

Research Article

Comparison of Two Approaches for Detection and Estimation of Radioactive Sources

Eugene Yee,¹ Ajith Gunatilaka,² and Branko Ristic³

¹ *Physical Protection Section, Defence Research and Development Canada-Suffield, P.O. Box 4000 Stn Main, Medicine Hat, AB, Canada T1A 8K6*

² *Human Protection and Performance Division (HPPD), Defence Science and Technology Organization, 506 Lorimer Street, Fishermans Bend, VIC 3207, Australia*

³ *Intelligence Surveillance and Reconnaissance Division (ISRD), Defence Science and Technology Organization, 506 Lorimer Street, Fishermans Bend, VIC 3207, Australia*

Correspondence should be addressed to Eugene Yee, eyee0309@gmail.com

Received 25 May 2011; Accepted 17 June 2011

Academic Editors: Y. Dimakopoulos and L. Guo

Copyright © 2011 Her Majesty the Queen in Right of Canada. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper describes and compares two approaches for the problem of determining the number of radioactive point sources that potentially exist in a designated area and estimating the parameters of these sources (their locations and strengths) using a small number of noisy radiological measurements provided by a radiation sensor. Both approaches use the Bayesian inferential methodology but sample the posterior distribution differently: one approach uses importance sampling with progressive correction and the other a reversible-jump Markov chain Monte Carlo sampling. The two approaches also use different measurement models for the radiation data. The first approach assumes a perfect knowledge of the data model and the average background radiation level, whereas the second approach quantifies explicitly the uncertainties in the model specification and in the average background radiation level. The performances of the two approaches are compared using experimental data acquired during a recent radiological field trial.

1. Introduction

There is growing concern in recent years regarding the risk of the smuggling and illicit trafficking of stolen radiological material. This state of affairs has heightened the spectre of radiological terrorism involving the use of an improvised radiological dispersal device (dirty bomb) to disperse radiological material over a large area using the force of conventional explosives (see, Panofsky [1]). In this context, the detection, localisation, and characterization of sources of radioactive material using radiation sensors (or, networks of such sensors) have important implications for safety and public security.

Various researchers have focussed on the problem of recovering the physical location and strength of radioactive sources using measurements obtained from a number of radiological detectors. Howse et al. [2] applied a nonlinear, recursive least-squares algorithm for tracking the positions of a moving radioactive point source. Stephens Jr. and Peurrung [3] conducted a cost-benefit analysis involving a variety of (conflicting) operational requirements that need to be considered in the design and deployment of a network of radiation sensors for use in the detection of a moving radioactive source. Nemzek et al. [4] and Brennan et al. [5] investigated the application of a Bayesian methodology for the detection and estimation of mobile radioactive sources, with a specific focus on examination of the signal-to-noise ratio envelope (and the concomitant detection limits) expected from utilization of a distributed sensor network for detection of a radiological source in a moving vehicle. Gunatilaka et al. [6] proposed the application of a maximum likelihood estimation algorithm, extended Kalman filter, and unscented Kalman filter for the localisation of a single static radiological point source. Morelande et al. [7, 8] applied a Bayesian methodology realized using a particle filtering technique with progressive correction to localize multiple static radiological point sources. Finally, Mendis et al. [9] applied binary and continuous genetic algorithms to solve a maximum likelihood estimation problem for recovery of the location and strengths of an *a priori* known number of static radiological point sources.

In this paper, we describe two approaches for radiological source localisation for the case when the number of sources is unknown *a priori* and compare their performance by applying them to an experimental data set collected in a field trial in Puckapunyal Military Area (Victoria, Australia). Both approaches use the Bayesian inferential methodology but sample the posterior distribution using different algorithms. In the first approach, the sampling algorithm uses importance sampling with progressive correction (PC), while the second approach uses the Markov chain Monte Carlo (MCMC) algorithm for the sampling. While the first algorithm uses the minimum description length (MDL) to estimate the unknown number of sources, the second algorithm uses the reversible-jump MCMC approach for this purpose. Finally, the measurement model for the radiation data differs for the two approaches in the following way: the first approach uses a form of the posterior distribution that implicitly assumes that the average background radiation count rate is exactly known and that the error in the model for the radiation data is negligible, whereas the second approach utilizes a form of the posterior distribution that relaxes these assumptions.

The experimental data for evaluating the two approaches were acquired using a Geiger-Müller detector in the presence of multiple radiological point sources of unequal strengths. Background radiation data in the absence of these radiation sources were also acquired to provide an estimate of the average background radiation level.

The organisation of the paper is as follows. The radiological source localisation problem is formulated in Section 2. Section 3 describes the two approaches we use for radiological source estimation. Section 4 describes the radiation field trial and experimental data sets. Some implementation aspects of the two algorithms are discussed in Section 5. Section 6 presents the results of the application of the two approaches for source reconstruction to the experimental data. Conclusions are drawn in Section 7.

2. Problem Formulation

The problem of the detection and estimation of the characteristics (e.g., location, activity) of an unknown number of static radiological point sources from a finite number of “noisy”

measurements (data) obtained from radiation sensors is addressed using the Bayesian formalism. This formalism involves the application of Bayes' rule:

$$p(\Theta | \mathbf{z}, I) \propto p(\Theta | I)p(\mathbf{z} | \Theta, I), \quad (2.1)$$

where Θ are the parameters that we are trying to determine and \mathbf{z} is the data. In (2.1), I is the background (contextual) information available in the problem, “|” denotes “conditional upon”, and the various factors that appear in this relation have the following interpretation: $p(\Theta | I)$ is the prior probability for the parameters Θ , $p(\mathbf{z} | \Theta, I)$ is the likelihood function which specifies the probability that we observe data \mathbf{z} when Θ is known exactly, and $p(\Theta | \mathbf{z}, I)$ is the posterior probability for the parameters Θ in light of the new information introduced through the acquired data \mathbf{z} . In essence, Bayes' rule informs us how to update our knowledge about the parameters Θ encoded in $p(\Theta | I)$ following the acquisition of data \mathbf{z} . To apply the Bayesian formalism to the radiological source reconstruction problem, we need to assign appropriate functional forms for the prior probability and the likelihood function, which in turn determines the posterior probability.

2.1. Assignment of Prior Probability

Here, we assume that a number r of point sources of gamma radiation are present in a flat open area that is free of any obstructions. The number of sources, r , is unknown. If $r > 0$, the i th point source of gamma radiation (where $i = 1, 2, \dots, r$) is parameterised by

- (i) its location $\mathbf{x}_{s,i} \equiv (x_{s,i}, y_{s,i})$ in a Cartesian coordinate system with $(x_{s,i}, y_{s,i}) \in \mathcal{A}$, where $\mathcal{A} \subset \mathbb{R}^2$ is some large region that is assumed to contain all the sources,
- (ii) its intensity rate (source strength), $Q_{s,i}$.

The parameter vector of source i is thus given by

$$\theta_i = (x_{s,i} \ y_{s,i} \ Q_{s,i})^\top, \quad (2.2)$$

where \top denotes the matrix transpose. The source parameter vectors are collected into a stacked vector $\theta(r) = (\theta_1^\top \ \dots \ \theta_r^\top)^\top$, and let $\Theta \equiv (r \ \theta(r)^\top)^\top$.

If we assume the logical independence of the various source parameters (or equivalently, the components of Θ), the prior probability for the parameters factorizes as follows:

$$p(\Theta | I) = p(r | I) \prod_{i=1}^r p(Q_{s,i} | I) p(\mathbf{x}_{s,i} | I). \quad (2.3)$$

Now, we need to assign explicit functional forms for each of the component prior probabilities in (2.3).

The prior distribution of r , $p(r | I)$, is assigned as a binomial distribution:

$$p(r | I) = \frac{(r_{\max} - r_{\min})!}{(r - r_{\min})!(r_{\max} - r)!} p^{*(r-r_{\min})} (1 - p^*)^{r_{\max}-r}, \quad (2.4)$$

($r = r_{\min}, r_{\min} + 1, \dots, r_{\max}$). In (2.4), $p^* \in (0, 1)$ is the binomial rate, and, r_{\min} and r_{\max} are, respectively, the minimum and maximum number of sources.

Two alternative functional forms are used for the prior distribution of $Q_{s,i}$ ($i = 1, 2, \dots, r$). The first of the two alternatives assigns the prior for the source strength $Q_{s,i}$ to be a gamma distribution:

$$p(Q_{s,i} | I) = Q_{s,i}^{\kappa-1} \frac{\exp(-Q_{s,i}/\psi)}{\Gamma(\kappa)\psi^\kappa}, \quad i = 1, 2, \dots, r, \quad (2.5)$$

where κ and ψ are the shape and scale parameters, respectively, and $\Gamma(x)$ is the gamma function. For the second of the two alternatives, the prior distribution for $Q_{s,i}$ is assigned a Bernoulli-uniform mixture model:

$$p(Q_{s,i} | I) = (1 - \gamma)\delta(Q_{s,i}) + \gamma \frac{\mathbb{I}_{(0, Q_{\max})}(Q_{s,i})}{Q_{\max}}, \quad i = 1, 2, \dots, r, \quad (2.6)$$

where γ is the probability that the source is active ($\Pr\{Q_{s,i} > 0\} = \gamma$), $\delta(x)$ is the Dirac delta function, and Q_{\max} is an *a priori* upper bound on the expected source intensity rate. In (2.6), $\mathbb{I}_A(x)$ denotes the indicator (characteristic) function for set A , with $\mathbb{I}_A(x) = 1$ if $x \in A$ and $\mathbb{I}_A(x) = 0$ if $x \notin A$.

Finally, the prior distribution for the source location $\mathbf{x}_{s,i}$ is assigned to be uniformly distributed over the region $\mathcal{A} \subset \mathbb{R}^2$ that is assumed *a priori* to contain the source:

$$p(\mathbf{x}_{s,i} | I) = \frac{\mathbb{I}_{\mathcal{A}}(\mathbf{x}_{s,i})}{\text{meas}(\mathcal{A})}, \quad i = 1, 2, \dots, r, \quad (2.7)$$

where $\text{meas}(\mathcal{A})$ is the area of the region \mathcal{A} .

2.2. Assignment of Likelihood Function

To prescribe a functional form for the likelihood function, we need to relate the hypotheses of interest about the unknown source(s)¹ as encoded in Θ to the available radiation data \mathbf{z} measured by detectors emplaced in the vicinity of the source(s). Towards this objective, we need to formulate a measurement model for the radiation data \mathbf{z} .

The radiation counts from nuclear decay obey Poisson statistics (see, Tsoulfanidis [10]), implying that the probability that a gamma radiation detector registers $z \in \mathbb{N} \cup \{0\}$ counts ($\mathbb{N} \cup \{0\}$ being the set of natural numbers *including zero*) in an exposure time of τ seconds, from a source that emits on average μ counts per second, is

$$\mathcal{P}(z; \lambda) = \frac{\lambda^z}{z!} \exp(-\lambda), \quad (2.8)$$

where $\lambda = \mu\tau$ is the parameter (both the mean and the variance) of the Poisson distribution.

The measurements of the radiation field are made using a Geiger-Müller (GM) counter. Let $z_j \in \mathbb{N} \cup \{0\}$ ($j = 1, \dots, m$) be a measured count from the GM counter, taken at (receptor) location $(x_{l,j}, y_{l,j}) \in \mathbb{R}^2$. The following assumptions are made: (1) the GM counter has a uniform directional response; (2) the radiation sources are point sources; (3) there is

negligible attenuation of gamma radiation due to air; (4) the radiation measurements are independently distributed, and, (5) the exposure time (or the sampling interval) τ is constant for all measurements.

2.2.1. Measurement Model: Known Background Rate

The likelihood function is simply the joint density of the measurement vector $\mathbf{z} \equiv (z_1 \dots z_m)^\top$ conditional on the parameter vector $\boldsymbol{\theta}$ and the knowledge that r sources are present. With this identification, the likelihood function can then be written as a product of Poisson distributions (Martin and Harbison [11]; Tsoufanidis [10]) as follows:

$$p(\mathbf{z} | \mu_b, \boldsymbol{\theta}, I) \equiv p(\mathbf{z} | \mu_b, r, \boldsymbol{\theta}, I) \equiv l(\mathbf{z} | \mu_b, r, \boldsymbol{\theta}) = \prod_{j=1}^m p(z_j; \lambda_j(\boldsymbol{\theta})), \quad (2.9)$$

where $\lambda_j(\boldsymbol{\theta})$ is the mean radiation count for the j th sensor location:

$$\lambda_j(\boldsymbol{\theta}) = \left(\mu_b + \sum_{i=1}^r \frac{Q_{s,i}}{d_{ji}^2} \right) \tau, \quad (2.10)$$

with

$$d_{ji} = \left((x_{l,j} - x_{s,i})^2 + (y_{l,j} - y_{s,i})^2 \right)^{1/2} \quad (2.11)$$

being the Euclidean distance between the i th source and the j th sensor.

