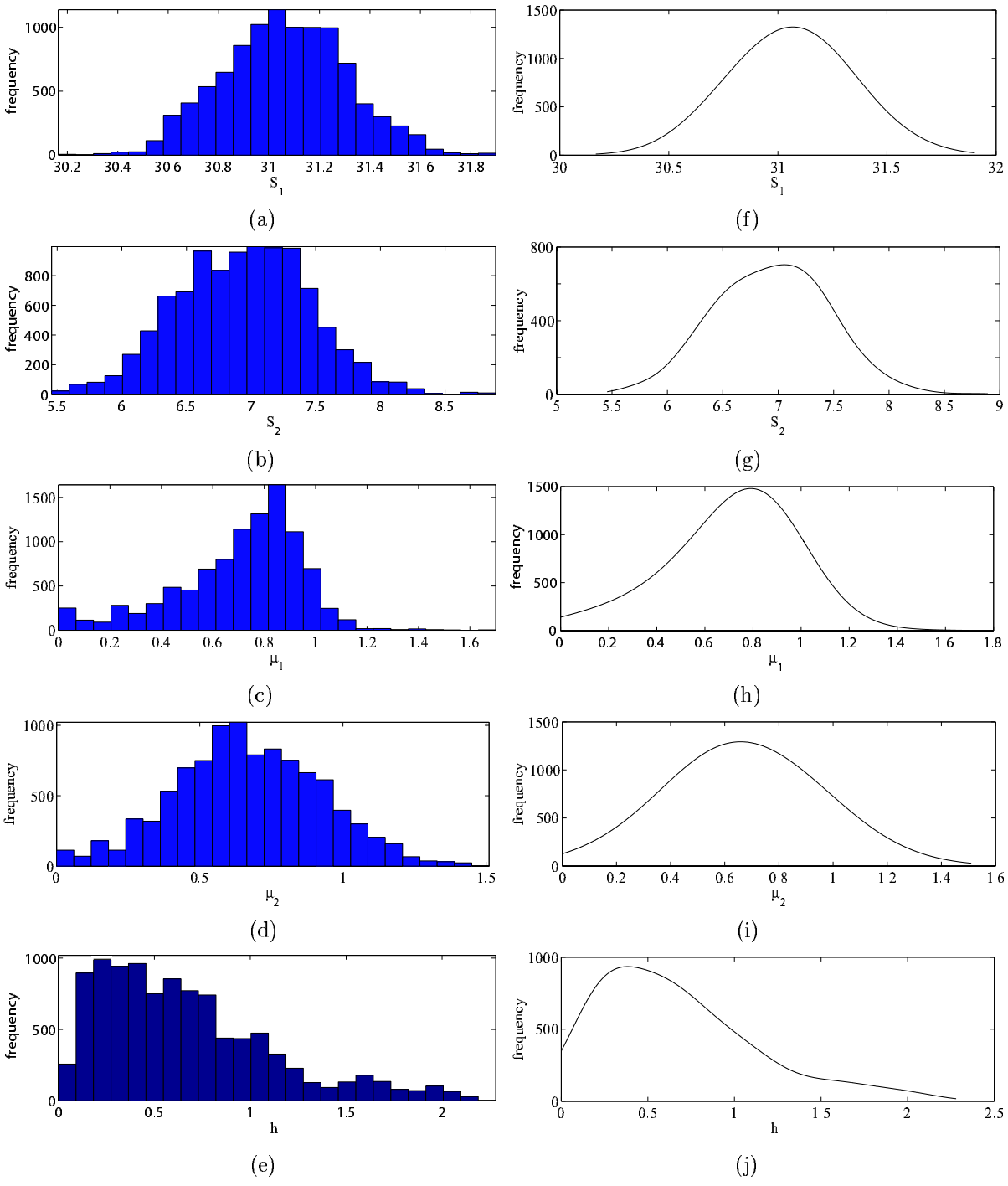


Figure 5: Examples of histograms posterior marginal distributions for (a) the source S_1, (b) the source S_2, (c) the attenuation coefficient μ_1, (d) and (e) the depth, h, of layer Λ_1; 10000 chain samples are employed. Figures (f) to (j) show the marginal distributions as determined by kernel smoothing.

Table 1: Regularization parameters $\lambda_k = 2\xi^2/\zeta_k^2$ and estimated values $\widehat{\boldsymbol{\beta}}$ with 90 % and 95 % credibility regions.

$\boldsymbol{\beta}$	λ_k	$\widehat{\boldsymbol{\beta}}_0$	$\widehat{\boldsymbol{\beta}}_M$	$\widehat{\boldsymbol{\beta}}_{MAP}$	90 % credibility region	95 % credibility region
S_1	8.0	33.57	31.06	31.10	(30.65,31.48)	(30.61,31.55)
S_2	2.0	5.82	6.94	7.02	(6.12,7.76)	(5.95,7.92)
μ_1	0.9	0.68	0.79	0.69	(0.21,1.01)	(0.06,1.06)
μ_2	8.0	0.89	0.67	0.61	(0.25,1.09)	(0.16,1.17)
h	0.5	0.41	0.66	0.46	(0.22,1.41)	(0.09,1.85)

the credibility regions is an extremely useful diagnostic tool since it provides a quantitative measure of the reliability of the estimation technique, which is often based upon very restricted or poor-quality data.

