Inference of Unexploded Ordnance (UXO) by Probabilistic Inversion of Magnetic Data

Wigh, Mark David; Hansen, Thomas Mejer; Døssing, Arne

Published in:
Geophysical Journal International

Link to article, DOI:
10.1093/gji/ggz421

Publication date:
2020

Document Version
Peer reviewed version

Citation (APA):
Inference of Unexploded Ordnance (UXO) by Probabilistic Inversion of Magnetic Data

Mark David Wigh¹, Thomas Mejer Hansen² and Arne Døssing¹

¹ Crustal Magnetism Technology and Research Group, DTU Space, Technical University of Denmark, Centrifugevej 356, Denmark. E-mail: mdwi@space.dtu.dk
² Department of Geoscience, Aarhus University, Høgh guldbergs-gade 2, Denmark

17 September 2019

SUMMARY

Magnetic modelling of unexploded ordnance (UXO) is a well documented method used to interpret magnetic anomalies occurring in UXO excavation surveys. By treating UXO as a ferrous spheroidal object, the induced dipole moment can be estimated by approximation of UXO characteristics such as shape, size, and orientation. Inversion of magnetic data with respect to UXO requires one to solve the equation for the induced dipole moment, while also determining the location and orientation of the object. This is a highly non-linear, non-unique problem, where many solutions often are present, which make it difficult for standard inversion methods, such as linearized approaches and maximum likelihood estimators, in assessing uncertainties and correlations in estimated model parameters that often result in an incomplete solution. In this study, we treat the problem concerning magnetic UXO inversion by a probabilistic approach using Markov Chain Monte Carlo (MCMC) sampling. To deal with the potential multimodality a combination of two well known MCMC sampling methods is employed in a single-chain approach: The extended Metropolis algorithm is used for efficient local sampling, and the Gibbs sampler is used to help exploring the possible multimodal density of the posterior. By adding a Gibbs step we significantly increase the efficiency of the extended Metropolis, making it viable to use as a single-chain sampler for this problem. We refer to the algorithm as the Gibbs-within-
Metropolis algorithm. We test and compare the proposed algorithm to a multi-chain McMC parallel tempering setup using the extended Metropolis algorithm. We then present synthetic cases and a real data case, where the algorithm is used. We demonstrate how the probabilistic approach allow full inference of the parameters describing an UXO, while also including the possible presence of remanent magnetization.

**Key words:** Inverse theory – Probability distributions – Magnetic anomalies: modelling and interpretation

1 INTRODUCTION

Contamination of Unexploded ordnance (UXO) is a world-wide problem and poses a major safety hazard in public and industrial aspects. UXO are remains of former warfares, where armed conflicts have resulted in an abundant use of bombs, mines, and projectiles. In Europe, the dominant part of UXO are remains from WWI and WWII, contaminating land areas as well as the seabed, which, besides posing a safety hazard, also hinders productive use of the affected areas. For instance, dredging of the seabed or cable laying from offshore production sites require an exhaustive search of UXO prior to construction. This increases the cost of such projects significantly in terms of delays and increased labour required.

Remote sensing is the dominant technique in UXO clean-up methods, since UXO generally are made from ferromagnetic material. Hence, it is often possible to model and measure the magnetic signature of UXO. Much effort has been put into research of this subject in terms of magnetostatics and electromagnetic induction (EMI) each with their particular strengths and weaknesses. Magnetic type of data is in general easier to obtain through measurements, is not sensitive to sensor orientation, and is better at detecting deeper targets (Billings et al. 2010b). Magnetometry is, however, more sensitive to magnetic geology in the survey area, and it measures both the induced and remanent part of the magnetization of ferrous objects, such as UXO. EMI data are better at detecting smaller objects, but are also sensitive to non-ferrous metals. EMI is better at detecting targets in areas with high magnetic geology and is not affected by remanent magnetization in the object, which improves the intrinsic discrimination ability (Billings et al. 2010a). Both methods can be used alone in an inversion procedure, or cooperatively to improve interpretation of data by sequentially or simultaneously collecting each data set. The viability of each method, however, is dependent on the survey environment, for instance, marine or land may pose different challenges in the data collecting and interpretation phase.

Interpretation of magnetic anomaly data to infer information on a point source such as (x,y,z)-position, shape, volume and orientation of the object is a highly nonlinear inverse problem, that often possesses non-unique features, due to model ambiguities and high correlation
between model parameters. Different solutions often fit the magnetic signal equally well, so that approximations to the full nonlinear inverse problem tend to fail in giving a realistic, consistent solution. In the past twenty years extensive research has been done in the subject of UXO discrimination and classification. In the region of magnetometry UXO are modelled as static dipoles, where the dipole moment can be characterized by a ferrous spheroid with a given susceptibility and aspect ratio. Modelling the UXO as spheroidal object was first made in the mid 90s (Altshuler 1996; Butler et al. 1998) and later were used in a discrimination procedure (Billings 2004; Billings 2006; Billings & Youmans 2006; Billings 2009; Butler et al. 2012), that has provided good results, given that the UXO did not contain significant remanent magnetization. In the aspect of positioning and depth estimates, automatic detection algorithms have previously been made with notable results using continuous wavelet transforms (Billings & Herrmann 2003), and a self-adaptive fuzzy c-means clustering algorithm (Yin et al. 2017).

EMI has in recent years had increased usage in the aspect of UXO discrimination. It has been used in classification algorithms due to the high intrinsic discrimination ability for EMI data, but it can also be used for target location and orientation estimates. (Beran et al. 2011a) use a robust inversion algorithm in a linearized least-squares approach to minimize an objective function. The non-linearity of the forward model, that produces local minima in the objective function, is dealt with by selecting several starting models with different locations and orientations to ensure proper exploration of the misfit function. A different approach is made in (Beran et al. 2011b), where the uncertainty of the feature vectors in a classification algorithm is evaluated by exploring the posterior density using MCMC sampling. Simplified means are used to explore the nonlinear probability density, which in many cases do not ensure an adequate sample of the true distribution. Recent examples with cooperative UXO inversion of EMI and total field magnetics can be seen in (Pasion 2007; Pasion et al. 2008; Billings et al. 2010b). Common for these studies, in particular, is the usage of magnetic data to estimate object depth and position, which is used as a constraint for inversion of the EMI data. Other examples of usage of MCMC sampling related to UXO processing and inversion can be seen in (Aliamiri & Miller 2007; Tantum et al. 2008), where MCMC is used to address the uncertainties related to sensor position errors.

We focus on magnetostatics, where a static dipole moment is fitted to the source by approximating it as an elongated shaped ferromagnetic (prolate) spheroid. We solve the problem by a probabilistic inversion on depth, volume, aspect ratio of the spheroid, orientation, x-,y- position of the ordnance and possible remanent magnetization present in the source. As stated, this is a highly non-linear problem, where many different solutions and correlations are present and a complete exploration of the full solution is often overlooked. In order to sufficiently solve the problem we present a hybrid sampling method in a probabilistic framework, that utilizes the Metropolis algorithm (Metropolis et al. 1953; Mosegaard & Tarantola 1995), allowing for efficient local sampling, and the Gibbs sampler (Geman & Geman 1984; Gelfand & Smith 1992), that allow exploring of multimodality. Both algorithms are used in one combined process to achieve a more efficient single-chain sampling method, compared to a regular extended Metropolis algorithm. Furthermore, we show that uncertainties of estimated UXO size and shape associated with remanent magnetization in ferrous materials can be accounted for with a probabilistic approach, once
M. D. Wigh, T. M. Hansen and A. Døssing.

certain hard constraints are set prior to the inversion. The inversion method will be tested on synthetic data generated from various realizations of a possible UXO in a real survey. Finally, we will test on real total field magnetic data of a measured anomaly from a 155mm Projectile buried at a test site. Inversion results will be shown for inversion on induced magnetization only, as well as the more complex scenario, where remanent and induced magnetization are simultaneously present.

2 METHODOLOGY

The magnetic field of a ferromagnetic induced object is often complex due to the strength of the demagnetization field, which is dependent on the shape and orientation of the object with respect to the ambient field. Calculating the demagnetizing field of all complex UXO shapes is practically impossible, so one can only approximate the induced magnetic response of simple shapes, such as prisms or spheroids.

In the current study we will focus on modelling the static dipole field of spheroids since UXO can generally be seen as compact spheres or compact elongated spheroids. The forward model is termed the 'spheroidal model' (Altshuler 1996). The idea is to model any UXO as a compact ferrous spheroidal object with a specified orientation in space. Dependent on the shape and orientation of the object, it will induce a magnetization, and if there is any remanent magnetization in the object, then the total magnetization will be a superposition of both. The dipole moment of the total magnetization is then given by (Billings 2006)

\[
m = \frac{V}{\mu_0} A^T \chi_d A b_0 + A^T m_r
\]

where \( m \) is the dipole moment of the anomaly, \( m_r \) is the remanent magnetization of the object, \( V \) its volume, \( A \) is the Euler rotation matrix, \( \chi_d \) is the effective susceptibilities (a diagonal matrix), and \( b_0 \) is Earth’s magnetic field. \( \mu_0 \) is the permeability of free space.