The constant μ_b in (2.10) is the mean background count *rate* (namely, the average count *rate* due to the background radiation only which includes the contribution from cosmic and terrestrial radiation). In the measurement model (and concomitant form of the likelihood function given by (2.9) and (2.10)), the mean background count rate μ_b and the number of radiological point sources r are assumed to be known *a priori*. In consequence, μ_b and r have been added as quantities to the right of the vertical bar in the likelihood function of (2.9) to indicate explicitly that these quantities are known exactly.

The mean signal count rate μ_s^j at the j th sensor due to the radioactive point sources only is modelled in (2.10) as

$$\tilde{\mu}_s^j = \sum_{i=1}^r \frac{Q_{s,i}}{(x_{l,j} - x_{s,i})^2 + (y_{l,j} - y_{s,i})^2}. \quad (2.12)$$

In (2.10), it is implicitly assumed that the model for the mean signal count rate in (2.12) is perfect (no model error) so $\mu_s^j = \tilde{\mu}_s^j$.

2.2.2. Measurement Model: Unknown Background Rate

The key assumptions in the formulation of the measurement model of (2.9) and (2.10) are that (1) the mean background count rate μ_b is known *exactly*, and (2) the model error associated

with the model for the mean signal count rate μ_s of (2.12) is *negligible*. Now, we will relax these critical assumptions and approach the (perhaps) more realistic case when we must incorporate both the model error and the effects of an *uncertain* mean background count rate in the measurement model.

Towards this objective, the following model is assumed for the (true though unknown) mean signal count rate μ_s^j ($j = 1, 2, \dots, m$). The mean signal count rate can be described by the radiation data \mathbf{z} and by the quantitative model (cf. (2.12)), so

$$\mu_s^j = \hat{\mu}_s^j + e_j^{(1)} = \tilde{\mu}_s^j + e_j^{(2)}, \quad j = 1, 2, \dots, m, \quad (2.13)$$

where $\hat{\mu}_s^j$ is an estimate for the mean signal count rate at the j th sensor obtained from the data \mathbf{z} (more specifically, z_j), $\tilde{\mu}_s^j$ is the mean signal count rate at the j th sensor obtained from the model given in (2.12), $e_j^{(1)}$ is the uncertainty in the estimate $\hat{\mu}_s^j$ of the mean signal count rate, and $e_j^{(2)}$ represents the model error incurred by using (2.12) to predict the mean signal count rate. The model errors consist of a number of contributions which include (1) the effects of absorption of the gamma radiation by the air and (2) the effects of backscattering from the ground surface which may (possibly) influence the inverse-square law dependence in the model of (2.12).

The measurement model of (2.13) can be rearranged to give

$$\hat{\mu}_s^j - \tilde{\mu}_s^j = e_j^{(2)} - e_j^{(1)} \equiv \epsilon_j, \quad j = 1, 2, \dots, m, \quad (2.14)$$

where ϵ_j is the composite error. If the expectation value $\langle \epsilon_j \rangle$ of the composite error is zero and the variance $\langle \epsilon_j^2 \rangle$ is assumed to be given by σ_j^2 , then application of the principle of maximum entropy (see, Jaynes [12]) provides the following explicit form for the likelihood function (assuming that at least one radiological source is present, so $r > 0$):

$$p(\mathbf{z} | \Theta, I) = \frac{1}{\prod_{j=1}^m \sqrt{2\pi}\sigma_j} \exp \left\{ -\frac{1}{2} \sum_{j=1}^m \left(\frac{\hat{\mu}_s^j - \tilde{\mu}_s^j}{\sigma_j} \right)^2 \right\}. \quad (2.15)$$

It is noted that the Gaussian likelihood (which is effectively the probability distribution of the “noise” ϵ_j) obtained from application of the maximum entropy principle is simply the most conservative noise distribution that is consistent with the given facts (namely, the first two moments of the “noise” are known). In the Gaussian likelihood function of (2.15), the variance of the composite error ϵ_j is obtained as

$$\sigma_j^2 = \sigma_{1,j}^2 + \sigma_{2,j}^2, \quad j = 1, 2, \dots, m, \quad (2.16)$$

where $\sigma_{1,j}^2$ is the variance of the uncertainty $e_j^{(1)}$ in the estimate $\hat{\mu}_s^j$ of the mean signal count rate (obtained from the data \mathbf{z}) and $\sigma_{2,j}^2$ is the variance of the model error $e_j^{(2)}$. In this paper, the model error standard deviation (or square root of the variance) is specified as $\sigma_{2,j} = (\tilde{\mu}_s^j)^{1/2}$.

To complete the specification for the likelihood function in (2.15), we need to provide

the estimate $\hat{\mu}_s^j$ along with the variance $\sigma_{1,j}^2$ in this estimate. To this purpose, we assume that we have available a measurement of background radiation (obtained in the absence of any point radiation sources) given to us as z^b counts obtained over an exposure time τ_b . Now, suppose that a radiation sensor at a fixed receptor location obtains a measurement of z counts in an exposure time τ . We will now show how to use this information to obtain an estimate for the mean signal count rate $\hat{\mu}_s$, as well as to characterize the uncertainty in this estimate in terms of the standard deviation σ_1 .

To accomplish this, we utilize a Bayesian solution for the extraction of weak signals in strong background interference described by Loredó [13]. We will briefly outline the application of this methodology to the problem in this paper. To begin, we can apply Bayes' rule to compute the joint probability for the mean signal count rate μ_s and the mean background count rate μ_b conditioned on the measurement z . This gives

$$p(\mu_s, \mu_b | z, I) \propto p(\mu_s | \mu_b, I) p(\mu_b | I) p(z | \mu_s, \mu_b, I). \quad (2.17)$$

It is stressed that the contextual information I in (2.17) includes the contextual information I_b concerning the measurement of the background radiation and the contextual information I_s concerning the measurement of the radiation field which may include the presence of one or more radiation sources. Hence, the contextual information I that is available in the problem can be decomposed as $I = I_b, I_s$.

We will assign a least informative prior probability for μ_s (the so-called Jeffreys prior; see Jaynes [12]) which has the following form (conditioned on a known mean background rate μ_b):

$$p(\mu_s | \mu_b, I) = \frac{1}{\mu_s + \mu_b}. \quad (2.18)$$

The prior probability for μ_b is chosen to be the following informative prior (using the information that is available, namely, that a measurement of the background radiation yielded z^b counts in an exposure time τ_b):

$$p(\mu_b | I) \equiv p(\mu_b | z^b, I_b, I_s) = p(\mu_b | z^b, I_b) \propto p(\mu_b | I_b) p(z^b | \mu_b, I_b), \quad (2.19)$$

since I_s does not have any bearing on the determination of μ_b . The prior $p(\mu_b | I_b)$ is assigned the nonsinformative Jeffreys prior so $p(\mu_b | I_b) = 1/\mu_b$ and the likelihood function $p(z^b | \mu_b, I_b) = (\mu_b \tau_b)^{z^b} \exp(-\mu_b \tau_b) / z^b!$, so

$$p(\mu_b | I) = p(\mu_b | z^b, I_b) \propto \tau_b (\mu_b \tau_b)^{z^b - 1} \exp(-\mu_b \tau_b). \quad (2.20)$$

The likelihood function in (2.17), associated with the measurement of z counts in exposure interval τ coming from a radiation field with mean count rate $(\mu_s + \mu_b)$, is described by the following Poisson distribution:

$$p(z | \mu_s, \mu_b, I) = \frac{\tau^z (\mu_s + \mu_b)^z}{z!} \exp(-(\mu_s + \mu_b) \tau). \quad (2.21)$$

Now, if we substitute (2.18), (2.20), and (2.21) in (2.17), the joint probability for the mean signal and background count rates is given by

$$p(\mu_s, \mu_b | z, I) \propto (\mu_s + \mu_b)^{z-1} \mu_b^{z\tau-1} \exp(-\mu_s \tau) \exp(-\mu_b(\tau_b + \tau)). \quad (2.22)$$

The marginal probability distribution for the mean signal count rate can be derived by marginalizing with respect to the (unknown) mean background count rate:

$$p(\mu_s | z, I) = \int_0^\infty p(\mu_s, \mu_b | z, I) d\mu_b, \quad (2.23)$$

which on using (2.22) and performing the integration finally gives (on normalization of the probability density function)

$$p(\mu_s | z, I) = \sum_{i=1}^z \Phi_i \frac{\tau(\mu_s \tau)^{i-1}}{(i-1)!} \exp(-\mu_s \tau), \quad (2.24)$$

where

$$\Phi_i = \frac{(1 + \tau_b/\tau)^i (z + z^b - i - 1)! / (z - i)!}{\sum_{j=1}^z (1 + \tau_b/\tau)^j (z + z^b - j - 1)! / (z - j)!}. \quad (2.25)$$

Note that (2.24) and (2.25) provide an explicit expression for the posterior probability of the mean signal count rate μ_s that is independent of the (unknown) mean background count rate μ_b . More specifically, the precise value of μ_b is not known, so it is treated as a nuisance parameter here and removed by integration (marginalization).

Given the posterior probability of μ_s summarized by (2.24) and (2.25), a best estimate for the mean signal count rate can be obtained as the posterior mean:

$$\hat{\mu}_s \equiv \langle \mu_s \rangle = \int_0^\infty \mu_s p(\mu_s | z, I) d\mu_s = \frac{1}{\tau} \sum_{i=1}^z i \Phi_i. \quad (2.26)$$

The second moment of μ_s about zero can be evaluated as

$$\langle \mu_s^2 \rangle = \int_0^\infty \mu_s^2 p(\mu_s | z, I) d\mu_s = \frac{1}{\tau^2} \sum_{i=1}^z i(i+1) \Phi_i, \quad (2.27)$$

from which the uncertainty σ_1 in the estimate for μ_s given in (2.26) can be determined as

$$\sigma_1 = \left(\langle \mu_s^2 \rangle - (\hat{\mu}_s)^2 \right)^{1/2}. \quad (2.28)$$

Equations (2.26), (2.27) and (2.28) when used in conjunction with (2.16) and (2.15) fully determine the form of the likelihood function (assuming implicitly that $r > 0$). This functional

form for the likelihood function accounts explicitly for the model error and the sampling error (the latter of which is compounded by an uncertain mean background count rate).

For the case $r = 0$, the likelihood function needs to be determined as follows (again assuming that our state of knowledge consists of a background radiation measurement yielding z^b counts in exposure time τ_b):

$$p(\mathbf{z} | z^b, I) = \prod_{j=1}^m \int_0^\infty \rho(z_j | \mu_b \tau) p(\mu_b | z^b, I_b) d\mu_b, \quad (2.29)$$

which on insertion of the normalized form of (2.20) for $p(\mu_b | z^b, I_b)$ and evaluation of the integral gives

$$p(\mathbf{z} | z^b, I) = \prod_{j=1}^m \frac{\tau_b^{z_j} \tau^{z_j}}{(\tau + \tau_b)^{z_j + z^b}} \frac{(z_j + z^b - 1)!}{z_j! (z^b - 1)!}. \quad (2.30)$$

In (2.29) and (2.30), the radiation measurements z_j are assumed to be made using the same exposure time τ (without any loss in generality).

Finally, the likelihood function of (2.15) valid for $r > 0$ and the likelihood function of (2.30) valid for $r = 0$ can be combined to give the following general likelihood function for the measurement model assumed in this section:

$$p(\mathbf{z} | z^b, \Theta, I) = \frac{\mathbb{I}_{\{r>0\}}(r)}{\prod_{j=1}^m \sqrt{2\pi} \sigma_j} \exp \left\{ -\frac{1}{2} \sum_{j=1}^m \left(\frac{\hat{\mu}_s^j - \tilde{\mu}_s^j}{\sigma_j} \right)^2 \right\} + \mathbb{I}_{\{r=0\}}(r) \prod_{j=1}^m \frac{\tau_b^{z_j} \tau^{z_j}}{(\tau + \tau_b)^{z_j + z^b}} \frac{(z_j + z^b - 1)!}{z_j! (z^b - 1)!}, \quad (2.31)$$

where the assumption of the availability of a background radiation count z^b (obtained over an exposure time τ_b) has been included in the conditioning of the likelihood function (and of the posterior distribution based upon this likelihood function).

3. Computational Framework

This section describes the computational procedures that were used for extracting the source parameter estimates for source reconstruction. To this purpose, two different methodologies for sampling from two different forms of the posterior distribution for the source parameters are described.