It can be extremely useful to examine the marginal posterior distributions graphically, as in Figure 5. It can be seen that the histogram of the posterior distribution of μ_1 in Figures 5(c) and (h) is unimodal with relatively little spread as compared with the distribution of, say h; Figures 5(e) and (j). The mode of the posterior marginal distribution of $\boldsymbol{\beta}_3 = \mu_1$ is a little larger than 'target' value of $\mu_1 = 0.67$.

Figures 5 (e) and (j) shows the histogram for the distribution of h. The spread is much larger than that observed for μ_1, but the mode corresponds approximately to the 'target' of $h = 0.50$. The graphical representation of the marginal distributions gives a qualitative feel for the distribution of the estimated values, and should be used as a visual reinforcement for the data in Table 1. Note that the distribution of h is skewed to the right. The computational effort for the numerical evaluation of the posterior is therefore justified.

It can be seen for example that although the bulk of the posterior distribution suggests that h has a relatively low value, around 0.5 m, it is credible that h could be much larger, around 1.25 m say, given the data collected. This has implications for strategy should it be necessary to retrieve waste from the vault.

Correlations between state variables can also be determined from the chain. This shows that there is little correlation between most variables (absolute value of the correlation is less than 0.1) except that the correlation between h and μ_1 which is high (0.6608). Consequently greater prior information concerning the attenuation in the upper layer would increase the precision with which h can be estimated.

5. CONCLUSIONS

A high-level model for the contents of an intermediate-level nuclear waste vault has been considered. The high-level approach is necessitated as a consequence of the available data being of limited quantity and quality. The resulting inverse problem then contains only a few key quantities which need to be identified.

It was shown that despite the high-level model, some quantities are difficult to obtain by traditional gradient-based optimization approaches, due to zero gradients with respect to two of the state variables.

A paradigm shift towards considering the quantities to be estimated as random variables allows Bayesian statistics to be employed. The Bayesian approach, which is implemented computationally by a Markov chain Monte-Carlo method allows greater robustness and flexibility in summarizing the reconstructed values. In particular the posterior distribution is skewed with respect to at least one of the key state variables (h). The presentation of credibility intervals and/or marginal distributions provides valuable information for industrial utilization.

Acknowledgements

BAC gratefully acknowledges the financial and technical support of British Nuclear Fuels plc., British Nuclear Group and the Smith Institute. We should like to thank the referees for their comments.

REFERENCES

1. J. V. Beck and K. J. Arnold, *Parameter Estimation in Engineering and Science*, John Wiley, New York, 1977.

2. B. A. Cattle, A. S. Fellerman and R. M. West, On the detection of solid deposits using gamma ray emission tomography with limited data. *Meas. Sci. Technol.* (2004) **15**, 1429–1439.

3. B. A. Cattle, *Integral Equations for Gamma Ray Transport and Detection*, Department of Applied Mathematics, University of Leeds: Internal Report, 2005.

4. L. R. Foulds, *Optimization Techniques*, UTM Series, Springer Verlag, New York, 1981.

5. W. Gander and W. Gautschi, Adaptive Quadrature: Revisited, *BIT* (2000) **40**, 84-101.

6. P. C. Hansen, *Regularization Tools: A Matlab Package for the Analysis of Ill-Posed Problems*, Denmark Technical University, 1992. Web URL: `www.imm.dtu.dk/pch`.

7. W. K. Hastings, Monte Carlo sampling methods using Markov chains, and their applications, *Biometrika* (1970) **57**, 97–109.

8. J. Huddleston and I. G. Hutchinson, *The Potential Use of Tomographic Techniques for the Quality Checking of Cemented Waste Drums*, Internal report: AERE Harwell, UK, 1985.

9. K. A. Hughes, and J. A. Lightfoot, RadScan 600: A portable instrument for the remote imaging of gamma contamination: Its design and use in aiding decommissioning strategy. *IEEE Nuclear Science Symp.*, San Francisco, 1996.

10. K. A. Hughes, and J. A. Lightfoot, Upgrading the RadScan 600 gamma scanner to produce dose maps in three-dimensions, *Proc. Fifth Int. Conf. on Nuclear Engineering*, 1997.

11. R. A. James, C. H. Zimmerman, N. H. Merrill, S. D. Vasey and G. Monk, Final Report on the Application of High Resolution Gamma Spectrometry to Quality Checking of Cement Encapsulated ILW and Vitrified HLW, *Sellafield Research and Development Research Memorandum* QCTF(88)P3, BNFL Sellafield, UK, 1988.

12. G. F. Knoll, *Radiation Detection and Measurement*, John Wiley and Sons, London, 1979.

13. J. S. Liu, *Monte Carlo Strategies in Scientific Computing*, Springer–Verlag, New York, 2001.

14. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, Equations of state calculations by fast computing machine. *J. Chem. Phys.* (1953) **21**, 1087–1091.

15. F. Natterer and F. Wübbeling, *Mathematical Methods in Image Reconstruction*, SIAM, Philidelphia PA, 2001.

16. J. A. Nelder and R. Mead, A simplex method for function minimization. *Computer J.* (1965) **7**, 308–13.

17. L. N. Trefethen, *Spectral Methods in Matlab*, SIAM, Philadelphia PA, 2000.

18. J. Wood, *Computational Methods in Reactor Shielding,* Pergamon Press, Oxford, 1982.