We adopt the three Euler angles as \((\gamma, \theta, \phi)\), where \( \gamma \) is the rotation about the spheroidal objects symmetry axis, \( \theta \) is the dip angle of the semimajor axis of the spheroid with respect to the xy-plane, and \( \phi \) is the azimuthal angle of the semimajor axis relative to true north (positive east).

Here we assume that all objects are symmetric around their rotation axis, thus \( \gamma = 0 \). The default setting of the forward model has the semimajor axis of the ordnance positioned along the y-axis (North) with a dip angle of 0 relative to the xy-plane, such that an orientation of \((\theta = 0, \phi = 0)\) corresponds to an elongated spheroid with the semi-major axis facing north, lying flat on the xy-plane. With the above conventions, the Euler rotation matrix is defined in eq. (2)

The coordinate system used in the modelling is illustrated in Fig. 1. It differs from the usual geomagnetic coordinate system, IGRF, such that the Y-axis is aligned with true north, X is positive east and Z is positive down. The declination of Earth’s field and declination of remanent magnetization is defined as the angle between the Y-axis (North) and the projected vector positive from north (+Y) to east (+X)
2.1 Estimation of the induced magnetic moment

When modelling within an area of no current density the magnetic field can be modelled as a potential field and thus it can be expressed as the gradient of a scalar potential, $V_p$. For a compact body with volume $V$ and a magnetization $M(x)$ the scalar potential is given by the

$$ V_p(x_o) = \frac{1}{4\pi} \int_V M(X) \cdot \nabla \left( \frac{1}{r} \right) \, dV $$

(3)

Where $r = ||r||$ is the distance from the observation point to the source. The usual method for estimating the scalar magnetic potential in this equation is by the method of multipole expansion of the potential. This can be expressed in terms of spherical harmonics, or it may be simplified and expressed as a sum of Taylor series expanding $(1/r)$ about the origin. Using Taylor expansion, the magnetic field can then be expressed as a sum of magnetic moments. The first term when performing a Taylor expansion of eq. (3) is the dipole moment which dominates the far-field of the source and is a simple three-element vector given by

$$ m_1 = \int_V M \cdot \nabla x_1 \, dV = \int_V M_1 \, dV. $$

(4)

Assuming the magnetization vector is constant throughout the volume, the dipole moment is the product of the magnetization with the volume, so eq. (4) can be reduced to

$$ m = M \int_V dV = MV. $$

(5)

The next two moments derived from the Taylor expansion are the quadropole and the octupole moments. For any compact body with

---

Figure 1. Coordinate system used in the modelling. Y-axis aligned with magnetic north, X positive east, Z positive down. The forward model is set up such that the $\phi$ angle is 0 when the semi-major axis of the spheroid is facing towards north, and $\theta$ angle is 0 when the spheroid is lying horizontal in the xy-plane.
M. D. Wigh, T. M. Hansen and A. Døssing.

three axes of symmetry (such as a cylinder or spheroid), the quadropole contributions will be zero. Asymmetry in the ambient field of the object can cause quadropole contributions, however this contribution dies out as the fourth power of distance. The last moment in consideration is the octupole moment, which dies out as the fifth power of distance. The octupole moment is generally nonzero, even for symmetric bodies such as a cylinder or spheroids. In this study, we are only modelling the dipolar effects of UXO, hence it is important that we only measure the dipole contributions of the magnetized source. As long as the observation point has a distance of at least twice the length of the object only the dipole moment is thought to be affecting the observer (Billings 2009).

In general and for weak magnetic fields, such as the geomagnetic field, the induced magnetization vector can be expressed as a product of the field amplitude and the magnetic susceptibility of the object as

\[ M_{in} = \frac{\chi}{\mu_0} b_0. \]  

However, self demagnetization effects of ferrous objects of fixed shapes are very significant and must be accounted for properly, when modelling the magnetic moment of ferrous objects (Tandon et al. 2004). The effects of demagnetization can be accounted for by calculating an effective susceptibility along all three axes. For an arbitrarily shaped body with shape \( \Omega \) and susceptibility \( \chi \), the effective susceptibility \( \chi \) will be a rank-two tensor as function of position, \( x \), shape and susceptibility \( \chi(\mathbf{x}, \chi, \Omega) \). By integrating over the volume of the object, one can define the average effective susceptibility \( \bar{\chi} \) as

\[ \bar{\chi} = \frac{1}{V} \int \chi(\mathbf{x}, \chi, \Omega) d\mathbf{x}. \]  

which has to be evaluated for all three principal axes, such that \( \bar{\chi} \) is a matrix containing the effective susceptibilities for \( (x,y,z) \)- along the diagonal of the matrix, respectively. Analytic determination of the average effective susceptibility of complex shapes is not possible, hence only simple shapes such as spheroids can be analytically determined. If we define the diameter of a spheroid as \( a \), and the length, \( L \), one may define the aspect ratio, \( e \), as \( L = ae \). If \( e = 1 \), we get \( L = a \) which is a perfect sphere. For \( e < 1 \) the shape resembles an elongated spheroid (Prolate spheroid) and for \( e > 1 \) it resembles an oblate spheroid. Here, we focus on prolate spheroids which are the most common structures of UXO. A solution to the boundary value problem in eq. (7) determines the effective susceptibility following the expression (Billings 2006)

\[ \bar{\chi}_i = \frac{\chi}{1 + \alpha_i \frac{1}{2}} \]  

where \( \alpha_1 = \alpha_2 \) and \( \alpha_3 \) are self-demagnetization factors that are dependent on the aspect ratio. These are given by

\[ \alpha_1 = \alpha_2 = \frac{e(e + E)}{e^2 - 1} \]
\[ \alpha_3 = \frac{-2e(e^{-1} + E)}{e^2 - 1}, \]

with \( E \) defined as (only for \( e < 1 \))

\[ E = \frac{\log(e - \sqrt{e^2 - 1})}{\sqrt{e^2 - 1}}. \]
From eq. (8) it can be seen, that when the susceptibility is large the effective susceptibility can be approximated by \( \tilde{\chi}_i = \frac{2}{\alpha_i} \). This means, that the induced magnetization becomes independent of the susceptibility for very high susceptibilities. Since steel usually has a susceptibility of several hundred (Billings 2006), it can be neglected in the modelling process.

### 2.2 Remanent Magnetization

All ferromagnetic metals have the ability to retain a magnetization in the absence of external magnetic fields. Large impacts on materials with a remanent magnetization can cause what is defined as shock demagnetization, which erases or changes the remanent magnetization of the material (Billings et al. 2014). This can cause significant changes for some UXO, such as projectiles, where shock demagnetization is possible when firing as well as on impact. However, it could be that the UXO never were affected by any shocks able to cause shock demagnetization that could erase its thermoremanent magnetization, which would be likely for UXO that are dropped from airplanes into the sea, or UXO that are deployed in the sea from ship.

The remanent magnetization serves to complicate any UXO discrimination and classification significantly since it is, in principle, uncorrelated with shape and size of the UXO. With remanent magnetization present the total magnetization is defined as the sum

\[
M = M_{rem} + M_{in}.
\]

### 2.3 Calculation of the anomalous magnetic field

The contribution of the estimated magnetic dipole moment given in eq. (1) can be calculated by

\[
b(x_0) = \frac{\mu_0}{4\pi r_0^3} \left( \frac{3}{r_0^2} [r_0 \cdot m] r_0 - m \right).
\]

Where \( \mu_0 = 4\pi \cdot 10^{-7} \text{H/m} \) is the permeability of free space and \( r_0 \) is the distance from the center of the object to the observation point. The scalar magnitude of the total field is interpreted as the projection of field \( b \) along the Earth’s main field with inclination \( I \) and declination \( D \) at the observation point. In Denmark the geomagnetic field is characterized by a magnitude of \( \approx 55.000 \text{nT} \) and inclination and declination of \( I = 70^\circ, D = 3.5^\circ \). The total field effect of the point dipole is computed by the dot product (Hinze 2013)

\[
\Delta F(x, y, z) = \hat{u} \cdot b(r) = \cos(I)\cos(D)b_x + \cos(I)\sin(D)b_y + \sin(I)b_z
\]

where \( \hat{u} = (\cos(I)\cos(D))\hat{e}_x + (\cos(I)\sin(D))\hat{e}_y + (\sin(I))\hat{e}_z \) is the unit geomagnetic field vector at the observation point.
2.4 Probabilistic inverse problem