3.1. Importance Sampling Using Progressive Correction

Importance sampling using progressive correction is a computational methodology for sampling from a posterior distribution. In this paper, this sampling procedure is applied to a form of the posterior distribution for the source parameters in which the number of sources

r and the mean background count rate μ_b are assumed to be known *a priori* (see (5.1)). To this purpose, the information contained in the prior distribution for the source parameters $p(\boldsymbol{\theta}(r) | r, I)$ and in the measurements \mathbf{z} assimilated through the likelihood function $p(\mathbf{z} | \mu_b, r, \boldsymbol{\theta}(r), I) \equiv l(\mathbf{z} | \mu_b, r, \boldsymbol{\theta}(r))$ is combined to obtain the posterior distribution for $\boldsymbol{\theta}(r)$ (assuming that the number of sources r and the background count rate μ_b are known):

$$p(\boldsymbol{\theta}(r) | \mu_b, r, \mathbf{z}, I) \propto l(\mathbf{z} | \mu_b, r, \boldsymbol{\theta}(r))p(\boldsymbol{\theta}(r) | r, I). \quad (3.1)$$

The minimum mean-squared error (MMSE) estimate of $\boldsymbol{\theta}$ is then the posterior expectation,

$$\widehat{\boldsymbol{\theta}(r)} \equiv \langle \boldsymbol{\theta}(r) | \mu_b, r, \mathbf{z} \rangle = \int \boldsymbol{\theta}(r) p(\boldsymbol{\theta}(r) | \mu_b, r, \mathbf{z}, I) d\boldsymbol{\theta}(r). \quad (3.2)$$

Because the posterior PDF $p(\boldsymbol{\theta}(r) | \mu_b, r, \mathbf{z}, I)$ and, hence, the posterior expectation cannot be found exactly for the measurement model of Section 2, an approximation of the integral in (3.2) is computed via importance sampling. This involves drawing N samples of the parameter vector from an importance density and approximating the integral by a weighted sum of the samples:

$$\widehat{\boldsymbol{\theta}(r)} \approx \sum_{n=1}^N w^n \boldsymbol{\theta}^{S,n}(r), \quad (3.3)$$

where $\{\boldsymbol{\theta}^{S,n}(r), n = 1, \dots, N\}$ is a ‘‘particle’’ set consisting of samples drawn from the importance density and $\{w^n, n = 1, \dots, N\}$ is the set of importance weights. The sum of these weights is equal to one (namely, $\sum_{n=1}^N w^n = 1$).

Because the prior distribution will often be more diffuse than the likelihood, many samples drawn from the prior may fall outside the region of parameter space favoured by the likelihood function. Therefore, a straightforward application of (3.1) may result in poor source parameter estimates. The progressive correction approach overcomes this problem by using a multistage procedure in which samples are obtained from a series of posterior distributions which become progressively closer to the true posterior distribution. This is achieved by adopting an approximate likelihood at each stage which is somewhat more diffuse than the true likelihood.

Let S denote the number of stages and $p_s(\boldsymbol{\theta}(r) | \mu_b, r, \mathbf{z}, I), s = 1, \dots, S$ denote the posterior PDF at stage s . A series of $p_s(\boldsymbol{\theta}(r) | \mu_b, r, \mathbf{z}, I)$ can be constructed by setting

$$p_s(\boldsymbol{\theta}(r) | \mu_b, r, \mathbf{z}, I) \propto l_s(\mathbf{z} | \mu_b, r, \boldsymbol{\theta}(r))p(\boldsymbol{\theta}(r) | r, I), \quad (3.4)$$

where the intermediate likelihood at stage $s = 1, \dots, S$ is $l_s(\mathbf{z} | \mu_b, r, \boldsymbol{\theta}(r)) = l(\mathbf{z} | \mu_b, r, \boldsymbol{\theta}(r))^{G_s}$ with $G_s = \sum_{\ell=1}^s \gamma_\ell$, such that $\gamma_s \in [0, 1)$, for $s = 1, \dots, S$ and $G_S = \alpha \leq 1$. Assume that a random sample $\{w^n, \boldsymbol{\theta}^{s-1,n}(r)\}$ from $p_{s-1}(\boldsymbol{\theta}(r) | \mu_b, r, \mathbf{z}, I)$ is available, and it is desired to produce a sample from $p_s(\boldsymbol{\theta}(r) | \mu_b, r, \mathbf{z}, I)$. The steps of the PC algorithm are given in Algorithm 1. Note that for $s = 1$, the sample $\{w^n, \boldsymbol{\theta}^{s-1,n}(r)\}$ is drawn directly from the prior $p(\boldsymbol{\theta}(r) | r, I)$ with $w^n = 1/N$.

The performance of the procedure depends somewhat on the number of steps S and the expansion factors $\gamma_1, \dots, \gamma_S$. The expansion factors γ_s were selected in line 8 of Algorithm 1

```

(1)  $s = 0, G_0 = 0, \mu_b = \mu_b^*, r = r_*$ ;
(2) for  $n = 1, \dots, N$  do
(3)   Draw  $\theta^{0,n}(r) \sim p(\theta(r) | r, I)$ ;
(4)   Compute  $l(\mathbf{z} | \mu_b, r, \theta^{0,n}(r))$ ;
(5) end for
(6) while  $G_s < \alpha$  and  $s < S$  do
(7)    $s \leftarrow s + 1$ ;
(8)   Select  $\gamma_s$ ;
(9)    $G_s = G_{s-1} + \gamma_s$ ;
(10)  for  $n = 1, \dots, N$  do
(11)   Weights:  $w^{s,n} = l(\mathbf{z} | \mu_b, r, \theta^{s-1,n}(r))^{\gamma_s} / \sum_{j=1}^N l(\mathbf{z} | \mu_b, r, \theta^{s-1,j}(r))^{\gamma_s}$ ;
(12)  end for
(13)  Resample  $\{\theta^{s-1,n}(r)\}_{n=1}^N$  with weights  $\{w^{s,n}\}_{n=1}^N$  to give  $\{\theta^{s,n}\}_{n=1}^N$  with uniform weights
(14)  for  $n = 1, \dots, N$  do
(15)    $\theta^{s,n}(r) = \theta^{s-1,n}(r) + \epsilon^{s,n}$ , where  $\epsilon \sim g_s$ ;
(16)  end for
(17) end while

```

Algorithm 1: Importance sampling with progressive correction algorithm.

using an adaptive scheme proposed in Musso et al. [14]. In this adaptive scheme, the expansion factors are selected after each step rather than being selected *a priori*. Line 13 in Algorithm 1 performs a resampling on the samples $\{w^{s,n}, \theta^{s-1,n}(r)\}_{n=1}^N$ to give samples $\{1/N, \theta^{s,n}\}_{n=1}^N$ that have uniform weights. This resampling is undertaken using the systematic resampling algorithm (see, Kitagawa [15]), which is summarized in the form of a pseudocode given in Table 3.2 of Ristic et al. [16]. Line 15 in Algorithm 1 performs regularisation (jittering) of samples in order to avoid their duplication. Here, g_s is a Gaussian regularisation kernel (see, Musso et al. [14]).

The case $\alpha = 1$ in the progressive correction algorithm should be used only when the likelihood function model in (2.9) is perfectly correct and its parameters precisely known. As we discussed in Section 2.2.2, this is unrealistic in practice: there are many potential errors in the model, such as the background radiation level, air attenuation, sensor location, and propagation loss. In order to make the PC algorithm robust against the modelling errors, it is recommended to adopt $\alpha < 1$ (this change affects the “while” loop in line 6 of Algorithm 1). In this way, the measurement likelihood is effectively approximated by a fuzzy membership function, but the Bayesian estimation framework holds as the weights in Step 9 of Algorithm 1 are normalised. This is similar to the treatment of the ambiguously generated unambiguous measurements in ([17], Chapter 7). Parameter α is a tuning parameter, which needs to be determined by field trials and calibration.

The PC estimate of (3.2) is approximated as (after the resampling step of line 13 and the regularization step of line 15 in Algorithm 1)

$$\widehat{\theta}(r) = \frac{1}{N} \sum_{n=1}^N \theta^{s,n}(r). \quad (3.5)$$

3.1.1. Source Number Estimation

The importance sampling algorithm using progressive correction described above assumes that the number of sources is known *a priori* (see line 1 of Algorithm 1) which initializes

$r = r_*$, where r_* is the number of sources. However, the number of sources will not be known in advance in practice, requiring this number also to be estimated from the data.

For this purpose, we use the minimum description length (MDL) criterion which chooses the value of r that minimises (Kay [18] page 225); Rissanen [19]):

$$\chi_r = -\log l(\mathbf{z} \mid \mu_b, r, \hat{\boldsymbol{\theta}}_{\text{ML}}(r)) + \frac{n_r}{2} \log m, \quad (3.6)$$

where $\hat{\boldsymbol{\theta}}_{\text{ML}}(r)$ is the maximum likelihood estimate of $\boldsymbol{\theta}(r)$ assuming r sources are present and $n_r \equiv 3r$ is the dimension of $\boldsymbol{\theta}(r)$ under the hypothesis $r \in \mathcal{M}$. Here, $\mathcal{M} = \{0, \dots, r_{\text{max}}\}$ denotes a set of candidate source numbers up to some maximum r_{max} . While MDL is derived based on the maximum likelihood estimate $\hat{\boldsymbol{\theta}}_{\text{ML}}(r)$ (see (3.6)), we will use the MDL in conjunction with the PC by simply replacing $\hat{\boldsymbol{\theta}}_{\text{ML}}(r)$ with the PC estimate $\widehat{\boldsymbol{\theta}}(r)$ in (3.6), so

$$\chi_r = -\log l(\mathbf{z} \mid \mu_b, r, \widehat{\boldsymbol{\theta}}(r)) + \frac{n_r}{2} \log m. \quad (3.7)$$

Finally, the estimate of the number of sources r is then determined as follows:

$$\hat{r} = \arg \min_{r \in \mathcal{M}} \chi_r. \quad (3.8)$$

To estimate the number of sources using the MDL criterion, we first run the importance sampling using progressive correction for all values of $r = 1, \dots, r_{\text{max}}$. For each of these values of r , the estimated value for $\widehat{\boldsymbol{\theta}}(r)$ in (3.5) is used to compute the corresponding value for χ_r from (3.7). The case $r = 0$ (namely, the hypothesis that there are no sources present) is treated separately. There is no need to run the importance sampling using PC in this case, as (3.7) reduces to

$$\chi_0 = -\log l(\mathbf{z} \mid \mu_b) \quad (3.9)$$

when $r = 0$.

3.2. Reversible-Jump Markov Chain Monte Carlo Algorithm

In Section 3.1, we described an importance sampling algorithm for sampling from the posterior distribution based on a sequence of proposal densities (involving, as such, a sequence of modified forms of a likelihood function that has been raised to powers $\gamma_j < 1$, $j = 1, 2, \dots, S$). In its application here, calculations with the importance sampling algorithm using progressive correction were made separately for each model structure (namely, for each value of r), and comparisons of the different models were realized using the MDL criterion. In this section, we describe an alternative for sampling from the posterior distribution that involves application of an MCMC algorithm that allows sampling over both model and parameter spaces simultaneously (namely, r which indexes the various model structures is treated simply as another parameter, so that it is estimated jointly with the other model parameters $\boldsymbol{\theta}(r)$).

Towards this objective, we apply a reversible-jump MCMC (RJMCMC) algorithm. The formalization of RJMCMC algorithms for dealing with variable dimension models has been

described in the seminal work of Green [20]. The application of the RJMCMC algorithm to the problem of inverse dispersion has been described previously by Yee [21, 22]. The algorithm involves constructing a Markov chain whose stationary distribution is the posterior distribution $p(\Theta | z^b, \mathbf{z}, I)$ of the source parameters $\Theta \equiv (r \ \boldsymbol{\theta}(r)^\top)^\top \in \mathbb{R}^{1+3r}$. To this purpose, let $\{\Theta^{(t)}\} \equiv \{(r^{(t)} \ \boldsymbol{\theta}(r^{(t)})^\top)^\top\}$ ($t = 0, 1, 2, \dots$) be the state vector of a Markov chain that is so constructed that its stationary distribution coincides with $p(\Theta | z^b, \mathbf{z}, I)$ (see (5.2) for the explicit form of the posterior distribution that is sampled from using the RJMCMC algorithm in this paper). The construction of the Markov chain uses an RJMCMC algorithm in which the MCMC moves over the model and parameter spaces are separated into two categories: (1) propagation moves which do not change the dimensionality of the parameter space; and, (2) transdimensional jump moves which change the source distribution by ± 1 radiological point source and, in so doing, changes the parameter space dimension by ± 3 .

In the first category of moves (which are dimension-conserving), the parameter vector $\boldsymbol{\theta}(r)$ (r fixed) is partitioned as follows: $\boldsymbol{\theta}(r) = (\boldsymbol{\theta}^{1\top} \ \boldsymbol{\theta}^{2\top})^\top$ where $\boldsymbol{\theta}^{1\top} \equiv (Q_{s,1} \ Q_{s,2} \ \dots \ Q_{s,r}) \in \mathbb{R}^r$ and $\boldsymbol{\theta}^{2\top} \equiv (x_{s,1}^\top \ x_{s,2}^\top \ \dots \ x_{s,r}^\top) \in \mathbb{R}^{2r}$. The parameters collected together in $\boldsymbol{\theta}^1$ are linearly related to the radiation “measurements” (cf. (2.12)). For these parameters, we apply a Gibbs sampler for the update (propagation) move involving drawing samples directly from the univariate full conditional distribution for $Q_{s,i}$ (which happens to be a truncated Bernoulli-Gaussian distribution in this case). The remaining parameters in $\boldsymbol{\theta}^2$ are related nonlinearly to the radiological measurements (cf. (2.12)). Owing to the fact that the univariate conditional posterior distribution for these parameters cannot be determined analytically, we apply a Metropolis-Hastings (M-H) sampler to update the parameters in $\boldsymbol{\theta}^2$. Towards this purpose, we update $\theta_k^2 \in \boldsymbol{\theta}^2$ ($k = 1, 2, \dots, 2r$) by generating a new value $\theta_k^{2'}$ from a proposal transition kernel that is chosen here to be a Gaussian mixture, with each component of the mixture having a mean θ_k^2 and different variances β_k^2 . The variances of the components in this Gaussian mixture are chosen to cover several orders of magnitude from 0.001 to 0.1 times the “length” of the domain of definition for the parameter (assigned using the prior distribution for the parameter).

