Bayesian Joint Inversions for the Exploration of Earth Resources

Alistair Reid[1], **Simon O'Callaghan**[1], **Edwin Bonilla**[1], **Lachlan McCalman**[1], **Tim Rawling**[2] and **Fabio Ramos**[3]

1. NICTA, 2. University of Melbourne, 3. University of Sydney

{ alistair.reid, simon.ocallaghan, edwin.bonilla, lachlan.mccalman }@nicta.com.au

tim.rawling@unimelb.edu.au, f.ramos@acfr.usyd.edu.au

Abstract

We propose a machine learning approach to geophysical inversion problems for the exploration of earth resources. Our approach is based on nonparametric Bayesian methods, in particular, Gaussian processes, and provides a full distribution over the predicted geophysical properties whilst enabling the incorporation of data from different modalities. We assess our method both qualitatively and quantitatively using a real dataset from South Australia containing gravity and drill hole data and through simulated experiments involving gravity, drill hole and magnetics, with the goal of characterizing rock densities. We show that our approach is more effective than the usual inversions in the geophysics community. The significance of probabilistic inversion extends to general exploration problems and can dramatically benefit the industry.

1 Introduction

The discovery of resources in the Earth's crust is a problem of crucial importance for continued global economic development, so modern investigation methods make use of vast amounts of geophysical data to improve the efficiency of resource exploration. In a geological *inversion* problem, properties such as temperature, conductivity, density, magnetic susceptibility, and permeability are inferred from related observations such as gravity, magnetics and seismic reflexion.

Two components are crucial in a geophysical inversion: the assessment of uncertainty over the predicted geophysical properties, and the joint reasoning from multiple sources of information. Uncertainty allows for principled decision-theoretic approaches to minimize the risk of acquiring more measurements. Joint reasoning through the estimation of statistical dependencies provides a natural mechanism to fuse many of the available data modalities.

In this paper we formulate geophysical inversion as a machine learning problem, and propose an approach that naturally provides a full distribution over the inverted quantities and enables stochastic fusion of different types of observations. Our approach is based on nonparametric Bayesian methods, in particular, Gaussian processes. We apply our method to a real dataset from South Australia containing gravity and drill hole data with the goal of characterizing rock densities for geothermal target exploration, and also to simulated validation data involving gravity, drill hole and magnetic observations. Our approach applies to general resource exploration problems where it could have a dramatic impact on the industry.

2 Related Work

One of the most popular methods to solve inversion problems in Geophysics is the UBC software developed at the University of British Columbia — Geophysical Inversion Facility, based on the work by Li and Oldenburg [1996] and Li and Oldenburg [1998]. This software is widely used in the industry and in academia, in particular for gravity inversions, magnetic inversions and mineral exploration [Oldenburg *et al.*, 1998]. These methods have the advantage of being relatively easy to understand and very fast to obtain an initial solution to an unconstrained inversion problem. However, one of the major drawbacks of these techniques is that they do not provide uncertainty estimates for the properties of interest.

General inverse problems have also been studied in the machine learning community [Carreira-Perpiñan, 2001], and the gravitational inverse problem has been discussed by Yurtsever *et al.* [2011]. From the geostatistics community, a comprehensive overview of stochastic process priors for inverse problems can be found in Tarantola [2005]. However, these methods seem to be underexploited in the mineral exploration area. Gaussian processes are, in fact, a Bayesian formulation of the nonparametric priors which have been used in Geostatistics for interpolation and regression problems under the name of Kriging, see e.g. Cressie [1993] and Stein [1999]. However, extensions of standard regression approaches are required for an inversion problem because the observations are indirectly related to the property of interest.

The importance of the joint analysis of multiple data types in realistic 3D geological modeling has been studied, for example, by Guillen *et al.* [2008] and Fullagar and Pears [2007]. Very recent work by Shamsipour *et al.* [2012] has carried out stochastic inversions of gravity and magnetic data. The main advantage of our approach is that we provide a consistent probabilistic model, where we can use principled marginal likelihood objectives for model selection and leverage all the machinery developed in the machine learning community in recent years to perform inference with large datasets.

3 Gaussian Processes for Joint Inversions

In this section we describe our approach to the problem of geophysical inversion. We start by introducing Gaussian process priors in standard regression settings. The formulation is then extended to inversion problems for which linear forward models exist. We then detail how our modeling approach can naturally be extended to multi-task settings. Finally, tractable implementation of this potentially costly model is discussed.