Consider some data, $d$, which are indirect measurements of given model parameters, $m$, describing a physical system of interest. Given error-free values of $d$, the relation between the data and the model parameters are related through the function $g$ as

$$d = g(m). \quad (15)$$

Eq. (15) describes the forward problem which is defined as the prediction of observed data given information on the model parameters. The inverse problem consists of inferring information about the model parameters, $m$, given some data observations $d$. In a probabilistic framework the inverse problem consists of combining all known information (quantified using a probability density) into a posterior probability density that describes the combined information. Typically available information consists of prior information about $m$ quantified through $\rho_M(m)$, and information from geophysical data quantified through a likelihood function, $L(m)$. Then the a posteriori probability density $\sigma_M(m)$, which is the solution to the inverse problem, is proportional to the product of the prior probability density and the likelihood function (Tarantola & Valette 1982a; Tarantola 2005)

$$\sigma_M(m) = k \rho_M(m) L(m) \quad (16)$$

Where $\rho_M(m)$ is the prior probability density (PD) that contains all prior knowledge about the model parameters. $k$ is a normalization constant, and the likelihood function, $L(m)$, describes the fit between observed data and the data computed from the model parameters through the forward relation following the equation

$$L(m) = \int_D d \rho_D(d) \frac{\theta(d|m)}{\mu_D(d)} \quad (17)$$

$\rho_D(d)$ describes the measurement uncertainties such as errors caused by uncertainties in the instruments or unwanted signals contaminated in the data, which often is an issue in magnetic surveys, $\theta(d|m)$ is the theoretical probability density and represents errors related solving the forward problem, such as errors due to imperfect parameterization or forward model. Finally, $\mu_M(m)$ is the homogeneous probability density, that ensures the parameterization is invariant to changes in the coordinate system.

Under the assumption of negligible theoretical error and uncorrelated Gaussian noise, which we will use here, the data misfit is calculated following a $l_2$ norm, so the likelihood function is defined as

$$L(m) = k \exp \left( - \sum_i \frac{(g'(m) - d_{i,\text{obs}})^2}{(\sigma^2)^2} \right) \quad (18)$$

where $k$ is a constant and $\sigma$ is the standard deviation of the uncorrelated noise.


2.5 Sampling the posterior probability density

When dealing with non-linear inverse problems with a large number of model parameters it gets inefficient, or, in some cases practically impossible, to solve the inverse problem analytically by computing the posterior PD, as given by eq. (16). Instead, sampling methods exist that allow sampling of the posterior PD, see for example (Mosegaard & Sambrigde 2002). If correctly applied such methods will generate a collection of models that will be distributed according to the posterior PD. Each of these models will be referred to as a ‘realization’, and the collection of realizations as a ‘sample’ of the posterior PD.

Perhaps the most widely used sampling method is the Metropolis algorithm and the more general Metropolis-Hastings algorithm (Metropolis et al. 1953; Hastings 1970) and variants thereof, see for example (Laloy & Vrugt 2012). These are Markov Chain Monte Carlo (McMC) methods that allow sampling of any probability distribution through a random walk. The only requirement is that the probability density that one wishes to sample can be evaluated for any given model proposal, and that the random walk satisfies ergodicity and detailed balance during sampling. In the present case this refers to the posterior PD.

While the Metropolis algorithm is guaranteed to sample the correct distribution over time, it may be computationally intractable depending on the problem at hand. If for example the posterior PD is multimodal, a traditional Metropolis algorithm may have difficulty in jumping between areas of high posterior probability. To address this issue a multitude of methods has been proposed. (Liu et al. 2000) propose to use multiple proposals model at each step iteration of the Markov chain, to allow taking both small and large jumps. This is utilized in (Laloy & Vrugt 2012) that present an efficient parallel implementation of the multiple-try Metropolis algorithm. The number of multiple tries, as well as the magnitude require some manual adjustment.

Parallel tempering refer to a class of a multi-chain Monte Carlo methods in which multiple chains of Metropolis algorithms are run in parallel, each with a unique and increasing temperature (related to the uncertainty of the data) (Geyer 1991; Earl & Deem 2005; Sambrigde 2013). Chains with higher temperature explore more of the model space (but fit the data less accurate). Pairs of chains change position at random iteration with a probability that ensures detailed balance, and hence, a correct sampling of the posterior. The exact number of chains to use, and the temperature profile, requires some manual adjustment for optimal performance.

In other cases the shape of the not necessarily multimodal PD is such that the PD is difficult to sample with conventional Metropolis-Hastings algorithm. In these cases the use of information about the gradient of the PD can make the sampling for efficient. Hamiltonian use information about the gradient (Duane et al. 1987; Neal et al. 2011), while stochastic Newton methods use information about the second order derivative (Martin et al. 2012), to achieve a better proposal as part of proposing an update at each iteration of the Markov chain. Both methods require that the derivative can be efficiently obtained in order to be efficient, and require some tuning of parameters controlling the inversion for optimal performance (Hoffman & Gelman 2014).
M. D. Wigh, T. M. Hansen and A. Døssing.

In the following we suggest using a combination of the Metropolis sampler that allow efficient local sampling of the posterior distribution, and the Gibbs sampler for a 2D subset of the full PD, that allow jumping between potential multimodal states of the PD. First we will briefly recall the two methods and then we will propose to use these together as an algorithm referred to as the Gibbs-within-Metropolis algorithm.

2.5.1 The extended Metropolis algorithm

The extended Metropolis (Mosegaard & Tarantola 1995) sampler is a variation of the Metropolis-Hastings (Metropolis et al. 1953). The extended Metropolis differs from the classical Metropolis algorithm in that one neither need to evaluate the posterior probability $\sigma_M(m)$, nor the prior probability $\rho_M(m)$ of a given model $m$, as given by eq. (16). It is enough that an algorithm exists that can sample the prior PD and a method exists for evaluating the likelihood. The only requirement when running the algorithm, is that one must be able to sample the prior PD such that apriodicity and irreducibility is ensured (Mosegaard & Sambrigde 2002).

If a realization of the prior PD is generated as $m_{\text{cur}}$ and the associated likelihood $L(m_{\text{cur}})$ is evaluated, then the extended Metropolis algorithm can be implemented by iterating over the following algorithm (Mosegaard & Tarantola 1995):

1. In the vicinity of $m_{\text{cur}}$, propose a new model candidate, $m_{\text{pro}}$, proportionally to the prior model.
2. Compute $L(m_{\text{pro}})$.
3. Accept the proposed model with probability $P_{\text{acc}} = \min([1, L(m_{\text{pro}})/L(m_{\text{cur}})])$
4. If the proposed model is accepted, then the transition from $m_{\text{cur}}$ to $m_{\text{pro}}$ is accepted and the proposed model becomes the current model as a realization in the sample. If not accepted, the random walker will stay at location $m_{\text{cur}}$ and $m_{\text{cur}}$ counts again.

This algorithm will end up sampling the posterior PD (Mosegaard & Tarantola 1995).

In the following we make us of (Hansen et al. 2013a) that provide a number of algorithms to sample a variety of prior models, from simple 1D to more complex 3D prior models. Note that we will only use simpler 1D prior models in which case the extended Metropolis differ only slightly in computational efficiency to the use of the Metropolis algorithm.

Typically the use of the Metropolis sampler is adjusted (through adjustment of the way new models are proposed) such that an acceptance rate of about 30%. This may be a good choice to sample local properties of the posterior PD, but may also make it difficult to move between different areas of high posterior probability.
2.5.2 Gibbs Sampling

The Gibbs sampler, closely related to the Metropolis algorithm, is a sampling algorithm that allows sampling of a probability density (Geman & Geman 1984). It differs from the Metropolis algorithm in that in each iteration of the Markov chain, typically a single model parameter, $m_i$, is updated by sampling $m_i^*$ from the conditional posterior distribution:

$$m_i^* \sim \sigma_M(m_i|m_c), \quad (19)$$

where $m_c = [m_1, m_2, ..., m_{i-1}, m_{i+1}, ..., m_M]$, meaning that all other model parameters are kept fixed.

Here we will use Gibbs sampler to draw two (randomly chosen) model parameters $[m_a^*, m_b^*]$ from the 2D conditional distribution $\sigma_M(m_a, m_b|m_c)$. To do this first the 2D conditional distribution is established, from which the realized values $[m_a^*, m_b^*]$ are then simulated.