In the second category of moves (which are dimension-changing), the source distribution is modified by ± 1 radiological point source by (1) a birth move that adds a single radiological point source and (2) a death move that removes a single radiological point source to/from the current source distribution. If a birth move is selected, the “coordinates” of the new radiological point source are generated by drawing a random sample from the prior distribution for each coordinate. Alternatively, if a death (reverse) move is selected, then a radiological point source in the current source distribution is randomly selected and removed. Following the recommendation of Green [20], the probabilities for the birth (p_b) and death (p_d) moves are specified as follows:

$$\begin{aligned} p_b(r) &= \frac{1}{2} \min \left\{ 1, \frac{p(r+1 | I)}{p(r | I)} \right\}, \quad r = r_{\min}, \dots, r_{\max} - 1, \\ p_d(r+1) &= \frac{1}{2} \min \left\{ 1, \frac{p(r | I)}{p(r+1 | I)} \right\}, \quad r = r_{\min}, \dots, r_{\max} - 1, \end{aligned} \quad (3.10)$$

ensuring that the probability of a jump move (either birth or death) $p_b + p_d$ lies in $[0.5, 1]$ at each iteration. We note that for $r = r_{\min}$, $p_d(r) = 0$ and for $r = r_{\max}$, $p_b(r) = 0$.

In practical terms, the RJMCMC algorithm as applied to the problem of source reconstruction can be set up as follows.

- (1) Specify values for the hyperparameters $(r_{\min}, r_{\max}, Q_{\max}, \gamma, p^*)$ and \mathcal{A} which define the prior distribution $p(\Theta | I)$, as well as a value for t_{upper} (maximum number of iteration steps to take).
- (2) Set the iteration counter $t = 1$ and choose an initial state $\Theta^{(t-1)}$ for the Markov chain by sampling from $p(\Theta | I)$.
- (3) Starting from $\Theta^{(t-1)}$, conduct the following sequence of moves to update the state vector to $\Theta^{(t)}$:

$$\Theta^{(t-1)} \xrightarrow{\mathcal{M}_{b,d}} \Theta_{\star} \xrightarrow{\mathcal{M}_1} \Theta_{\star\star} \xrightarrow{\mathcal{M}_2} \Theta^{(t)}, \quad (3.11)$$

where Θ_{\star} and $\Theta_{\star\star}$ denote some intermediate transition states between iterations $(t - 1)$ and t .

- (4) Change the counter from t to $t + 1$, and return to step 3 until a maximum number of steps (t_{upper}) has been taken.

In Step 3, $\mathcal{M}_{b,d}$ denotes a jump move involving either the birth (addition) of a radiological point source at a random locations or the death (removal) of an existing radiological point source with probabilities p_b and p_d , respectively; \mathcal{M}_1 denotes the update of the source intensity rates using a Gibbs sampler; and, \mathcal{M}_2 denotes the update of the source locations using an M-H sampler.

3.2.1. Simulated Annealing

To improve convergence of the Markov chain, we have implemented a simulated annealing procedure. This procedure is very similar to the progressive correction that is used in conjunction with the importance sampling (as described in Section 3.1). In a nutshell, simulated annealing mimics the quasistationary evolution of a system through a sequence of equilibrium samplings, involving an infinitely slow switching between the initial and final states. More concretely, let us consider an ensemble of N_{\star} different source distributions (system) that has been randomly drawn from a modified posterior having the following form: $p_{\lambda}(\Theta | z^b, \mathbf{z}, I) \propto p(\Theta | I)p^{\lambda}(\mathbf{z} | z^b, \Theta, I)$, where $\lambda \in [0, 1]$ is a ‘‘coolness’’ parameter. These samples will be labelled $\Theta_k(\lambda)$ ($k = 1, 2, \dots, N_{\star}$). When $\lambda = 0$ (initial state), the likelihood function is switched off and the modified posterior distribution reduces exactly to the prior distribution. On the other hand, when $\lambda = 1$ (final state) the modified posterior distribution is exactly the posterior that we wish to sample from. In between these two extremes, with $\lambda \in (0, 1)$, the effects of the radiation measurements \mathbf{z} are introduced gradually through the modified (or ‘‘softened’’) likelihood function $p^{\lambda}(\mathbf{z} | z^b, \Theta, I)$.

The parameter λ can be interpreted as an inverse temperature parameter T ($\lambda \equiv 1/T$), so $\lambda \in [0, 1]$ implies that $T \in [1, \infty]$. The modified posterior $p_{\lambda}(\Theta | z^b, \mathbf{z}, I)$ ($\lambda \in [0, 1]$) corresponds to ‘‘heating’’ the posterior distribution to a temperature $T = 1/\lambda > 1$. This heated distribution is flattened relative to the posterior distribution $p(\Theta | z^b, \mathbf{z}, I)$, with the result that moves required to adequately cover the parameter space become more likely. In other words, this facilitates Markov chain mobility in the remote regions of the posterior distribution, ensuring a more reliable relaxation of the chain into the potentially (exponentially) small regions of the hypothesis space where the posterior distribution is large.

We initialize the stochastic sampling scheme at $\lambda = 0$ (infinite temperature) by randomly drawing N_{\star} samples of source distributions $\Theta_k(0)$ ($k = 1, 2, \dots, N_{\star}$) from the prior

distribution $p_0(\Theta | z^b, \mathbf{z}, I)$. Given an ensemble of N_* samples $\Theta_k(\lambda)$ that has achieved equilibrium (at temperature $T = 1/\lambda$) with respect to the modified posterior $p_\lambda(\Theta | z^b, \mathbf{z}, I)$, an ensemble of N_* samples $\Theta_k(\lambda + \delta\lambda)$ that is consistent with $p_{\lambda+\delta\lambda}(\Theta | z^b, \mathbf{z}, I)$ (at the reduced temperature $T = 1/(\lambda + \delta\lambda)$, $\delta\lambda > 0$) can be obtained by using the weighted resampling method (see, Gamerman and Lopes [23]) applied to $\Theta_k(\lambda)$ ($k = 1, 2, \dots, N_*$). To this purpose, each sample $\Theta_k(\lambda)$ is assigned a *normalized* importance weight w_k as follows:

$$\begin{aligned} w_k &= \frac{p_{\lambda+\delta\lambda}(\Theta_k(\lambda) | z^b, \mathbf{z}, I)}{p_\lambda(\Theta_k(\lambda) | z^b, \mathbf{z}, I)} \times \left(\sum_{j=1}^{N_*} w_j \right)^{-1} \\ &= p^{\delta\lambda}(\mathbf{z} | z^b, \Theta_k(\lambda), I) \times \left(\sum_{j=1}^{N_*} w_j \right)^{-1}, \end{aligned} \quad (3.12)$$

for $k = 1, 2, \dots, N_*$. We note that this procedure of computing weights is analogous to line 11 in Algorithm 1 for the importance sampling algorithm using progressive correction.

Next, a resample is drawn from the discrete distribution concentrated at the samples $\{\Theta_k(\lambda)\}_{k=1}^{N_*}$ with weights $\{w_k\}_{k=1}^{N_*}$. This resampling involves mapping the random discrete measure defined by $\{w_k, \Theta_k(\lambda)\}_{k=1}^{N_*}$ into the new random discrete measure defined by $\{1/N_*, \Theta_{k^*}\}_{k^*=1}^{N_*}$ having uniform weights (namely, $\Pr\{\Theta_{k^*} = \Theta_k(\lambda)\} = w_k$ with the result that $\Theta_{k^*} \sim p_{\lambda+\delta\lambda}(\Theta | z^b, \mathbf{z}, I)$ where \sim denotes “is distributed as”). The resampling conducted here is analogous to line 13 in Algorithm 1 for the importance sampling algorithm using progressive correction. The resampling used to produce the weighted replica ensemble acts as a *selection mechanism* in the sense that the number of replicas (copies) of a particular sample $\Theta_k(\lambda)$ is increased for those samples that are assigned a large weight w_k (which, in turn, enhances the number of *relevant* configurations in the ensemble). Note that the weight determined according to (3.12) provides a measure as to whether the state $\Theta_k(\lambda)$ contains new information about areas of the “energy” landscape that are favourable (reflected in larger values of the likelihood function).

There are many ways to implement this resampling (with replacement) from $\{\Theta_k(\lambda)\}_{k=1}^{N_*}$, but in this paper we use the systematic resampling scheme described by Kitaga-wa [15] (namely, the same resampling scheme that is used in the PC algorithm described in Section 3.1). This scheme requires $O(N_*)$ time to execute and minimizes the Monte Carlo variation in the resample. For completeness, we briefly describe this resampling scheme as applied to the current problem. A vector $(n_1, n_2, \dots, n_{N_*})$ of copies of the samples $\Theta_k(\lambda)$ ($k = 1, 2, \dots, N_*$) is obtained by computing a vector $(\rho_1, \rho_2, \dots, \rho_{N_*})$ of the cumulative sums of $N_* \times (w_1, w_2, \dots, w_{N_*})$, generating a random draw from the uniform distribution $u \sim \mathcal{U}([0, 1])$, and determining n_k ($k = 1, 2, \dots, N_*$) from

$$\begin{aligned} n_k &= \lfloor \rho_k + u \rfloor - \lfloor \rho_{k-1} + u \rfloor, \quad k = 2, 3, \dots, N_* - 1, \\ n_1 &= \lfloor \rho_1 + u \rfloor, \quad n_{N_*} = N_* - \lfloor \rho_{N_*-1} + u \rfloor, \end{aligned} \quad (3.13)$$

where $\lfloor \cdot \rfloor$ denotes the “integer part of”. After resampling, the weights for each member of the resample are reset to $w_k = 1/N_*$ (namely, an equal weight is assigned to each member of the resample). After each resampling step to give an ensemble of N_* samples $\Theta_k(\lambda + \delta\lambda)$ that is in equilibrium at temperature $T = 1/(\lambda + \delta\lambda)$ (approximately or better) with respect to

$p_{\lambda+\delta\lambda}(\Theta | z^b, \mathbf{z}, I)$, we apply N_r (typically 10) Markov chain transitions $T_{\lambda+\delta\lambda}$ to each member of the ensemble obtained from the resampling operation. These Markov chain transitions utilize the update scheme given by (3.11) and leave $p_{\lambda+\delta\lambda}(\Theta | z^b, \mathbf{z}, I)$ invariant.

It is necessary to specify the sequence of $\delta\lambda$'s used to move the ensemble of samples $\{\Theta_k(\lambda)\}_{k=1}^{N_*}$ from $\lambda = 0$ to $\lambda = 1$. This sequence defines the annealing schedule for $\lambda \in [0, 1]$. Taking a hint from the physical process of annealing, it is useful to allow the system to cool slowly. This allows the system to transition through a sequence of quasiequilibrium states and, in so doing, minimizes the probability of the system being trapped into a metastable state along the path of annealing. To mimic this process for the case of simulated annealing, at each step of the annealing schedule (starting from $\lambda = 0$ where the stochastic sampling scheme is exploring the prior) the step size $\delta\lambda$ is chosen so that

$$\frac{1}{N_*} \sum_{k=1}^{N_*} |w_k - 1| \approx R, \quad (3.14)$$

where w_k are the weights computed in accordance to (3.12) and R is a small constant ($R \ll 1$). The constant R is chosen to provide a limiter on the maximum permissible size of $\delta\lambda$ in each step of the annealing schedule. Note that unlike the PC algorithm, the sequence of $\lambda_i \in [0, 1]$ ($i = 0, 1, 2, \dots, K$) with $\lambda_i < \lambda_j$ for $i < j$ and $\lambda_0 = 0$, $\lambda_K = 1$ that define the annealing schedule need not sum to α (namely, $\sum_{i=0}^K \lambda_i \neq \alpha$ in general).

When $\lambda = 1$, the annealing phase is complete (corresponding, as such, to the burn-in phase of the algorithm) and probabilistic exploration of the hypothesis space proceeds using the RJMCMC algorithm described at the end of Section 3.2. In the application of this algorithm following simulated annealing, Step 2 is modified as follows: instead of sampling $\Theta^{(0)}$ from the prior distribution $p(\Theta | I)$, these initial states are simply chosen to coincide with the samples from the ensemble of source distributions $\{\Theta_k(\lambda = 1)\}_{k=1}^{N_*}$ obtained at the end of the annealing phase.