3.1 Gaussian Processes for Regression

Gaussian processes are flexible nonparametric priors that have been used successfully in various machine learning tasks such as regression and classification (see e.g. Rasmussen and Williams [2006]). The probability distribution of a function $f(\mathbf{x})$ is a Gaussian process (GP) if for any finite subset of points $\mathbf{x}_1, \ldots, \mathbf{x}_N$, the function values $f(\mathbf{x}_1), \ldots, f(\mathbf{x}_N)$ follow a Gaussian distribution. We denote a Gaussian process with:

$$f(\mathbf{x}) \sim \mathcal{GP}(\mu(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')), \tag{1}$$

$$\mu(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})], \tag{2}$$

$$k(\mathbf{x}, \mathbf{x}') = \mathbb{E}[(f(\mathbf{x}) - \mu(\mathbf{x}))(f(\mathbf{x}') - \mu(\mathbf{x}'))], \tag{3}$$

where $\mu(\mathbf{x})$ and $k(\mathbf{x}, \mathbf{x}')$ are the mean function and the covariance function of the GP respectively. A Gaussian process is completely determined by its mean function and its covariance function. In a regression setting it is customary to assume a GP prior over the latent functions $f(\mathbf{x})$ and a likelihood model where these latent functions are corrupted by Gaussian noise. In this case, the predictive distribution at a test point $\mathbf{x}_*$ is Gaussian with simple analytical forms for its mean and variance.

The parameters of the covariance function and the parameters of the noise process are usually referred to as the *hyper-parameters* of the GP. These can be learned from data by optimization of the marginal likelihood. Under the GP prior and the Gaussian likelihood assumptions, it is possible to marginalize the latent functions analytically, and the resulting marginal likelihood is also a Gaussian distribution.

3.2 Gaussian Processes for Inversion

Here we focus on inversion problems for which a linear forward model that relates the parameter of interest (i.e. geophysical property) to the observations can be formulated. For explanatory purposes, we take the problem of gravity inversion as a running example. In this case, it is well known that the density of a body at a particular location and the observed vertical component of the gravitational field are related via an integral operator (see e.g. Kearey *et al.* [2002], Ch 6).

Additionally, as in many popular geophysical inversion approaches, we work upon a discretized version of the forward model, where the 3-dimensional region of interest has been partitioned into cells, each having a constant value of the parameter of interest. In the case of our gravity example, physical principles imply that the vertical component of the gravitational field measured by a sensor is a linear combination of the contributions of all the cells as a function of their positions and densities. For details on how these physical models over voxel cells are obtained see Li and Oldenburg [1998].

Problem Formulation

Let $\{\phi^{(j)}\}$ denote the unknown values of the geophysical parameter at locations $\{\mathbf{z}^{(j)}\}_{j=1}^{M}$, and let $\{y^{(i)}\}$ be the related observations at the locations $\{\mathbf{x}^{(i)}\}_{i=1}^{N}$, with $\mathbf{x}, \mathbf{z} \in \mathbb{R}^3$. Our goal is to reason about the geophysical parameter in the region of interest $(\{\phi^{(j)}\})$ given the related observations $(\{y^{(i)}\})$. In the gravity inversion problem, $\{y^{(i)}\}$ are measurements of the variations of the vertical component of the gravitational field and $\{\phi^{(j)}\}$ are the values of the density of the anomalous body responsible for these variations.

Forward Model

Assuming a linear forward model we have that:

$$\mathbf{f} = \mathbf{G}\boldsymbol{\phi}, \tag{4}$$

where $\mathbf{f}$ is the vector of noiseless observations; $\boldsymbol{\phi}$ is the vector of unknown parameters at the 3D locations; $\mathbf{G}$ is a known $N \times M$ *sensitivity* matrix that relates the values of the geophysical property at different locations to the observations. For a gravity forward model the values of the matrix $\mathbf{G}$ can be determined analytically, assuming simple shapes such as prisms (see e.g. Nagy *et al.* [2000]). In particular, for a prism determined by its original and end coordinates $\mathbf{z}^o$ and $\mathbf{z}^e$, we can compute the corresponding sensitivity at location $\mathbf{x}$ by using:

$$G_{\mathbf{x},\mathbf{z}} = \gamma g(\mathbf{z}) \Big|_{\bar{z}_1^o}^{\bar{z}_1^e} \Big|_{\bar{z}_2^o}^{\bar{z}_2^e} \Big|_{\bar{z}_3^o}^{\bar{z}_3^e}, \tag{5}$$

where γ is the gravitational constant;

$$g(\mathbf{z}) = z_1 \log(z_2 + r) + z_2 \log(z_1 + r) - z_3 \arctan \frac{z_1 z_2}{z_3 r}; \tag{6}$$

$\bar{\mathbf{z}}^o = \mathbf{z}^o - \mathbf{x}$; $\bar{\mathbf{z}}^e = \mathbf{z}^e - \mathbf{x}$; and $r = \sqrt{z_1^2 + z_2^2 + z_3^2}$.