This is done according to the following pseudo-code:

1. Randomly select 2 model parameters for Gibbs sampling, $[m_a, m_b]$, from the current model parameters $m_{\text{cur}}$.
2. Generate $N_r$ sets of points $[m_a^i, m_b^i]$, $i = 1...N_r$, in the $[m_a, m_b]$ space from $\rho(m_a, m_b)$, and evaluate $\sigma(m_a = m_a^i, m_b = m_b^i|m_c)$ for all $N_r$ set of data points.
3. Interpolate to the $N_r$ data points to estimate the conditional distribution $\sigma(m_a, m_b|m_c)$ using nearest neighbour interpolation.
4. Generate one realization from the conditional density distribution $(m_a, m_b|m_c) \rightarrow m_a^*, m_b^*$.

Interpolating over the 2D conditional in order to draw a realization from result in an approximate 2D marginal distribution, may result in accepting models that will decrease the likelihood significantly. This jump can be mitigated by increasing the number of proposed points, $N_r$, for better interpolation of the conditional. However, we found it to be efficient when $N_r$ is kept low, so the accepted model will be likely to cause a jump in the likelihood, kicking the chain slightly out of place, which may help the exploration nature of the extended Metropolis algorithm it will be used with.

Alternatively, one can simplify the Gibbs step by accepting one model, from the $N_R$ generated points from $\rho(m_1, m_2)$, proportionally to the associated likelihood $L(m_a, m_b|m_c)$, and thus avoiding interpolating the 2D conditional density distribution. The accepted models will be more exact regarding the likelihood, but often results in small steps being accepted by the Gibbs step. A large step could be made this way by increasing the number of Gibbs iterations, however, it would be a time consuming process and significantly increase the computation time of the algorithm. For this reason, we only use the method with nearest neighbour interpolations, as explained in the above pseudo-code.

2.5.3 Gibbs-Within-Metropolis

We propose a sampling strategy that combines the strengths of the Gibbs sampler and the extended Metropolis sampler. The idea is to densely sample an area of interest (using the extended Metropolis sampler) and allow larger jumps between significant different posterior models (by
Figure 2. Flowchart of Gibbs-within-Metropolis algorithm. MOD command stands for modulo operation, meaning that every time the loop has carried out \(i_{\text{gibbs}}\) iterations, then it will perform a Gibbs step as explained in section 2.5.2, otherwise it will continue with the extended Metropolis algorithm as explained in section 2.5.1

performing Gibbs sampling on a 1D- or 2D- marginal of selected parameters). The Gibbs sampler has been implemented in combination with the extended Metropolis that allows simulating from both a 1D and 2D conditional distribution, thus allowing to update one or two model parameters simultaneously conditional to the other parameters. This may help the sampling algorithm to improve exploration of the posterior distribution, especially between correlated model parameters where the absence of uniqueness of the fitted dipole is a problem, and where the extended Metropolis might struggle to explore solutions of interest.

Both the (extended) Metropolis sampler and the Gibbs sampler described above are iterative Markov chain Monte Carlo methods that will sample the posterior distribution. Therefore, any combination of the two methods will also sample the posterior distribution. We therefore propose the hybrid MC/MC algorithm, called Gibbs-within-Metropolis, in which each iteration is mostly performed using the Metropolis rule, but once in a while using the Gibbs sampler. The algorithm is illustrated in the flowchart shown in Fig. 2. The algorithm starts by initializing the extended Metropolis, hereby drawing a realization from the prior. Every \(i_{\text{gibbs}}\)’th iteration a Gibbs step will be performed instead of the extended Metropolis. After either a Gibbs step or extended Metropolis the current model will be saved, and the loop will continue. It is up to the user to specify when a saved model should be saved as a realization of the posterior distribution. The algorithm will continue to sample until it reaches the user specified amount of iterations as given in the variable \(mcmc\_n\.)

The idea to combine different McMC algorithms into one hybrid is not new. Discussion in the subject of constructing hybrid algorithms by combination of the rejection sampler, Metropolis and Gibbs samplers can be found in (Tierney 1994). Similarities can be drawn to other
methods such as Metropolis-within-Gibbs, albeit the circumstances are different. Metropolis-within-Gibbs is used when the conditional
distribution cannot be sampled directly for all parameters, such that the Gibbs sampler is substituted with a Metropolis step to perturb the
parameter. In our case we do not use the Metropolis because we can’t sample the conditionals, we use the Gibbs sampler to improve the
efficiency of the extended Metropolis. Hence, the algorithm is termed Gibbs-within-Metropolis.

The hybrid sampling algorithm has shown good capabilities in sampling the posterior PD for our specific problem at hand. As mentioned
above other methods do exist, and in section 3.2 we compare the results of our algorithm with results derived using parallel tempering, a multi-
chain McMC technique that facilitates jumps between independent extended Metropolis chains, each sampling with a specified temperature.
Parallel tempering is implemented according to (Sambrigde 2013), where the method is explained in greater detail.

2.6 Quantifying prior information

A strong attribute in the probabilistic method of inverse problems is the ability to incorporate extensive prior knowledge during an inver-
sion. Usually, as in the current study, this consists of prior knowledge on the distribution of the model parameters, $\rho_m(m)$, which assigns a
probability to each possible model value. In total we have up to ten model parameters. The distribution of all model parameters are assumed
to be independent of each other and we may sufficiently describe each by a 1-dimensional probability distribution. The range of the prior
distributions will constrain the sampling process within the model space, while affecting the posterior PD. Thus, prior information can make
the inversion significantly easier, but also bias the solution if the information is wrong. The ten model parameters are listed in Table 1 for a
full-scale inversion on shape, size, position and remanent magnetization in the UXO.

Due to the complex nature of determining a point source dipole moment from a measured anomaly, some parameters are heavily correlated
with each other (The 1D prior distributions are still assumed independent of each other, however), so any constraint in the model space, such as max-
and minimum values defined as a priori constraints, are advantageous to confine the region of solutions within a specified
area. Spheroids of different shapes, sizes, and orientations can produce the same dipole moment so non-unique solutions are frequent. Ad-
ditionally, when inverting for both induced and remanent magnetization there can exist infinite solutions, unless constraints on the remanent
magnetization are set.

Characteristics of the prior distribution for each model parameter are also shown in Table 1. Most 1D prior distributions are approximated
by a simple uniform- or Gaussian-distribution while some require complex prior model distributions in order to properly incorporate prior
information. Examples of prior models that are used in the study are seen in Fig. 3, showing different types of distributions that will be
explained in detail below.

The prior model for depth of the source is created as a complex distribution in order to resemble burial depth information, in this case
of typical WWII UXO in the North Sea between UK and Denmark. The depth distribution is made from empirical observations provided by
the Royal Danish Navy EOD Service, stating that most UXO are found on the seabed or partly buried, while few are found at depths of 1.5 m or further below the seabed (Commander L. Møller. Royal Danish Navy EOD service, pers. comm., May 2017). The prior model for the aspect ratio is based on the most likely aspect ratio for some types of UXO is $\approx 4$ (Billings 2004). We assume that the majority of UXO has an aspect ratio close to 4, while other ratios are possible, but not as frequent. Thus, we believe that in a general case the aspect ratio can appropriately be defined as a Gaussian distribution centered at 4.

Regarding orientation ($\Theta, \Phi$), only the distribution for $\Theta$ can be approximated and constrained by utilizing the fact, that elongated bombs rarely have the long axis pointing up/down, but often lie relatively flat on or beneath the seabed. This will, however, be dependent on the actual seabed conditions at the survey site. According to the Royal Danish Navy EOD service the majority of elongated WWII UXO in the North Sea have $\Theta < \pm 20^\circ$ (Commander L. Møller. Royal Danish Navy EOD service, pers. comm., May 2017).

The remanent signal will be constrained by the Koenigsberger ratio (Q-ratio), where $Q = \frac{M_{\text{rem}}}{M_{\text{ind}}}$. The remanent magnetization is proportional to the induced magnetization, such that the remanent magnetization will never be present without any induced magnetization. This simplifies the sampling process of any inversion with remanent magnetization assumed. If there is no constraint on the remanent signal strength, then there will be infinite solutions to the problem. It is difficult to get any statistical insight into the amplitude of remanent magnetization present in UXO, since it is non diagnostic of size and shape of the UXO, but instead is defined by the object’s history. The prior model for the remanent signal will, thus, be a simple uniform distribution that is based empirically through collected data among different projects concerned with UXO inversions.
Figure 3. Examples of prior models that are used for the inversion of UXO in this study. The shown distributions consist of 50k samples of the target distributions. The depth distribution (m1) is an example of a complex distribution that cannot be described by simple analytic expressions and thus has to be sampled quantitatively. The red line shows the median of the distribution and the red dotted lines indicate the 5% and 95% confidence intervals, respectively.