4. Experimental Trial

A radiological field trial was conducted on a large, flat, and open area without obstacles within a military site in Puckapunyal, Victoria, Australia. An AN/PDR-77 radiation survey meter equipped with a gamma probe containing two Geiger-Müller tubes to cover both low and high ranges of dose rates was used to collect the radiation data. When connected to the AN/PDR-77 radiation survey meter, this gamma probe is capable of measuring gamma radiation dose rates from background to 9.99 Sv h^{-1} without saturating [24], and it has a fairly flat response ($\sim \pm 10\%$) from 0.1 to above 1 MeV (see [24]; page 7). Measured dose rate data were recorded in $\mu\text{Sv h}^{-1}$. We converted these dose rate data into raw count measurements $z_j \in \mathbb{N} \cup \{0\}$ by multiplication by a conversion factor (which was 0.21^{-1} for our probe) to convert to count rate data and subsequently by multiplication by the exposure time to convert to raw count data (measured over the exposure time).

Three radiation sources were used in the field trial: two cesium sources (^{137}Cs) and one cobalt (^{60}Co) source (see Table 1). While the strengths of the radiation sources are typically characterised in terms of their activity in GBq (cf. Table 1), for simplicity we will characterise the strength (intensity rate) of a radiation point source by the expected dose rate at a distance of one meter from the source. The strengths of our three radiation sources calibrated in this

Table 1: Radiation sources used in the field trial.

Source	Type	Activity (MBq)
1	^{137}Cs	26×10^3
2	^{137}Cs	5×10^3
3	^{60}Co	0.2×10^3

Table 2: The locations of radiation sources in the field trial.

Test set	Source 1 Location	Source 2 Location	Source 3 Location
1	(11, 10) m	—	—
2	(11, 10) m	(3, 50) m	—
3	(11, 10) m	(3, 50) m	(41, 5) m

manner were $Q_s = 1912, 392.28, \text{ and } 98.07 \mu\text{Sv m}^2\text{h}^{-1}$, for Sources 1, 2, and 3, respectively. We will use these values as the “true” source strengths in our source estimations.

The sources were mounted at the same height above the ground as the gamma probe. To ensure that radiation sources appear as isotropic in the horizontal plane, they were placed in a vertical configuration so that the handling rods were pointing up. Radiation dose measurements were collected on a grid of surveyed points that were carefully measured and marked beforehand in the local Cartesian coordinate system on the asphalt surface of the airfield. The data were acquired when the trolley-mounted gamma probe was positioned over individual grid points. During data collection at each grid point, the gamma probe remained stationary until sixty measurements were acquired. The exposure time for each radiation dose measurement (effectively the sampling interval) was kept constant at about $\Delta = 0.8 \text{ s}$. Multiple measurements were collected at each location so that the candidate algorithms could be tested using several different batches of experimental data, or different total exposure intervals can be used in the source reconstruction by choosing the appropriate number of data points at each grid point (e.g., to test the source reconstruction algorithms for raw counts obtained over an exposure interval τ , we can simply select n_* data points at each grid position such that $\tau \approx n_* \Delta$).

Data sets were collected with one, two, and three radiation sources emplaced, respectively. These data sets are referred to as test sets 1, 2 and 3. The sources used when collecting these data sets and their locations in the local Cartesian coordinate system are listed in Table 2.

An areal picture of the Puckapunyal airfield site where the field trial was conducted is shown in Figure 1. The three red stars show the locations where the three radiation sources were emplaced. The white cross symbols indicate the grid points where data were collected. The x - and y -axis markings on the plot show the distances along these directions in meters (as well as the local Cartesian coordinate system that is used here to refer to locations of sources and grid positions).

In order to estimate the background radiation level in the field trial, measurements were collected at a number of grid positions in the *absence* of radiation sources. At each of these grid positions, the detector was held stationary until 120 measurements were acquired. The background radiation counts were found to be spatially homogeneous and verify a Poisson distribution with mean background count rate of $\mu_b \approx 0.9 \text{ counts s}^{-1}$.

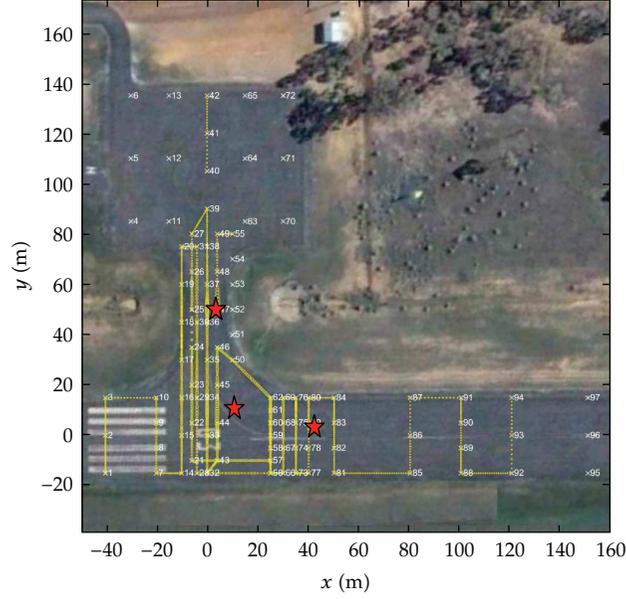


Figure 1: Aerial image of the Puckapunyal airfield site where the field trial was conducted. The red stars located at coordinates (11, 10) m, (3, 50) m, and (41, 5) m of the local Cartesian coordinate system indicate where the three sources were emplaced for Test 3. The numbered grid points marked with crosses are the measurement points.

5. Implementational Aspects

5.1. Importance Sampling Algorithm Using PC

The importance sampling algorithm using progressive correction was applied to the experimental field trial data for source reconstruction. Towards this purpose, this algorithm was used to sample from the following posterior distribution of the source parameters:

$$\begin{aligned}
 p(\boldsymbol{\theta} \mid \mu_b, r, \mathbf{z}, I) &\propto p(\boldsymbol{\theta} \mid r, I) p(\mathbf{z} \mid \mu_b, r, \boldsymbol{\theta}, I) \\
 &\propto \prod_{j=1}^m p(z_j; \lambda_j(\boldsymbol{\theta})) \\
 &\quad \times \prod_{i=1}^r Q_{s,i}^{\kappa-1} \frac{\exp(-Q_{s,i}/\psi)}{\Gamma(\kappa)\psi^\kappa} \times \frac{\mathbb{I}_{\mathcal{A}}(\mathbf{x}_{s,i})}{\text{meas}(\mathcal{A})},
 \end{aligned} \tag{5.1}$$

which combines the prior distribution for source strengths given by (2.5) and the prior distribution for source location given by (2.7) with the likelihood function given by (2.9). The quantities $\lambda_j(\boldsymbol{\theta})$ are determined in accordance with (2.10). In this form of the posterior distribution, it is assumed that the mean background count rate μ_b and the number of sources r are known *a priori*.

The importance sampling algorithm using progressive correction is applied to the experimental field trial data using the following values for the hyperparameters that define the prior distribution in (5.1). The gamma distribution that defines the prior for the source strengths has a shape parameter $\kappa = 1.5$ and a scale parameter $\psi = 8000 \mu\text{Sv m}^2 \text{h}^{-1}$. This

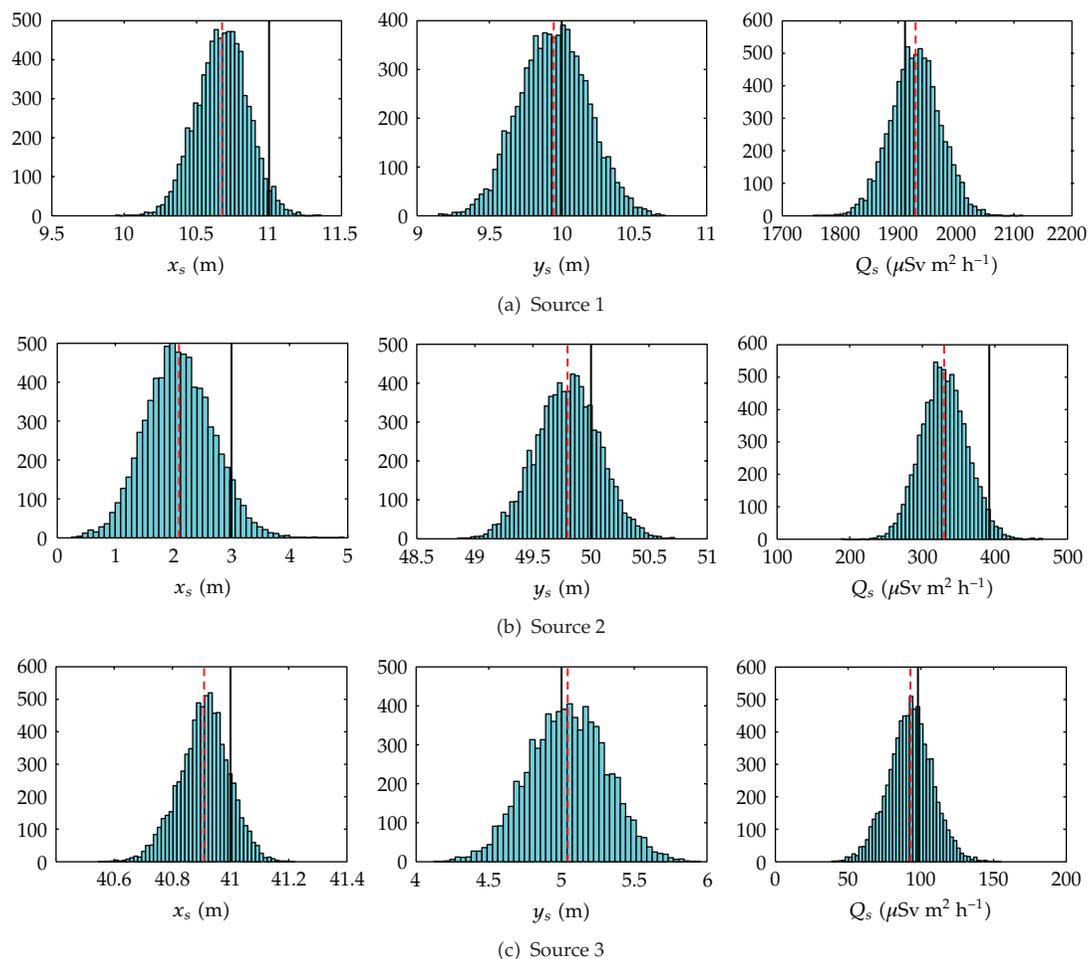


Figure 2: Inference of the parameters of the three radiological point sources obtained from samples drawn from the posterior distribution $p(\boldsymbol{\theta} \mid \mu_b, r, \mathbf{z}, I)$ for $\tau = 16$ s and $r = 3$ using the importance sampling algorithm with PC correction. (a, b, c) Histograms for the three parameters, namely x_s -coordinate of source, y_s -coordinate of source, and source intensity rate Q_s that characterize sources 1, 2, and 3. In each frame, the solid vertical line indicates the true value of the parameter and the dashed vertical line corresponds to the best estimate of the parameter obtained as the posterior mean of the marginal posterior distribution for the parameter.

choice of the hyperparameters for the gamma distribution provides a prior for the source strengths that is broad enough to cover all likely source strength values. The prior for the source location is a uniform distribution over a specified rectangular-shaped area \mathcal{A} , typically containing all measurement points. For the current application, the area \mathcal{A} is selected as the following rectangular region: $\mathcal{A} \equiv \{(-50, 200) \otimes (-20, 150)\}$, where \otimes denotes Cartesian product.

The PC algorithm is initialized by drawing a random sample in $3r$ -dimensional parameter space from $p(\boldsymbol{\theta} \mid r, I)$. The number of samples in the PC algorithm is set to $N = 2500r$.

The expansion factors and the actual number of stages of the PC algorithm are data dependent. We have tuned the PC algorithm so that on average the number of stages is directly proportional to the number of sources. Parameter α was set to 1.