Prior and Likelihood Models

In this work we assume a Gaussian process prior over the rock properties of interest $\boldsymbol{\phi}$ and an isotropic likelihood model:

$$\phi(\mathbf{z}) \sim \mathcal{GP}(0, \kappa_\phi(\mathbf{z}, \mathbf{z}')), \tag{7}$$

$$\mathbf{y} = \mathbf{f} + \boldsymbol{\eta} \quad \text{with} \tag{8}$$

$$\boldsymbol{\eta} \sim \mathcal{N}(\boldsymbol{\eta} | \mathbf{0}, \sigma^2 \mathbf{I}), \tag{9}$$

where $\kappa_\phi(\cdot, \cdot)$ is the covariance function in the parameter space and σ^2 is the noise variance. The GP prior over the functions ϕ translates into a Gaussian prior over the geophysical quantities at the locations of interest:

$$\boldsymbol{\phi} \sim \mathcal{N}(\boldsymbol{\phi} | \mathbf{0}, \mathbf{K}_{\phi\phi}), \tag{10}$$

where $\mathbf{K}_{\phi\phi}$ is the $M \times M$ covariance matrix obtained by evaluating the kernel $\kappa_\phi(\cdot, \cdot)$ at the locations of interest $\{\mathbf{z}^{(j)}\}_{j=1}^{M}$.

Inversion

Inversion within a Bayesian framework is straightforward as it corresponds to computing the posterior distribution of the parameters of interest $\boldsymbol{\phi}$ given our observations $\mathbf{y}$. This is easily obtained by conditioning, where we need to compute the covariance structures:

$$\mathbf{K}_{y\phi} = \mathbf{G}\mathbf{K}_{\phi\phi}, \tag{11}$$

$$\mathbf{K}_{yy} = \mathbf{G}\mathbf{K}_{\phi\phi}\mathbf{G}^T + \sigma^2 \mathbf{I}. \tag{12}$$

Hence we obtain that the predictive distribution is given by:

$$\boldsymbol{\phi}|\mathbf{G},\mathbf{y} \sim \mathcal{N}(\boldsymbol{\phi}|\boldsymbol{\mu}_{\phi|y},\boldsymbol{\Sigma}_{\phi|y}) \quad \text{with:} \tag{13}$$

$$\boldsymbol{\mu}_{\phi|y} = \mathbf{K}_{y\phi}^T\mathbf{K}_{yy}^{-1}\mathbf{y} \tag{14}$$

$$\boldsymbol{\Sigma}_{\phi|y} = \mathbf{K}_{\phi\phi} - \mathbf{K}_{y\phi}^T\mathbf{K}_{yy}^{-1}\mathbf{K}_{y\phi}. \tag{15}$$

3.3 Multi-task settings

Our ultimate goal in addressing geophysical inversion problems is to fuse different data sources within a single probabilistic framework. As an initial step, we have adopted a multi-task learning approach based on the model of Bonilla *et al.* [2008]. This model assumes that the covariance between two different geophysical quantities (tasks) at two distinct locations decomposes as:

$$\kappa(\chi(\mathbf{x}),\phi(\mathbf{x}')) = k_{\chi\phi}\kappa(\mathbf{x},\mathbf{x}'), \tag{16}$$

where χ and ϕ denote distinct geophysical quantities, such as density and magnetic susceptibility, and $k_{\chi\phi}$ is the covariance between these quantities. This means that instead of having uncorrelated priors with distinct $\kappa_\phi(\cdot,\cdot)$ and $\kappa_\chi(\cdot,\cdot)$ as in Equation (7), we share "statistical strengh" by using the correlated prior in Equation (16). Despite the simplicity, our experiments show that, under certain assumptions on the dependency of the quantities of interest, this model performs well, and observing data from an additional source does improve the quality of the inversion results, especially when the observations are very sparse.