(Billings 2009) has made an extensive search for remanent magnetization on smaller UXO (76-90 mm projectiles, 4.2 inch mortars), where 95% of all UXO had a Koenigsberger ratio $Q < 0.5$, and only very few had $Q > 1$. (Goodrich 2007) did additional research on the topic using the same measurement equipment as (Billings 2009) and obtained similar results. Both projects were carried out in America, while WWII UXO frequent in European waters tend to have larger volumes. In addition, they are usually air dropped bombs or mines deployed from ships, meaning that shock demagnetization might be less common and that remanent magnetization therefore is more significant. Based on the above, we have created a uniform model of the remanent amplitude ranging from 0% to 130% of the induced magnetization.

### 3 SYNTHETIC STUDY

To illustrate the inversion methodology and potential we initially consider 2 synthetic case studies. We have generated synthetic examples of different anomalies using the forward code. First, we consider an example of 2D marginal analysis of the likelihood function prior to inversion. Then we assess the inversion results for a case with no remanent magnetization, and finally we demonstrate an example where remanent magnetization is added to the source. Uncorrelated Gaussian noise has been added to the synthetic data sets, such that the signal to noise ratio (SN) is SN=6.

#### 3.1 2D model space sensitivity analysis

The strong non-linearity and high correlation between model parameters significantly increase the complexity of the inverse problem. Non-uniqueness is frequent due to the ambiguity in the modelling of prolate spheroids, where many different shapes and orientations can reproduce the same dipole moment. The posterior PD is often complex shaped with multimodal features and "small" steep regions of optimal solutions, where only a very narrow part of it is of any interest. There can be several local optima with great distances apart, which makes sampling of
the posterior difficult. Different 2D-Likelihood conditional distributions (all other parameters are fixed) are seen in Fig. 4 in order to try and visualize the intricacy of the sampling process, and to quantify the chosen sampling strategy for the problem. The log-likelihood marginal for \( \text{Theta} \) and \( \text{Volume} \) for an example with no remanent magnetization is shown in Fig. 4(a), and the normalized likelihood marginal is seen in Fig. 4(b). Another example of the likelihood marginal of \( \text{Volume} \) and \( M_{\text{rem}} \) is illustrated in Figs(4c and d), where remanent magnetization is added to the source. The noise is exaggerated by a factor of 20 in order to better visualize the structure of the likelihood, so the non-Gaussian structure of the marginals distributions are easily recognized. The small steep regions of interest are apparent in the likelihood plots in Figs(4b and d) showing the normalized likelihood, where 1 is equal to a perfect fit given the noise level. This means that correlations are certain in the model solution, for example when inspecting Fig. 4(d) it can be seen how a decrease in volume can be accounted by a change in the amplitude of the remanent magnetization, while still fitting the signal equally. Further, the 2D conditional in Fig. 4(b) indicates bimodal features, which could pose a challenge when only using an extended or traditional Metropolis sampler.

The 2D marginals in Fig. 4 is the main reason we opted to develop the Gibbs-within-Metropolis algorithm described above. The Metropolis sampler alone would be prone to take very small steps in order to sample locally narrow densities, and would be be prone to get stuck in local neighborhoods. The Gibbs sampler would in principle allow larger jumps, but would be very inefficient by itself to sample local variations. The hybrid Gibbs-within-Metropolis algorithm will in principle allow efficient local sampling, as well as occasional large jumps from one local neighborhood to another.
Figure 5. a) Anomaly generated with forward model. Contour lines are plotted in 0.5 nT intervals b) Prior (black) and posterior (red) distributions for inversion on induced magnetization only on synthetic data set shown in a). 1.5 million hybrid Metropolis-Gibbs iterations. Reference model used to generate the anomaly is indicated by the green circle.

3.2 Results - Synthetic data (No Remanence)

Synthetic magnetic anomalies are computed in a 7x7 m area as depicted in Fig. 5(a) where data points are made with a resolution of dX = 1.00 m. and dY = 0.14 m, thus simulating a magnetic survey with 1 m line spacing and a spatial sampling frequency every 0.14 meter.

It should be noted that all inversion procedures in this study are done on data points and not gridded data. The gridded data shown in the following sections are for visual purposes of the anomaly.

1D-marginals of the posterior PD acquired from the inversion is presented in Fig. 5(b) for all model parameters (in red) and the corresponding 1D marginal prior (in black). The green dots refer to the correct reference values for the 7 model parameters. The orientation of the modelled UXO is set to Theta = 20°, Phi = -5°. This specific orientation and aspect ratio results in a model ambiguity, where many different UXO sizes and orientations fit the signal equally well. The model non-uniqueness is evident in the 1D marginals of Volume, Aspect ratio and Theta. The marginal distribution of the aspect ratio has a complex density with a large peak around 3.1, while the marginal of Volume and Theta has bimodal features with two distinctive peaks. It is clear, that the large model variability in this case, makes it hard to conclude on any specific UXO from the posterior PD. The variance can be attenuated by incorporating complex prior information on specific parameters, such as volume and aspect ratio, in order to reduce the amount of possible outcomes which would reduce model ambiguity.

The relative frequency of running the extended Metropolis algorithm and the Gibbs sampler in each iteration is manually chosen. We found that running the Gibbs sampler for every 500 extended Metropolis iterations, in general provided good results, given the problem at hand and a SN ratio of 6. We also found that the efficiency of the Gibbs sampler was best when sampling from a 2D conditional (as opposed to a 1D conditional) for specific sets of model parameters, the ones for which we expect a high degree of correlation.

The 1-D marginals of the posterior in Fig. 5(b) only tell part of the story and one should be careful when interpreting these distributions. For instance, if we simply choose the most probable aspect ratio it would correspond to the most likely aspect ratio at a particular orientation and volume, thus disregarding many possible solutions of different UXO that can fit the model equivalently. This was illustrated in the logL 2d marginal for Theta + Aspect ratio in Fig. 4(b) where correlated structures are evident. To illustrate this, 2-D sampling points are shown for selected parameters in Fig. 6(a) for the inversion run illustrated in Fig. 5. The ‘parabola’ structure of the posterior PD for Volume-Aspect ratio and Theta-Aspect ratio are easily recognized.

In order to quantify the behaviour of the chosen sampling algorithm, several tests were carried out using multi-chain parallel tempering, with an extended Metropolis algorithm. Both methods should sample the same posterior distribution, and differ only in numerical efficiency.
Figure 6. Derived 2D-marginals of the posterior density for example illustrated in Fig. 5. 1500 model realizations achieved with both sampling methods are shown. Blue: Sampled with Gibbs-within-metropolis. Green: Sampled with extended Metropolis using multi-chain parallel tempering. For convenience, only selected model parameters with the highest degree of correlation are plotted. It is evident that both sampling methods are sampling the same posterior distribution.

The parallel tempering was carried out using a seven-chain tempering with temperatures of $T = [1, 3, 5, 7, 9]$ and a jump frequency between modes of 0.1. The derived distribution is plotted in Fig. 6(b) and is quite similar to Fig. 6(a). The optimal sampling settings was found by trial and error, every time it failed sampling the proper density, we changed the sampling size and temperature of the chains, until an agreeable posterior density was derived. Sampling with parallel tempering was carried out over 500k iterations with a sample taken every 1000 iterations. During the 0.5 million iterations in parallel sampling 16675 swaps between the 7 chains were made. The same inversion carried out with Gibbs-within-Metropolis was made with 1.5 million extended Metropolis iterations with Gibbs sampling executed every 100 iteration, and a model sample taken every 1000 iterations.

Comparison of selected 1D-marginals derived from both algorithms are shown in Fig. 7. The top panel presents the results of Gibbs-within-Metropolis, whereas the bottom panel presents the results derived with parallel tempering. The 1D-marginal posterior distributions depicts the same similarity as is apparent in the 2D-marginals, however, some differences can be seen in the two posterior densities. The bimodal peak densities for Volume and Theta are approximately equal for the posterior derived with Gibbs-within-Metropolis, while the 1D-marginals derived with parallel tempering have modes that are unequal in frequency. This could indicate, that the collected models might not be independently sampled, which could explain the difference in the two derived posterior distributions.

Convergence of each McMC chain was assessed by inspection of the log-likelihood curve during the sampling phase. The log-likelihood for the inversions with parallel tempering and Gibbs-within-Metropolis are shown in Fig. 8, together with a run carried out with extended
Figure 7. Derived 1D-marginals of the posterior density for example illustrated in Fig. 5, achieved with different sampling methods. Blue: 1500 model realizations sampled with Gibbs-within-metropolis. Green: 5000 model realizations sampled with extended Metropolis using multi-chain parallel tempering. For convenience, only selected model parameters are plotted.