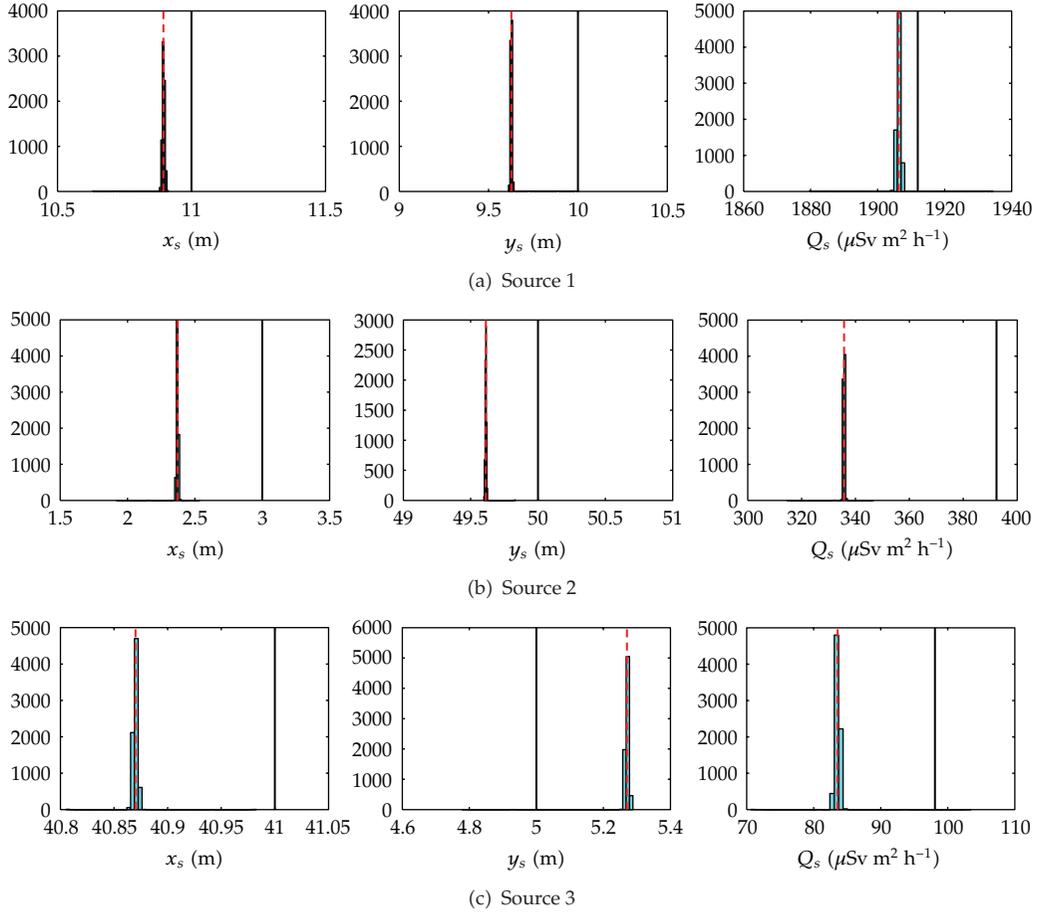


Figure 3: Inference of the parameters of the three radiological point sources obtained from samples drawn from the posterior distribution $p(\theta | \mu_b, r, z, I)$ for $\tau = 48$ s and $r = 3$ using the importance sampling algorithm with PC correction. (a, b, c) Histograms for the three parameters, namely, x_s -coordinate of source, y_s -coordinate of source, and source intensity rate Q_s that characterize sources 1, 2, and 3. In each frame, the solid vertical line indicates the true value of the parameter and the dashed vertical line corresponds to the best estimate of the parameter obtained as the posterior mean of the marginal posterior distribution for the parameter.

5.2. Reversible-Jump MCMC Algorithm

The reversible-jump MCMC algorithm (applied in conjunction with simulated annealing) was used to infer an unknown number of radiological point sources from the experimental field trial data. This involves drawing samples of radiological point source distributions from the following posterior distribution for the source parameters Θ :

$$\begin{aligned}
 p(\Theta | z^b, z, I) &\propto p(\Theta | I)p(z | z^b, \Theta, I) \\
 &\propto \frac{\mathbb{I}_{\{r>0\}}(r)}{\prod_{j=1}^m \sqrt{2\pi}\sigma_j} \exp\left\{-\frac{1}{2}\sum_{j=1}^m \left(\frac{\hat{\mu}_s^j - \tilde{\mu}_s^j}{\sigma_j}\right)^2\right\}
 \end{aligned}$$

$$\begin{aligned}
& \times \frac{(r_{\max} - r_{\min})!}{(r - r_{\min})!(r_{\max} - r)!} p^{*(r-r_{\min})} (1 - p^*)^{r_{\max}-r} \\
& \times \prod_{i=1}^r (1 - \gamma) \delta(Q_{s,i}) + \frac{\gamma \mathbb{I}_{(0, Q_{\max})}(Q_{s,i})}{Q_{\max}} \\
& \times \frac{\mathbb{I}_{\mathcal{A}}(\mathbf{x}_{s,i})}{\text{meas}(\mathcal{A})} \\
& + \mathbb{I}_{\{r=0\}}(r) \prod_{j=1}^m \frac{\tau_b^{z_j} \tau^{z_j}}{(\tau + \tau_b)^{z_j + z^b}} \frac{(z_j + z^b - 1)!}{z_j! (z^b - 1)!} \\
& \times (1 - p^*)^{r_{\max}},
\end{aligned} \tag{5.2}$$

which combines the likelihood function given by (2.31) with the prior for the number of sources given by (2.4), the prior for the source strength given by (2.6), and the prior for the source location given by (2.7). For the source reconstruction, the hyperparameters that define the prior distribution in (5.2) are assigned the following values: $r_{\min} = 0$, $r_{\max} = 4$, $p^* = 0.25$, $\gamma = 0.25$, $Q_{\max} = 75,000 \mu\text{Sv m}^2 \text{h}^{-1}$, and $\mathcal{A} = \{(-50, 200) \otimes (-20, 150)\}$ which is used to define the prior bounds for the location (x_s, y_s) of any radiological point source. Note that the prior bounds for the source location are exactly those assigned for the application of importance sampling using PC.

The stochastic sampling algorithm was executed with $N_* = 50$ members of an ensemble of distributions of radiological point sources (originally, randomly drawn from the prior distribution $p(\Theta | I)$). During the burn-in (or annealing) phase of the algorithm, the control strategy used to determine the annealing schedule (see (3.14)) was applied with $R = 0.03$. After termination of the annealing at $\lambda = 1$, a further $t_{\text{upper}} = 1000$ iterations of the RJMCMC procedure were applied to each of the $N_* = 50$ members of the ensemble of distributions of radiological point sources obtained at the end of the annealing phase. This gave 50,000 samples of distributions of radiological point sources drawn from the posterior distribution $p(\Theta | z^b, \mathbf{z}, I)$ given in (5.2) (corresponding, as such, to the probabilistic exploration phase of the algorithm).

The samples of radiological point source distributions drawn from the posterior distribution of (5.2) can be used to make various inferences on the values of the source parameters. For example, the number of sources can be obtained from the maximum a posteriori (MAP) estimate as follows:

$$\hat{r} = \arg \max_r p(r | z^b, \mathbf{z}, I), \tag{5.3}$$

where the (marginal) posterior probability for the number of sources is estimated from the samples as follows:

$$\hat{p}(r | z^b, \mathbf{z}, I) = \frac{1}{N'} \# \{t : r^{(t)} = r\} = \frac{1}{N'} \sum_{t=1}^{N'} \mathbb{I}_{\{r\}}(r^{(t)}). \tag{5.4}$$

Here, N' is the number of samples $\Theta^{(1)}, \Theta^{(2)}, \dots$ drawn from a sampled realization of the Markov chain. Similarly, using the Markov chain samples drawn from the posterior distribution,

the posterior means of the various source parameters $\theta^i(\hat{r})$ (various components of $\boldsymbol{\theta}(\hat{r})$; see (2.2)) can be estimated by

$$\widehat{\theta^i(\hat{r})} = \frac{\sum_{t=1}^{N'} \theta_{r^{(t)}}^{i(t)} \mathbb{I}_{\{\hat{r}\}}(r^{(t)})}{\sum_{t=1}^{N'} \mathbb{I}_{\{\hat{r}\}}(r^{(t)})}. \quad (5.5)$$

Similar estimates can be constructed for the posterior standard deviation of $\theta^i(\hat{r})$ which can be used as a measure of uncertainty in the recovery of this quantity. Alternatively, a $p\%$ credible (or highest posterior distribution (HPD)) interval that contains the source parameter $\theta^i(\hat{r})$ with $p\%$ probability, with the upper and lower bounds specified such that the probability density within the interval is everywhere larger than outside it, can be used as a measure of the uncertainty in the determination of this quantity.

In addition to these statistical quantities, the RJMCMC algorithm applied in conjunction with simulated annealing permits the normalization constant (or evidence) Z for the posterior distribution of (5.2) to be determined². This can be accomplished using an approach known as thermodynamic integration (see, von der Linden [25]).³ In this approach, Z can be evaluated as follows:

$$\log Z = \int_0^1 \left\langle \log \left[p(\mathbf{z} | z^b, \boldsymbol{\Theta}, I) \right] \right\rangle_{\lambda} d\lambda, \quad (5.6)$$

where $\langle \cdot \rangle$ denotes the expectation operation taken with respect to the modified posterior $p_{\lambda}(\boldsymbol{\Theta} | z^b, \mathbf{z}, I)$. This quantity can be computed during the simulated annealing phase of the stochastic simulation procedure. More specifically, the integral in (5.6) is discretized using the following Riemann sum:

$$\log Z \approx \sum_{i=1}^K \left\langle \log \left[p(\mathbf{z} | z^b, \boldsymbol{\Theta}, I) \right] \right\rangle_{\lambda_i} \delta\lambda_i, \quad (5.7)$$

where $\delta\lambda_i \equiv (\lambda_i - \lambda_{i-1})$ is determined according to (3.14); and, the annealing schedule so defined provides the set of quadrature points ($\lambda_0 = 0, \lambda_1, \dots, \lambda_{K-1}, \lambda_K = 1$).

The evidence Z can be used to determine the relative entropy (or, information gain) corresponding to the assimilation of the radiation measurements \mathbf{z} , which involves updating our state of knowledge about the source embodied in the prior distribution to that embodied in the posterior distribution. To this end, the relative entropy H can be determined as follows (see, Cover and Thomas [26]):

$$\begin{aligned} H &\equiv \int \log \left(\frac{p(\boldsymbol{\Theta} | z^b, \mathbf{z}, I)}{p(\boldsymbol{\Theta} | I)} \right) p(\boldsymbol{\Theta} | z^b, \mathbf{z}, I) d\boldsymbol{\Theta} \\ &= \left\langle \log \left[p(\mathbf{z} | z^b, \boldsymbol{\Theta}, I) \right] \right\rangle - \log(Z). \end{aligned} \quad (5.8)$$

The relative entropy (of the posterior with respect to the prior) is a measure of the information gain provided by the receipt of the radiation measurements \mathbf{z} . Geometrically, this can be visualized as the logarithm of the volumetric factor by which the prior has been compressed

to give the posterior in the hypothesis space (the greater this compression, the greater the information gain embodied in \mathbf{z}).

6. Results

To test the usefulness of the two approaches for source reconstruction, we have applied them to test data sets 1, 2, and 3 corresponding to one-, two-, and three-source examples, respectively. However, owing to space limitations, we will show only the results of application of the two approaches for test case 3. This test is arguably the most difficult case, but the results from the test are representative also for the other two test cases.

The results for test case 3 will be shown for two different applications: namely, the radiation measurements at the grid positions (see Figure 1) were obtained over exposure intervals of $\tau = 16$ and 48 s. In the application of importance sampling using PC, the mean background count rate μ_b was assumed to be exactly known ($\mu_b = 0.9$ counts s^{-1} was used in the algorithm). For the application of RJMCMC used in conjunction with simulated annealing, the first 24 s ($\tau_b = 24$ s) of a measurement of the background counts z^b at receptor location $(x_l, y_l) = (-40, -15)$ m (or, at grid position 1 as shown in Figure 1) was used in the source recovery.

Firstly, we show the results of the two applications of the importance sampling algorithm using PC for test case 3. For each of these two applications, the algorithm (summarized in Algorithm 1) was executed with *fixed* values of $r = r_*$ with $r_* = 1, 2,$ and 3. Essentially, the assumption in the algorithm is that the number of sources is exactly known *a priori*, with $r = r_*$. At the completion of this process, the MDL was computed according to (3.7) for $r \in \{0, 1, 2, 3\}$ and a “best” estimate for the number of sources was found using (3.8). In both applications of the importance sampling algorithm using PC here (namely, for $\tau = 16$ and 48 s), the minimum value of χ_r was found to correspond to $\hat{r} = 3$, an estimate that coincides with the correct number of sources for test case 3.

The results of the source recovery for the two applications of the importance sampling algorithm are summarized in Table 3 for $r = 3$ (assumed correct number of sources). Here, the posterior mean, posterior standard deviation, and lower and upper bounds for the 95% HPD interval of the source parameters (x_s, y_s, Q_s) are shown for $\tau = 16$ s (upper portion of Table 3) and for $\tau = 48$ s (lower portion of Table 3). In addition, histograms of the source parameters (x_s, y_s, Q_s) for the three sources are exhibited in Figures 2 and 3, respectively, for $\tau = 16$ and 48 s.