3.4 Hyper-parameter Learning

Let the set of hyper-parameters $\boldsymbol{\theta}$ include the parameters of the covariance function $\kappa_\phi(\cdot,\cdot)$ and the noise variance σ^2. We learn these hyper-parameters by maximization of the log marginal likelihood. As in the regression case, we can integrate out $\boldsymbol{\phi}$ analytically and obtain the marginal likelihood:

$$\mathbf{y}|\mathbf{G},\boldsymbol{\theta} \sim \mathcal{N}(\mathbf{y}|\mathbf{0},\mathbf{G}\mathbf{K}_{\phi\phi}\mathbf{G}^T + \sigma^2\mathbf{I}), \tag{17}$$

where, for notational simplicity, we have omitted the implicit dependency of $\mathbf{K}_{\phi\phi}$ on the kernel hyper-parameters.

3.5 Computational Considerations

The flexibility of our approach comes at the expense of a high computational cost in space and time (see equations (11) to (15)). While inversion of the covariance matrix over *observations* is the dominant cost in a standard GP regression, our inversion formulation uses the output grid in the computation of $K_{yy} = \mathbf{G}\mathbf{K}_{\phi\phi}\mathbf{G}^\top + \sigma^2 I$ making the *computation* of K_{yy} far more costly than its inversion. For example, the size of the grid M is typically $100 \times 100 \times 100$ output cells, while the number of observations N is typically 100×100 as an array on the surface.

To address this problem, we have assumed that the covariance function $\kappa_\phi(\cdot,\cdot)$ is stationary; that our region of interest is divided into a regular grid; and that our observations are arranged in a plane above the region of interest such that each observation lies vertically above the center of a column of cells. Under these restricted conditions we may lever significant computational savings reducing the run time in the typical case above from CPU years to CPU seconds.

Firstly, the stationarity and the gridded arrangement reduces the number of distinct entries in $\mathbf{K}_{\phi\phi}$ from $M(M+1)/2$ to $8M$ (any displacements within $\pm$ width, $\pm$ bredth, $\pm$ height of the grid). The assumption that observations lie above cell columns reduces the number of distinct entries in $\mathbf{G}$ from NM to $4M$ (we need only compute the sensitivity of a central observation to a grid of cells with displacements $\pm$ width, $\pm$ bredth, 0-height below it).

The distinct values are cached in memory, and are easily visually interpreted. For example, Figure 1 depicts a cached representation of a stationary squared exponential covariance function, a gravity sensitivity function, and their use in computing a covariance-vs-displacement lookup table. In this case, the covariance table $\grave{K}$ depicts a squared exponential covariance where the zero displacement case lies in the centre of the grid. The cached sensitivity table $\grave{G}$ contains evaluations of $\mathbf{G}$ for a gravity sensor over the same regular output grid, this time only only defined for cells that lie below a central observation point.

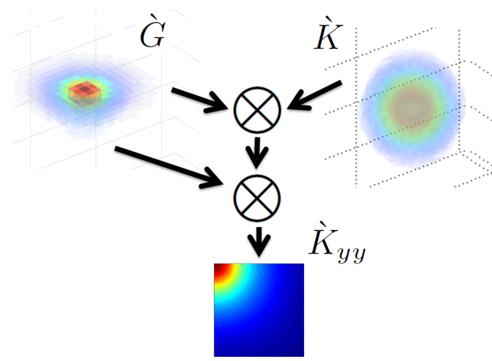


Figure 1: The computation of K_{yy} over discrete displacements. Values of $\mathbf{K}_{\phi\phi}$ are cached as a function of unique displacements into a grid $\grave{K}$. The values of $\mathbf{G}$, the sensitivity of an observation to cells below it, is also computed over the grid to form $\grave{G}$. Significant computational savings can be obtained by using frequency space convolutions to obtain a covariance vs sensor displacement table $\grave{K}_{yy}$.

Significant computational savings are obtained by expressing the multiplications in Equations (11) and (12) as discrete frequency space convolutions. Instead of computing $\mathbf{G}\mathbf{K}_{\phi\phi}\mathbf{G}^\top$, we apply the operator $\bigotimes$ defined as a zero-padded, three dimensional frequency space convolution (using the fast fourier transform). The convolution theorem states that multiplying the frequency space coefficients element-wise is equivalent to convolving the spatial functions, allowing the covariance over all possible spatial displacements to be computed in $\mathcal{O}(M \log M)$:

$$\grave{K}_{gg} = \grave{G} \bigotimes \grave{K} \bigotimes \grave{G}. \tag{18}$$

The discrete grid output of the convolutions, $\grave{K_y y}$ is therefore a lookup table of K_{gg} as a function of sensor-displacement. As we have only considered horizontal sensor