Metropolis only, for illustrative purposes. It can be seen how convergence is reached for parallel tempering and Gibbs-within-Metropolis during the first 2000-3000 iterations, while the extended Metropolis in this example never reaches convergence. The burn in time was set to 5000 iterations for all instances of using the extended Metropolis in order to ensure a proper step length, and the acceptance rate was set to 30% for all model parameters. In practice, the acceptance rate generally has a large variation for each parameter, since it is affected by the correlations between model parameters and the initial starting model in the extended Metropolis sampling. For example, the extended Metropolis chain shown in Fig. 8 (black) never reaches convergence, and thus, it must compensate for a very low acceptance rate by reducing the step length significantly. This results in a very low exploration rate, and hence a minimal chance of ever reaching convergence. This problem is mitigated by usage of either parallel tempering or Gibbs-within-Metropolis as long as the burn in time is set sufficiently high.

3.3 Results - Synthetic Data with Remanent magnetization

An example of an ordnance with remanent magnetization will be shown here. The position, orientation, volume and aspect ratio of the ordnance will be the same as in Fig. 5 but now significant remanent magnetization has been added to the source.

The addition of remanent magnetization increases the number of model parameters to ten. Hence, an increase in iterations is necessary to explore the posterior distribution. The sampling length is increased to 10 million iterations, and the Gibbs sampler is set to sample in 2D space of model parameters from sets of two of the following parameters: Aspect ratio, Volume, Theta, Phi, $M_{\text{rem}}$, $M_{\text{incl}}$, $M_{\text{decl}}$ every 500th iterations of the extended Metropolis. The spatial setup and noise model is the same as previous examples.
The remanent signal complicates the inversion process. It causes additional non-uniqueness and correlations between the remanent and induced signal. In many cases, a change in orientation, size or shape of the UXO can be accounted for by changing the remanent magnetization. In the worst case, the remanent magnetization will be aligned with the induced dipole moment or the ambient field, so the addition of the remanent signal only will cause a slight change in the direction of the dipole moment. In such cases, the change due to remanent magnetization can be accounted for by a change in the UXO parameters. However, if the remanent magnetization is perpendicular to the induced dipole moment, it would cause a change in the direction of the dipole moment that can not be resolved by the induced model alone, which would reduce the correlation between the induced and remnant magnetization. In the following, we show one synthetic example with remanent magnetization. We focus upon the case, where the remanent magnetization is aligned with the ambient field, since this is where we expect the highest degree of correlation.

### 3.3.1 Remanent magnetization aligned with ambient field

Remanent magnetization aligned with the ambient field mostly affects the amplitude of the induced signal, which at most times, can be accounted for by a change in aspect ratio, orientation and/or volume. This is illustrated by the increased spread of the posterior PD for both volume, aspect ratio and orientation in Fig. 9. The distribution for the volume is generally higher than the true value, however, it has sampled in the whole region of $M_{rem}$, meaning that it has included the possibilities of an ordnance with up to 130% $M_{rem}$ in the assessment of the ordnance size. Regarding aspect ratio it is apparent how the remanence complicates any procedure of discriminating between ordnance by the aspect ratio of the magnetized spheroid. 2D-marginals of the posterior distribution is seen in Fig. 10 showing interesting correlated structures between many parameters, in particular, between $M_{rem}$, $M_{decl}$, and Volume, Aspect ratio, and Phi. It is apparent, how the estimated parameters regarding ordnance orientation are no longer well resolved in the inversion, compared to an inversion
Figure 9. Remanent magnetization aligned with ambient field: a) Anomaly generated with forward model. Contour lines are plotted in 0.5 nT intervals b) Prior (black) and posterior (red) distributions for inversion on induced and remanent magnetization on synthetic data set. Total number of iterations: 10 million hybrid Metropolis-Gibbs iterations. Reference model used to generate the anomaly is indicated by the green circle.

with no remanent magnetization, as in Fig. 5. This is a consequence of inverting for both the induced and remanent magnetization, where a change in the induced dipole moment can now be accounted for, to some extent, by a change in the induced magnetization.
22 M. D. Wigh, T. M. Hansen and A. Døssing.

Figure 10. Remanent magnetization aligned with ambient field: 2D Marginals for inversion results on the generated anomaly shown in Fig. 9. The red dot shows the reference model value used to generate the anomaly. Total number of iterations: 10 million Gibbs-within-Metropolis iterations.

4 REAL DATA INVERSION

The inversion scheme for induced and induced + remanent magnetization will be carried out on a real example, where the magnetic anomaly from a 155 mm projectile shell has been measured. The ordnance was buried ≈ 0.8 – 1 m beneath the ground surface in a large old agricultural field consisting of glacial moraine deposits. A picture of the UXO is seen in Fig. 11.

The measurements were taken with two scalar magnetometers of the type GEM GSM-19 High Precision Overhauser Magnetometers with an absolute accuracy of ± 0.1 nT. The magnetometers were used to measure the total field and vertical gradient, while a third was used as base-station to correct for diurnal variations during the survey. The two survey magnetometers were mounted on a backpack with a vertical separation of 0.6 m functioning as a walking magnetometer + vertical gradiometer. Measurements were taken with a frequency of 5 Hz along the survey direction, while the survey lines were separated with 0.5 m close to the source, and 1 m further away from the source. Survey lines were taken in approximately N-S direction. The processed total field data are illustrated in Fig. 12 with a red box indicating the selected data for inversion. Only the area where the signal strength from the ordnance is highest will be selected, in order to minimize the correlated noise from other sources, while still obtaining sufficient information on the characteristics of the dipole anomaly. Processing of the data set involved correction of diurnal variation using the base station, applying a running mean low pass filter and removal of the main

Figure 11. 155 mm grenade buried at a depth of approx. 80-100 cm. The arrow points towards geomagnetic north, so its position (with respect to the magnetic sensor) and orientation can be approximated to (Depth,Theta,Phi)≈ (2.8 m, 0°, 35°). UXO properties : Aspect ratio ≈ 3.8, Volume ≈ 0.004 m³ steel (30 kg)
Figure 12. Contour plot of processed magnetic data for the survey on a 155mm projectile. Contour lines are shown in 1 nT intervals and the white dot is the measured GPS position of the buried UXO. The red box indicates the data used for inversion - Measurements outside of the box contains only a weak signal from the ordnance and stronger signals from other magnetic sources that gives unwanted correlated noise structures. This is apparent on the left side of the anomaly, where there is a clear unwanted signal with a peak of approx 5 nT.

First, an inversion of the induced signal alone was carried out on the total field data. Next, an inversion was made on the same data assuming both induced and remanent magnetization. 1D Marginals of the posterior pdf with induced parameters only is shown in Fig. 13. It is noticeable how well resolved the depth and orientation are and that they are in close range of the expected values. The estimated posterior volume is higher than the expected value, with an posterior mean of approximately 90 kg, which is 60 kg higher than measured reference value. This is also the case for the aspect ratio, where the posterior distribution is centered far from the expected value. The latter can be due to one or more of the following reasons; (i) The UXO has remanent magnetization that cannot be ignored. (ii) There is near surface geology causing magnetic disturbances interfering with the signal of the ordnance. (iii) Accumulation of noise/error along the survey direction. Since the aspect ratio defines the ratio of magnitude of the magnetization between the axial and radial component of the spheroid it is unlikely that the change of aspect ratio is due to errors in the sampled data. Hence, significant correlated noise present in the agricultural field inferring with the measured anomaly, or, remanent magnetization in the ordnance, is believed to be the cause, which could explain the offset in the estimated volume and aspect ratio.
The posterior PD for the second inversion on the 155mm projectile with induced and remanent magnetization is seen in Fig. 14. The setup of the sampling algorithm is similar to the setup used for the synthetic examples, but with a higher number of iterations that is increased to 20 million. As depicted in the 1D marginals, the posterior volume has a clear peak at the expected value and a gradually decreasing probability for higher volumes. The variance of the volume, however, is higher than the synthetic example shown in Fig. 9. The posterior depth is well resolved and have a peak density close (within 20 cm) to the estimate for the inversion with induced magnetization only, which both agrees well with the expected value. Regarding aspect ratio and $\theta$ the posterior is similar to the prior indicating that no information can be extracted with respect to these components, when significant remanent magnetization is assumed to be present. The posterior distribution for $\Phi$ has a peak density close to the expected value, but the increased variance due to the remanent magnetization is also apparent here.

The belief that there is significant remanent magnetization in this UXO agrees with the estimated $M_{\text{rem}}$ being at least above $Q = 0.2$, which can explain the offset in the parameters of volume and aspect ratio estimated, when inverting on induced alone as shown in Fig. 13.

Structures of the posterior PD is illustrated in Fig. 15 with 2D marginals on the model parameters of interest. Some features can be derived from the 2D-marginals, for instance, models accepted with the highest volume possible is only accepted when the remanent magnetization is highest and anti-aligned with the ambient field. This reduces the magnitude of the combined magnetization, such that it is
Probabilistic Inversion of UXO

Figure 14. Inversion 155mm projectile induced + remanent. Prior and posterior distributions for inversion on induced and remanent magnetization. 20 million Gibbs-within-Metropolis iterations

equivalent to the induced field of a small ordnance with no remanence. The probability of this being the case in a real situation is unlikely, however it currently cannot be excluded from the ensemble of accepted models.