For the application with $\tau = 16$ s, the algorithm successfully recovered the true parameters for all three sources to within the stated errors (e.g., the true parameters were all contained within a three-standard deviation interval about the best estimates of the parameters, and/or the true parameters were contained within the 95% HPD interval). However, this is not the case for the application of the algorithm with $\tau = 48$ s. More specifically, in this case it is seen that neither the three-standard deviation interval about the best estimates (based on the posterior mean) nor the 95% HPD interval for (x_s, y_s, Q_s) for all three sources enclose the true values for these parameters. The reason for this is that the likelihood function in (5.1) that is used in the importance sampling algorithm with PC does not incorporate model error (namely, the model for the radiation measurements given by (2.12) is assumed to be exact). In consequence, for $\tau = 48$ s, the sampling error is smaller than the model error, so the latter becomes the dominant component in the error specification. The neglect of the model error in this case leads to underestimates for the precision at which some of the source parameters are recovered by the importance sampling algorithm.

Table 3: The posterior mean, posterior standard deviation, and lower and upper bounds of the 95% HPD interval of the parameters $x_{s,i}$ (m), $y_{s,i}$ (m), and $Q_{s,i}$ ($\mu\text{Sv m}^2 \text{h}^{-1}$) for $i = 1, 2,$ and 3 calculated from samples of distributions of radiological point sources drawn from the posterior distribution of $\theta(r)$ ($r = 3$) using importance sampling with progressive correction. The upper and lower parts of the table show the results of the source recovery using radiological measurements with exposure times of $\tau = 16$ and 48 s, respectively.

Parameter	Mean	Standard deviation	95% HPD	Actual
$\tau = 16$ s				
$i = 1$				
x_s	10.7	0.2	(10.3, 11.0)	11.0
y_s	9.9	0.2	(9.5, 10.4)	10.0
Q_s	1930	43	(1848, 2014)	1912
$i = 2$				
x_s	2.1	0.6	(1.0, 3.2)	3.0
y_s	49.8	0.3	(49.3, 50.3)	50.0
Q_s	330	31	(269, 392)	392
$i = 3$				
x_s	40.9	0.1	(40.7, 41.1)	41.0
y_s	5.0	0.3	(4.5, 5.6)	5.0
Q_s	92.9	15.2	(62.3, 122.7)	98.1
$\tau = 48$ s				
$i = 1$				
x_s	10.9	0.01	(10.89, 10.90)	11.0
y_s	9.6	0.01	(9.62, 9.63)	10.0
Q_s	1906	1.3	(1905, 1907)	1912
$i = 2$				
x_s	2.3	0.01	(2.36, 2.38)	3.0
y_s	49.6	0.01	(49.60, 49.62)	50.0
Q_s	336	0.6	(335.3, 336.2)	392
$i = 3$				
x_s	40.8	0.00	(40.87, 40.87)	41.0
y_s	5.3	0.01	(5.26, 5.28)	5.0
Q_s	83.6	0.73	(83.0, 84.2)	98.1

Next, we show the results for source reconstruction for test case 3 using the RJMCMC algorithm (applied in conjunction with simulated annealing) for the two applications with $\tau = 16$ and 48 s. Figure 4(a) displays a trace plot for the number of radiological point sources r in source distribution models sampled from the posterior distribution of Θ (for $\tau = 48$ s). The comparable result for $\tau = 16$ s is very similar (and hence, not shown). The trace plot here corresponds to the first 500 iterations of the RJMCMC sampler obtained during the probabilistic exploration phase (post-burn-in iterations following the completion of the annealing phase). Note that death moves for source distribution models, involving transitions from $r = 3$ to 2 (or from $r = 2$ to 1 and from $r = 1$ to 0) do not occur. Furthermore, birth moves involving transitions from $r = 2$ to 3 do not occur, either. On the other hand, dimension-changing moves (births or deaths) involving transitions from $r = 3$ to 4 (and vice versa) are seen to occur. The residence time for occupation of a state with $r = 3$ is seen to be significantly longer than that for a state with $r = 4$. This is confirmed in Figure 4(b) which

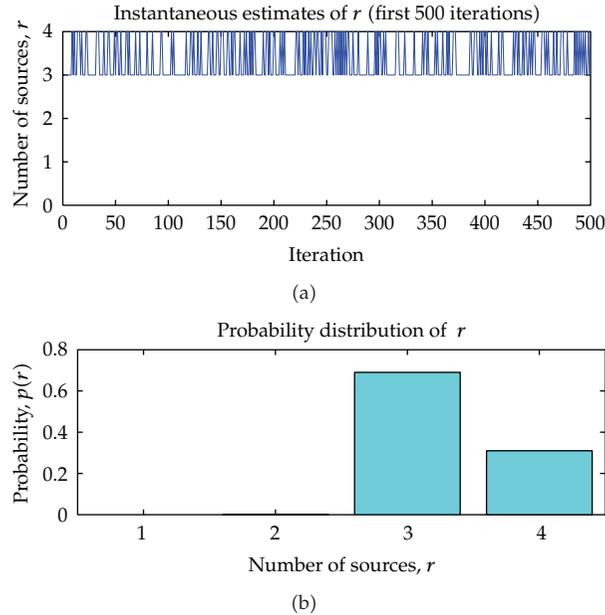


Figure 4: Trace plot (a) of the number of radiological point sources r in the first 500 samples drawn from $p(\Theta | z^b, \mathbf{z}, I)$ during the probabilistic exploration phase using RJMCMC (after the completion of the annealing phase) and the corresponding posterior distribution (b) for the number of sources, $p(r) \equiv p(r | z^b, \mathbf{z}, I)$.

exhibits the probability distribution $p(r) \equiv p(r | z^b, \mathbf{z}, I)$ for the number of sources (estimated as a histogram from the 50,000 post-burn-in samples in accordance with (5.4)).

Figure 4(b) shows that the most probable number of sources (MAP estimate) for this test case is 3 ($\hat{r} = 3$), which corresponds to the correct number of sources for this test. More specifically, $\hat{r} = 3$ is favored with a probability of about 0.7. Interestingly, it is noted that many of the distributions having four radiological point sources (which from Figure 4(b) are favored with probability of about 0.3) have one of the point sources “turned off” (namely, the source strength Q_s for this source is identically zero). In fact, these distributions of radiological point sources should be considered as being identical to $r = 3$ sources (and not $r = 4$ sources owing to the fact that $Q_s = 0$ for one of the sources implies that that source is not present, at least as far as the detection of that source by a radiological sensor is concerned). If this classification for the number of radiological sources is used instead, then $\hat{r} = 3$ is now favored with a probability greater than about 0.995.

Given the fact that our best estimate for the number of sources is $\hat{r} = 3$, we can now estimate the source parameters (x_s, y_s, Q_s) corresponding to each of these three radiological point sources. Towards this purpose, we extracted all samples of distributions having exactly three sources (including all four-source samples, with one of the sources “turned off”). It is noted that an important issue in the interpretation of the results is the identifiability problem, which arises owing to the fact that the posterior distribution of Θ (cf. (5.2)) is invariant under a relabelling of the identifiers used for each of the radiological point sources in the distribution. In consequence, we relabel the radiological point sources so that sources 1, 2, and 3 correspond to the (arbitrary) labelling of sources used in Table 2. After this re-labelling, Table 4 summarizes the posterior mean and standard deviation, as well as the lower and

Table 4: The posterior mean, posterior standard deviation, and lower and upper bounds of the 95% HPD interval of the parameters $x_{s,i}$ (m), $y_{s,i}$ (m), and $Q_{s,i}$ ($\mu\text{Sv m}^2 \text{h}^{-1}$) for $i = 1, 2,$ and 3 calculated from samples of distributions of radiological point sources drawn from the posterior distribution of Θ using the RJMCMC algorithm. The upper and lower parts of the table show the results of the source recovery using radiological measurements with an exposure times of $\tau = 16$ and 48 s, respectively.

Parameter	Mean	Standard deviation	95% HPD	Actual
$\tau = 16$ s				
$i = 1$				
x_s	8.7	0.8	(7.2, 10.6)	11.0
y_s	10.6	0.8	(9.1, 12.1)	10.0
Q_s	1404	152	(1109, 1704)	1912
$i = 2$				
x_s	3.0	0.6	(3.0, 5.0)	3.0
y_s	49.8	0.6	(49.0, 50.9)	50.0
Q_s	385	81	(228, 548)	392
$i = 3$				
x_s	40.9	0.5	(39.9, 41.6)	41.0
y_s	5.0	0.9	(3.5, 6.4)	5.0
Q_s	194	46	(105, 286)	98.1
$\tau = 48$ s				
$i = 1$				
x_s	9.3	0.9	(7.6, 11.0)	11.0
y_s	10.2	0.8	(8.7, 11.9)	10.0
Q_s	1433	156	(1130, 1742)	1912
$i = 2$				
x_s	3.8	0.6	(3.0, 4.9)	3.0
y_s	49.8	0.6	(49.0, 50.9)	50.0
Q_s	382	80	(225, 543)	392
$i = 3$				
x_s	41.1	0.5	(40.0, 41.8)	41.0
y_s	4.8	0.8	(3.5, 6.3)	5.0
Q_s	204	52	(104, 309)	98.1

upper bounds for the 95% HPD interval of the source parameters, for each of the three radiological point sources for radiological measurements with exposure times $\tau = 16$ and 48 s.

From Table 4, it is seen that the estimated values for the locations of the three sources agree well with the true locations of the sources in the reconstruction undertaken for both $\tau = 16$ and 48 s. Comparing these estimated values for the source locations with the actual source locations, we see also that the algorithm has adequately recovered the source locations to within the errors as represented by the 95% HPD interval for $\tau = 16$ and 48 s. Furthermore, the intensity rates of the sources are estimated quite well for both $\tau = 16$ and 48 s, although the accuracy in the recovery of these parameters is not as good as for the source location. Although the 95% HPD interval does not contain the true estimate for Q_s for sources 1 and 3 for both $\tau = 16$ and 48 s, it should be noted that the 99% HPD interval does enclose the true value of Q_s for these sources.

An examination of Table 4 shows that the recovery of the source parameters for $\tau = 16$ and 48 s using the RJMCMC algorithm gives results that are very similar, both in terms of their accuracy and their precision. This is in contrast with the results obtained using the importance

sampling algorithm where it was found that there was an underestimate in the precision of recovery of the source parameters for $\tau = 48$ s. This difference in performance arises because the importance sampling algorithm was applied to the simpler posterior distribution of (5.1) which does not incorporate the effects of model error, whereas the RJMCMC algorithm was applied to the posterior distribution of (5.2) where the effects of model error are explicitly incorporated.

A comparison of Tables 3 and 4 shows that the source parameters are generally recovered with greater accuracy and better precision using the importance sampling algorithm than those recovered using the RJMCMC algorithm. The reason for this is because the importance sampling algorithm uses a posterior distribution (cf. (5.1)) that is much more informative than the posterior distribution (cf. (5.2)) used by the RJMCMC algorithm. In particular, the number of sources is assumed known *a priori* in the former, but is an unknown parameter in the latter; the mean background rate is assumed to be known in the former, but is unknown in the latter (where it is assumed that only a background count measurement z^b taken over an exposure time τ_b is available), and the effects of model error on the source recovery are taken into account in the latter, but not in the former.

Histograms of the source parameters (approximating the marginal probability distributions for the parameters), associated with each of the three identified radiological point sources, are exhibited in Figure 5 for $\tau = 48$ s. Again, the comparable results for $\tau = 16$ s are not shown owing to the fact that they are very similar. A perusal of this figure suggests that the localisation of the three sources in both the x_s - and y_s -directions is generally very good. The recovery of the source intensity rates Q_s is quite good, although both the accuracy and precision in this recovery are poorer than the recovery of the source locations.

Finally, the information gain H (see (5.8)) obtained from the radiation measurements \mathbf{z} for $\tau = 16$ and 48 s (for test case 3) was found to be $H = 53.4$ and 60.1 natural units (nits), respectively.⁴ This implies that the information contained in the radiation measurements with exposure times $\tau = 16$ and 48 s allowed the “posterior volume” of the hypothesis space (volume of hypothesis space of reasonably large plausibility after receipt of the radiation data) to decrease by a factor of $\exp(H) \approx 1.55 \times 10^{23}$ and 1.26×10^{26} , respectively, relative to the “prior volume” of the hypothesis space (volume of the hypothesis space of reasonably large plausibility before the receipt of the radiation data). It is seen that using radiation data measured with exposure time of $\tau = 48$ s provided about 7 nits more information gain than that using radiation data measured with an exposure time of $\tau = 16$ s.

7. Conclusions

This paper presented two approaches for estimating an unknown number of point radiation sources and their parameters employing radiation measurements collected using a gamma radiation detector. The first approach used importance sampling with PC to sample the posterior distribution, and the second approach used the reversible-jump MCMC technique in conjunction with simulated annealing for this purpose. The two approaches were compared by applying them to experimental data collected in a recent field trial.