5 DISCUSSION

We have tested a hybrid sampling algorithm that combines two well known algorithms in a simple manner, to solve complex highly non-linear problems in a probabilistic procedure. The algorithm provides us with means for assessing the uncertainties in magnetic inversion regarding (x,y,z)-position, shape, and size of the UXO, even when significant remanent magnetization is present in the ordnance. While the results presented indicate that the inversion problem can be addressed and properly solved within a probabilistic framework, certain aspects and issues arise that will be discussed below.

5.1 Probabilistic approach to UXO inversion

By adopting a probabilistic approach in UXO inversion we have the ability to achieve further insight into the complex problem of modelling magnetic UXO sources. We have shown that, by using a suitable sampler such as the Gibbs-within-Metropolis sampler, we are capable of solving the dipole equation in non-unique situations. In practice, we can gain insight into a possible broad spectrum of solutions, even
Figure 15. 2D Marginals for UXO 12 155mm projectile induced + remanent magnetization. 20 million Gibbs-Within-Metropolis iterations. The Gibbs sampler perturbs \( m = [2 \ 3 \ 5 \ 8 \ 9 \ 10] \), while the extended Metropolis perturbs all model parameters evenly

when only minimum prior information, such as seabed constraints, is available. Any prior constraints available will, however, decrease the inversion complexity, but contrary to other inversion approaches we do not need a library of ordnance prior to an inversion. We can sample the full non-linear solution with very few prior constraints that are not tuned specifically.

Other UXO inversion methods concerning magnetometry (e.g. (Billings 2006)) use fixed model parameters in a library of different ordnance of which they know the aspect ratio and volume of every ordnance, thus simplifying the inversion process significantly. In such cases they do not have to consider the variations in aspect ratio and volume and the possible correlations associated to these parameters.

Attempts made to explore the full posterior solution, in the aspect of EMI inversion, have only been done before by simplified means to a full exploration. Optimization algorithms used (e.g. (Billings et al. 2010b; Pasion et al. 2008)) or sampling (Beran et al. 2011b), all address the problem of local minima by executing the algorithm on a range of different starting points. These starting points will be varied for model parameters with high correlation to ensure the algorithm does not get stuck in a local minima. This can often be a laborious and time consuming process, which is avoided when sampling the full posterior PD. The full posterior PD, however, is often complex and can not be approximated by Gaussian distributions, which makes it hard to quantify in a statistical classification scheme. Interpretation of the full posterior PD is out of the scope of this paper and will be addressed as future work.
5.2 Inversion of synthetic and real data

It has been shown throughout the study, how we are capable of dealing with UXO inversion with increasing difficulties and the problems that arise with it. In the synthetic studies, we demonstrated that, by using our proposed hybrid sampling algorithm, we can solve a case with induced magnetization, even when several solutions exist due to the ambiguity in the spheroidal forward model. The inversion complexity was increased significantly by adding remanent magnetization to the synthetic source and inverting for this, while fitting the induced signal as well. At the cost of extra computation time, we are able to sample the solution when dealing with induced and remanent magnetization together, but it is clear that the inversion problem becomes highly non-unique and undoubtedly is subject to further research. Some specific points regarding inversion with combined induced and remanent magnetization are achieved from this study. (i) modelling the strength of remanent magnetization in terms of percentage of the induced dipole moment indicates a feasible method to gain sufficient information on presence of remanent magnetization, while being able to constrain the problem, so that it can be sampled in a finite amount of time. (ii) It is in most cases impossible to distinguish between the induced and remanent dipole moment, when modelling a point dipole without no hard prior constraints on the orientation, shape, or size of the UXO. However, there are cases when the forward model cannot fit the anomaly with the induced signal alone, giving some resolution between estimated remanent and induced magnetization, which will be evident in the posterior distribution.

In the synthetic studies it was demonstrated, that we can sufficiently solve the inversion problem, given good data with uncorrelated noise and knowledge about the noise level in the data. The noise added in the synthetic data sets consisted of uncorrelated Gaussian noise, but in a real magnetic data survey the noise is generally spatially correlated, thus complicating the problem significantly. Furthermore, we do not have much knowledge on the noise present and can only give a reasoned approximation. When inverting on the real data case we assume uncorrelated noise, but defining this noise floor can be hard and may have large impact on the resulting posterior model. If we set the assumed noise level too low the algorithm is likely to fit the noise and correlated unwanted signals, which may hinder the sampling algorithm in full exploration of the posterior PD, while also obtaining false models. However, if we assume the noise to be significantly higher, we can reduce the chances of fitting unwanted correlated signals and simplify the sampling process, but at the cost of posterior model resolution.

5.3 Correlated noise

The aspect of noise present in magnetic data and the type of noise model used has not been a high priority in this study, but remains to be a subject to further research. Throughout the study we have only tested and assumed uncorrelated Gaussian noise, with a relatively high signal to noise ratios. In practice however, significant amount of correlated noise is expected to be present. In real surveys the presence of spatially correlated noise may be due to other magnetic sources, such as metallic debris, or near surface geology features. The measured
signal will be a superposition of all sources, for instance, if there are two sources close to each other, it will result in overlapping anomalies, and if they are similar in signal wavelength it will be difficult to distinguish between the two sources. Similarly, if there are geologic features causing magnetic signals with a spatial frequency comparable to the UXO, it can affect the inversion process so that the optimal models will be an inaccurate answer that compensates by fitting unwanted magnetic signals. The presence of correlated noise is evident in the real data example, as shown in Fig. 12, where we see magnitudes of up to 5 nT in the outer edges of the survey (that are clearly not originating from the UXO source). Hence, the uncorrelated instrumental noise with a standard deviation of 0.1 nT is generally insignificant, when dealing with magnetic anomalies of at least a few nT, and there is near-surface geology that distorts the anomaly with unwanted magnetic signals.

In our real data study case, however, we have information on the true model parameters due to the experimental setup of the UXO, so we can adjust uncorrelated noise by trial-and-error, until we obtain inversion results that are within the range of the expected model values, and have the best resolution. The most reasonable results were obtained at SN=5.5, which is equivalent to a noise standard deviation of 1.2 nT. This is significantly higher than the instrumental noise with a standard deviation of 0.1 nT, as we would expect, due to other unwanted magnetic signals that result in spatial correlated noise. It must be noted, however, that the optimal uncorrelated noise model of 1.2 nT standard deviation is probably far from reality, but the derived posterior results agree well with the expected. If the correlated noise can be modelled and quantified, for example as a modelling error, through a probability distribution, it may be possible to account for it (see e.g. (Cordua et al. 2008; Hansen et al. 2014; Köpke et al. 2018) for discussion on this). We leave this analysis for future research. In a real survey case, where the expected values are unknown and the noise is unknown as well, then the optimal noise floor cannot be determined by trial-and-error. A common approach to determine the noise level, can be made by hierarchical Bayes, where parameters such as the properties related to the noise model, can be inferred as part of the inversion. For example, if the noise model is assumed to be an uncorrelated Gaussian model, with a fixed variance, the variance of the noise model can be treated as a hyperparameter with a specified hyperprior, whose properties can be inferred as part of inversion. (see e.g. (Gelman et al. 2014) for further details). However, if the noise is in fact a type of correlated noise, then one need to know the shape of the correlated noise in order to infer information about the magnitude of the noise. (Madsen et al. 2017) demonstrated that inferring information about the magnitude of an assumed uncorrelated noise model, when in fact the noise is correlated, can lead to severe underestimation of the noise level, and hence this leads to fitting of noise as data, as part of the inversion. Therefore, before inverting for the noise level using for example hierarchical Bayes, we suggest to first establish and quantify a correlated noise model. This is left as a part for future work.

5.4 Remanent magnetization and model ambiguities

One of the remaining issues is the capability to verify the modelling of the induced dipole moment of large WWII UXO by compact spheroidal objects. This is due to the presence of remanent magnetization of which it is hard to gain any statistical insight, since it depends
on the objects history, and can generally be assumed as being independent of size and shape of the UXO. (Billings & Youmans 2006; Billings 2009) and (Butler et al. 2012) have made several detailed studies on the spheroidal forward model and shown acceptable modelling errors for smaller UXO, while also measuring the actual remanent magnetization on different UXO. (Billings 2009) achieved this by using a measurement platform consisting of six three-axis fluxgate magnetometers symmetrically distributed around a rotating sample holder. With enough samples of the UXO at different orientations, it is possible to distinguish between the induced and remanent magnetization, but such platform will in practice have a limitation on UXO size. Here, we focus on large (>25 kg) UXO common in the European regime.

Determining the remanent magnetization pattern of a large UXO is difficult and currently no attempts have been made for this, since we cannot distinguish between induced and remanent magnetization without a thorough survey of the UXO itself. This complicates any justification of the forward model in modelling the true induced dipole moment of larger UXO. Currently, the majority of the misfit when fitting the anomaly with induced magnetization is explained by the presence of remanent magnetization. Hence, it is of high interest to measure the effective susceptibilities of larger UXO, in order to validate the spheroidal model for larger sources, while also gaining more knowledge on the presence of remanent magnetization. Even if we are not capable of distinguishing properly between the induced and remanent magnetization, we can assess all possible solutions of UXO with the variations due to remanent magnetization included. This means, that we can start to consider a discrimination procedure of UXO and non-UXO objects, that takes into account the uncertainties related to UXO discrimination, when significant remanent magnetization can be present.

5.5 Sampling of the \textit{a posteriori} distribution

Sampling the problem for both induced and remanent magnetization, in order to achieve a deeper understanding of the uncertainties in classification of a UXO by its magnetic signature, is a difficult, highly non-linear, and non-unique problem. The high correlation between parameters makes it hard to quantify statistical dependency of proposed models in a probabilistic framework by means of autocorrelation. The underlying theory is, that enough samples in the posterior PD should be statistically independent, when interpreting the posterior probabilities, in order to ensure that the posterior distribution represents statistics of independent realizations. For instance, if we have 2000 saved models but only 1 independent realization that can be analogous as to say, we only have found 1 independent UXO to fit the signal, but varied its orientation or remanence 2000 times. Assuring that we have generated independent realizations is usually done by calculating the autocorrelation between the proposed model and the previously selected models, to ensure that the next proposed model is randomly chosen, independently from the previous sampled model. This notion of distance and movement within a high dimensional non-linear space of correlated parameters is, perhaps, not a sufficient way of analyzing the posterior statistics. The Gibbs-within-Metropolis sampler forces the sampling algorithm to jump around in the model space, hereby bypassing the fixed step length of the extended Metropolis, which should in practice reduce the autocorrelation of parameters. The autocorrelation of a series of posterior model parameters reveal, that most model
parameters have a very high autocorrelation. This is especially true when inverting for remanent magnetization. As an example, the autocorrelation time of the \( \Theta \) parameter when inverting for remanent and induced magnetization is generally high. However, this does not seem to be consistent, but appears to depend on the actual path taken in the sampling process. Thus, to avoid analyzing the autocorrelation time of all model parameters during a sampling, in order to justify the statistical independence of the posterior models, we instead run the same inversion several times with an identical sampling setup and compare the posterior PD (Not shown here). This is also done to ensure convergence to a stationary posterior distribution.

In terms of efficiency the algorithm converges to the region of interest usually within a few thousand iterations that can be achieved within a minute on a modern day CPU. It is, however, time consuming in terms of fully exploring the posterior distribution, in order to obtain many independent realizations to infer information on the anomaly and all possible models of UXO that can fit the signal equally. The proposed algorithm was compared to a multi-chain MCMC approach, parallel tempering in combination with an extended Metropolis algorithm, that was run on the synthetic data case as illustrated in Figs 6 and 7. It was possible to get good estimates of the posterior PD using both methods, but each method has its own advantages and disadvantages. Using parallel tempering we are, in a quite simple manner, capable of sampling a complex problem with no or little knowledge into the nature of the problem, such as correlation between parameters. By scaling the temperature we allow for a change in exploration rate for the multiple chains, however, selecting the optimal scaling temperature is often hard, and if set too high or too low, it reduces the efficiency. Furthermore, there were problems with sampling the posterior density robustly using the setup with parallel tempering. It was not each time it succeeded, indicating that each chain could get stuck, in which case means, that one might just as well run 7 independent single-chains.

The Metropolis-within-Gibbs has some different aspects. One should have a decent degree of knowledge about the nature of the problem in order to sample efficiently. Gibbs sampling should be carried out on model parameters that have high correlation, in order to optimize the sampling efficiency. When Gibbs sampling is executed on the right parameters with high correlation and/or multimodal likelihood, it shows good capabilities of exploring as a single-chain MCMC approach.

There are other recent examples of sampling algorithms that utilize multi-chain sampling, in particular the DREAM algorithm (Vrugt et al. 2008; Vrugt 2016). This is a more advanced algorithm with more steps as part in the sampling process, such as outlier chain correction and subspace sampling of chains to increase exploration in between nodes (which can draw some similarities to a Gibbs step in Gibbs-within-Metropolis, when we want to jump in between nodes). This approach is dependent on utilizing many independent chains in the sampling process, whereas the Gibbs-within-Metropolis works as a simple single-chain sampler, that is made by combining two well known sampling algorithms.

We have discussed a number of sampling methods (Metropolis et al. 1953; Hastings 1970; Geman & Geman 1984; Liu et al. 2000; Neal et al. 2011; Martin et al. 2012; Sambrigde 2013) that can all in principle be used to solve the inverse problem of sampling from the posterior
distribution as defined above. They differ in their numerical efficiency and in the ease of use; some methods require more tuning parameters to make the sampling run optimally than others. Here, the key problem has been sampling the posterior probability distributions related to probabilistic inversion of UXO data. We have found that the use of the Gibbs-Within-Metropolis sampling algorithm provides robust inversion in reasonable time, with little to no manual selection of hyperparameters. Therefore, it is our intention to make use of this algorithm for similar problems in the future. We do not claim that our algorithm is superior in efficiency compared to state-of-the-art multi-chain samplers, that properly utilize parallel usage of CPUs. This is left as future work, to compare these sampling algorithms in efficiency.

There should be no difficulty in adapting this sampling algorithm to other subjects within the field of probabilistic inversion. Anyone familiar with extended Metropolis and Gibbs sampling can implement this in a simple manner. In order to optimize it, however, one should have sufficient knowledge on the nature of the problem and the parameterization used prior to the sampling phase. In such cases, one can make an analysis of the likelihood or, for instance, use a simple ‘naïve’ parallel tempering approach with extended Metropolis to achieve a first impression of the posterior distribution, which can then be used to configure and optimize the Gibbs-within-Metropolis algorithm. In any general case, where multi-modality and strong correlations in 2D marginals of a high dimensional distribution are present, then Gibbs-within-Metropolis should be applicable as a single-chain sampler. However, there should be no problem in utilizing Gibbs-within-Metropolis in a multi-chain approach, which is also of high interest for future work.

6 CONCLUSION

In this paper, we have investigated a probabilistic approach to UXO magnetometry inversion. It has been successfully tested on a full-scale inversion of UXO by interpreting location, size, shape, and orientation from the measured anomaly. Such inversion procedures are usually taken in different iterative steps, or by utilizing multi-chain McMC approaches. We have tested a single-chain hybrid algorithm of the extended Metropolis and Gibbs samplers, which has, compared to a normal extended Metropolis with or without parallel tempering, shown to be a good alternative to sample the posterior distribution. Our method presented here is not sensitive to a selected starting point, and does not have to be repeated iteratively in order to avoid local minima, which can be frequent in a highly non-linear model space. By sampling the full posterior distribution we assess all possible outcomes with respect to the forward model, noise model and prior constraints. Thus, it is a viable method when a thorough analysis of the uncertainties related in UXO discrimination and classification is needed. We have shown, that we can account for the presence of remanent magnetization and the forward model ambiguity, when using a probabilistic inversion approach.
REFERENCES


Aliamiri, Alireza & Miller, Eric 2007, Random Walk/Markov Chain Model for Sensor Positional Uncertainty with Application to UXO Discrimination IEEE international geoscience and remote sensing symposium


Hastings, W Keith 1970, Monte Carlo sampling methods using Markov chains and their applications, *Oxford University Press*


Mosegaard, Klaus. & Sambrige, Malcolm. 2002, Monte Carlo analysis of inverse problems, *Inverse problems* 18(3) pp. 29-54

Mosegaard, Klaus. & Tarantola, Albert, 1995, Monte Carlo sampling of solutions to inverse problems, *Journal of geophysical research* 100(B7) pp. 12431-12447


M. D. Wigh, T. M. Hansen and A. Døssing.


Stalnaker, Jack. & Miller, Eric. 2007, Particle swarm optimization as an inversion tool for a nonlinear UXO model, *IEEE International geoscience and remote sensing symposium*


Tarantola, Albert. and Valette, Bernard. 1982a, Inverse problems = quest for information, *Journal of geophysics* 50(3) pp. 150-170


Jasper A. Vrugt, . , Volume 75, 2016, .