It was demonstrated that both approaches perform well on the experimental data: in general, the number of sources is correctly determined and the parameters (e.g., location and intensity rates) that characterize each radiological point source are reliably estimated, along with a determination of the uncertainty (e.g., standard deviation, credible intervals) in the

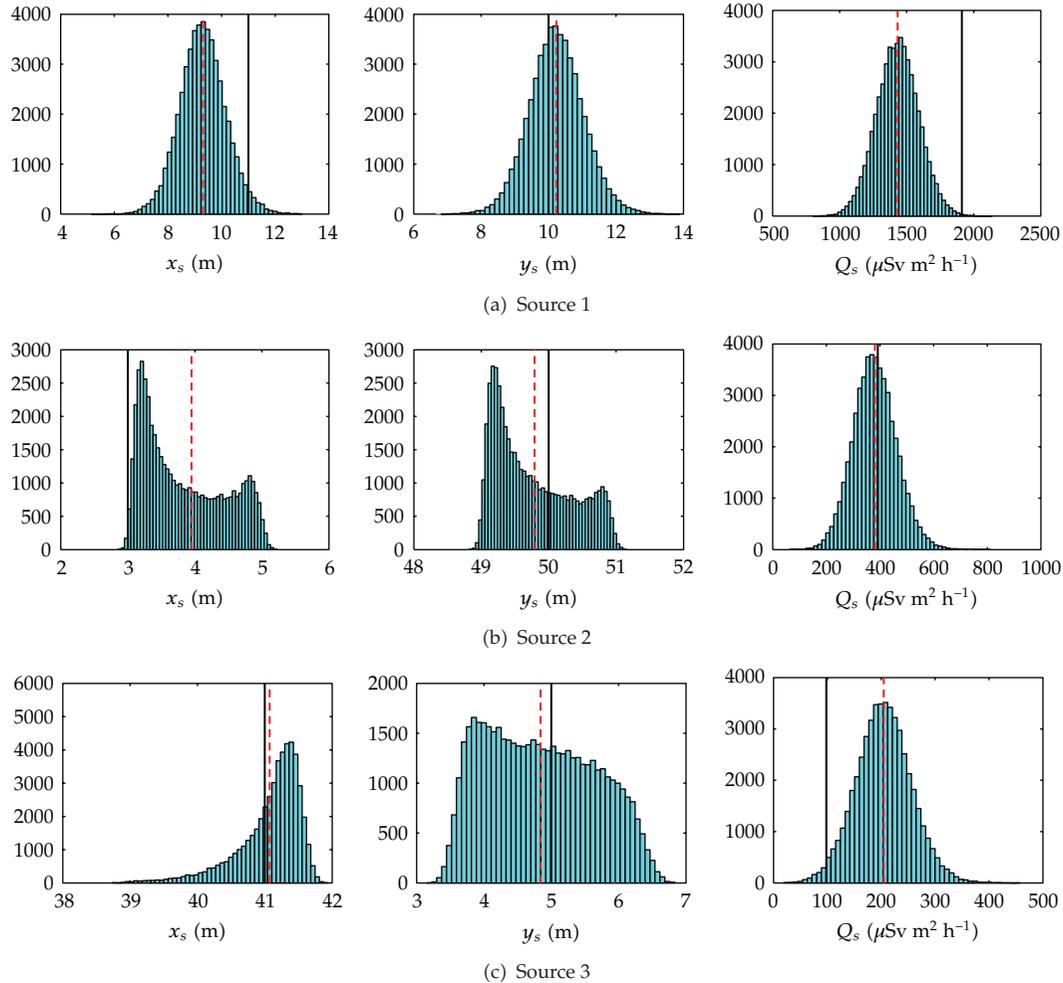


Figure 5: Inference of the parameters of the three radiological point sources obtained from samples drawn from the posterior distribution $p(\Theta | z^b, \mathbf{z}, I)$ for $\tau = 48$ s using the RJMCMC algorithm. (a, b, c) Histograms for the three parameters, namely, x_s -coordinate of source, y_s -coordinate of source, and source intensity rate Q_s that characterize sources 1, 2, and 3. In each frame, the solid vertical line indicates the true value of the parameter and the dashed vertical line corresponds to the best estimate of the parameter obtained as the posterior mean of the marginal posterior distribution for the parameter.

inference of the source parameters. However, it was shown that application of the first approach with $\alpha = 1$ did not adequately characterize the uncertainty in the source parameter estimates for the radiation measurements obtained with an exposure time $\tau = 48$ s. The defect here was not due to the progressive correction algorithm *per se*, but rather the result of the fact that the simple form of the posterior distribution of θ used with this algorithm assumed that the model error was negligible. This turned out not to be the case, and application of the RJMCMC with a more sophisticated form of the posterior distribution of Θ that incorporated explicitly the model error was shown to correct this defect.

The results of this paper suggest that Bayesian probability theory is a powerful tool for the formulation of methodologies for source reconstruction. It should be stated that

the application of importance sampling using PC to the posterior distribution of (5.1) is fully Bayesian in character (assuming that r is known *a priori*), as is the application of the RJMCMC algorithm to the posterior distribution of (5.2). However, the use of MDL in conjunction with the importance sampling algorithm to determine an estimate of the number of sources is not consistent with a strict Bayesian formalism.

From this perspective, it should be noted that if the Bayes factor is used instead to determine the number of sources (namely, to determine which model of a distribution of radiological point sources to use), then the methodology would be fully consistent with a Bayesian formalism (see, Morelande and Ristic [8]). In fact, as discussed in Jaynes [12], application of the Bayes factor for model selection automatically incorporates Ockham's razor (namely, selecting the simplest model for a distribution of radiological point sources that "best" fits the available radiation data).

Importance sampling using PC may also be carried out over the joint space of r and θ in the same way as the RJMCMC to estimate the number of sources and parameters associated with these sources simultaneously and in a manner that is fully consistent with the Bayesian formalism (see Ristic et al. [27]).

Acknowledgments

The authors would like to thank A. Eleftherakis, D. Marinaro, P. Harty, and I. Mitchell for their involvement in the field trial and A. Skvortsov, R. Gailis, and M. Morelande for useful technical discussions. This work has been partially supported by Chemical Biological Radiological-Nuclear and Explosives Research and Technology Initiative (CRTI) program under project number CRTI-07-0196TD.

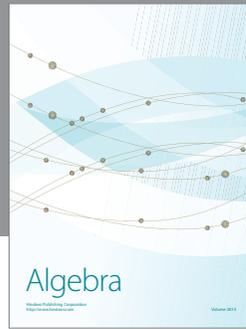
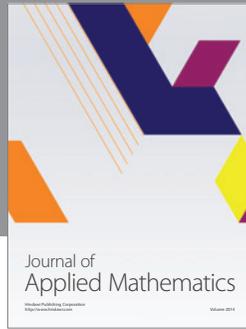
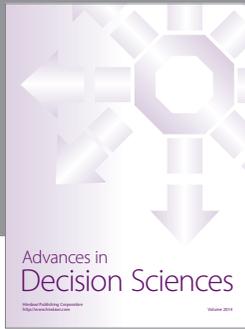
Endnotes

1. This hypothesis is of the form "the unknown value of the source parameters is Θ ", so the numerical value of Θ indexes a set of hypotheses of this form.
2. The evidence Z is the normalization constant in (5.2), required to ensure that the posterior distribution is properly normalized (namely, $Z \equiv \int p(\Theta | I)p(\mathbf{z} | z^b, \Theta, I) d\Theta$, so that $p(\Theta | z^b, \mathbf{z}, I) = p(\Theta | I)p(\mathbf{z} | z^b, \Theta, I)/Z$). The value for the evidence Z should not be too small, as this would imply that the "overlap" of $p(\Theta | I)$ and $p(\mathbf{z} | z^b, \Theta, I)$ in the hypothesis space (involving hypotheses concerning Θ) is minimal. This would suggest that the prior distribution of Θ and the measurements \mathbf{z} assimilated through the likelihood function are inconsistent).
3. Actually, the technique of thermodynamic integration dates back to Kirkwood [28] who first used it within the context of the statistical mechanics of fluid mixtures to compute free energy differences.
4. The relative entropy is measured in natural units (nits) rather than the more usual binary units (bits) because a natural logarithm was used in the definition of the relative entropy, rather than a logarithm to the base 2 (see (5.8)).

References

- [1] W. K. H. Panofsky, "Nuclear proliferation risks, new and old," *Issues in Science and Technology*, vol. 19, no. 4, pp. 73–75, 2003.

- [2] J. W. Howse, L. O. Ticknor, and K. Muske, "Least squares estimation techniques for position tracking of radioactive sources," *Automatica*, vol. 37, no. 11, pp. 1727–1737, 2001.
- [3] D. L. Stephens Jr. and A. J. Peurrung, "Detection of moving radioactive sources using sensor networks," *IEEE Transactions on Nuclear Science*, vol. 51, no. 5 I, pp. 2273–2278, 2004.
- [4] R. J. Nemzek, J. S. Dreicer, D. C. Torney, and T. T. Warnock, "Distributed sensor networks for detection of mobile radioactive sources," *IEEE Transactions on Nuclear Science*, vol. 51, no. 4, pp. 1693–1700, 2004.
- [5] S. M. Brennan, A. M. Mielke, and D. C. Torney, "Radioactive source detection by sensor networks," *IEEE Transactions on Nuclear Science*, vol. 52, no. 3, pp. 813–819, 2005.
- [6] A. Gunatilaka, B. Ristic, and R. Gailis, "On localisation of a radiological point source," in *Proceedings of the Information, Decision and Control, (IDC '07)*, pp. 236–241, Adelaide, Australia, February 2007.
- [7] M. Morelande, B. Ristic, and A. Gunatilaka, "Detection and parameter estimation of multiple radioactive sources," in *Proceedings of the 10th International Conference on Information Fusion, (FUSION '07)*, Quebec, Canada, July 2007.
- [8] M. Morelande and B. Ristic, "Radiological source detection and localisation using Bayesian techniques," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4220–4231, 2009.
- [9] C. Mendis, A. Gunatilaka, B. Ristic, S. Karunasekera, and A. Skvortsov, "Experimental verification of evolutionary estimation algorithms for radioactive source localisation," in *Proceedings of the 5th International Conference on Intelligent Sensors, Sensor Networks and Information Processing, (ISSNIP '09)*, pp. 151–156, Melbourne, Australia, December 2009.
- [10] N. Tsoulfanidis, *Measurement and Detection of Radiation*, Taylor & Francis, 1995.
- [11] A. Martin and S. A. Harbison, *An Introduction to Radiation Protection*, Chapman & Hall, 3rd edition, 1986.
- [12] E. T. Jaynes, *Probability Theory: The Logic of Science*, Cambridge University Press, Cambridge, UK, 2003.
- [13] T. J. Loredo, "From Laplace to supernova SN 1987A: Bayesian inference in astrophysics," in *Maximum Entropy and Bayesian Methods*, P. F. Fougère, Ed., pp. 81–142, Kluwer Academic Publishers, 1990.
- [14] C. Musso, N. Oudjane, and F. LeGland, "Improving regularized particle filters," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, N. de Freitas, and N. J. Gordon, Eds., Springer, New York, NY, USA, 2001.
- [15] G. Kitagawa, "Monte Carlo filter and smoother for non-Gaussian nonlinear state space models," *Journal of Computational and Graphical Statistics*, vol. 5, no. 1, pp. 1–25, 1996.
- [16] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*, Artech House, 2004.
- [17] R. Mahler, *Statistical Multisource Multitarget Information Fusion*, Artech House, 2007.
- [18] S. M. Kay, *Statistical Signal Processing: Detection Theory*, Prentice-Hall, 1998.
- [19] J. Rissanen, "Modelling by shortest data description," *Automatica*, vol. 14, pp. 465–471, 1978.
- [20] P. J. Green, "Reversible jump Markov chain Monte Carlo computation and Bayesian model determination," *Biometrika*, vol. 82, no. 4, pp. 711–732, 1995.
- [21] E. Yee, "Inverse dispersion of an unknown number of continuous sources," in *Proceedings of the 15th Joint Conference of Air Pollution Meteorology with the A&WMA*, p. 17, New Orleans, La, USA, 2008, paper no. 7.1.
- [22] E. Yee, "Theory for reconstruction of an unknown number of contaminant sources using probabilistic inference," *Boundary-Layer Meteorology*, vol. 127, no. 3, pp. 359–394, 2008.
- [23] D. Gamerman and H. F. Lopes, *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, Chapman & Hall, 2nd edition, 2006.
- [24] *AN/PDR-77 User's Guide*, Canberra Industries, Meriden, Conn, USA, 2000.
- [25] W. von der Linden, R. Fischer, and V. Dose, "Evidence integrals," in *Maximum Entropy and Bayesian Methods*, K. M. Hanson and R. N. Silver, Eds., Kluwer Academic Publishers, 1996.
- [26] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, NY, USA, 1991.
- [27] B. Ristic, M. Morelande, and A. Gunatilaka, "Information driven search for point sources of gamma radiation," *Signal Processing*, vol. 90, no. 4, pp. 1225–1239, 2010.
- [28] J. G. Kirkwood, "Statistical mechanics of fluid mixtures," *The Journal of Chemical Physics*, vol. 3, no. 5, pp. 300–313, 1935.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