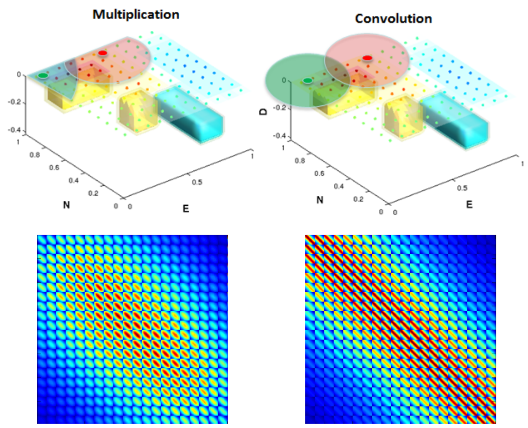


Figure 2: Top: The conceptual difference between computing the covariance function over finite bounds (left) and a stationary template (right). The reigemes are equivalent in the centre of the region, but observations near the edges 'see' less common dependence due to the restricted boundaries. Bottom: The impact on covariance matrices computed over a grid.

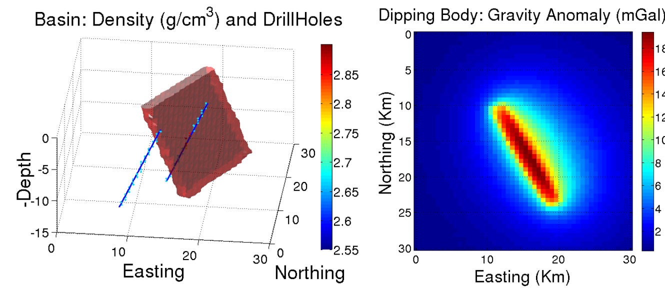
displacement, only its central horizontal slice is needed. This lookup table image is also shown in Figure 1.

A subtle conceptual difference is now present between the multiplicative result and the convolution result. $\mathbf{G}\mathbf{K}_{\phi\phi}\mathbf{G}^\top$ is not actually stationary (even for stationary G and $K_{\phi\phi}$ functions) because of the finite volume of the grid causing pairs of observations near the edges to less common dependence. By convolving, we have assumed that the grid around the sensors extends to the grid's farthest dimensions regardless of position - this actually removes the edge effects from the covariance matrix (and the inversion results), behaving as if the simulation had padding cells. This concept is illustrated in Figure 2. The corresponding difference between the covariance functions is shown in Figure **??** below, where the envelope induced from the finite grid is visible in the multiplication case (left image) only.

4 Outline of Experiments

Our experiments focus on the 3D inversion problem for determining the density contrast of the earth subsurface. This is crucial in many applications, including the characterization of of hot dry rocks up to 5km deep when exploring geothermal energy targets (see e.g. Huenges and Ledru [2010], Ch 2). However, gravity alone provides a very poor depth resolution so it is necessary to fuse additional sources of information into the inversion models in a principled way.

Our experiments investigate gravity inversion on both a real dataset and a simulated scenario. On the real dataset, we study a region in the Cooper Basin in South Australia, fusing ground-based gravity observations with drillhole core samples that provide very sparse direct observations of the density. On the simulated dataset, we perform joint inversions with gravity and drill hole observations on a realistic geological structure, and also investigate the inclusion of magnetic susceptibility observations.



(a) (b)

Figure 3: (a) The density of the dipping body along with the location of the two simulated drill-holes. (b) Simulated gravity observations corresponding to the dipping body.

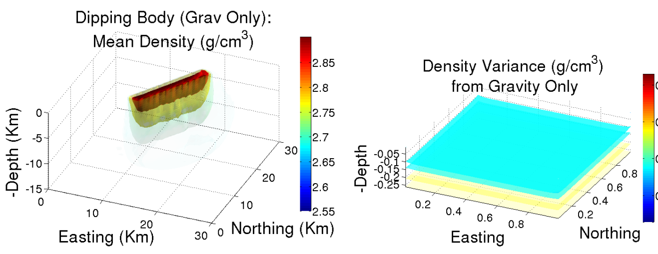
4.1 Simulated Data

The simulated structure is a *dipping body*, where a slab of dense, magnetically susceptible material is present. It is inclined from the surface and has properties typical of igneous rock that may have intruded into the crust. The geometry of this scenario is shown in Figure 3a where the density of the dipping body is shown, along with two simulated drill holes — one drill hole passes through the body, while the nearer one misses the body.

The background material in this volume has been assigned the average density of the earth's crust $(2.67g/cm^3)$. The dipping body was assigned a density of $2.9g/cm^3$ and a magnetic susceptibility of 2×10^{-4} in SI (which is used in a later scenario). A voxel grid of $50 \times 50 \times 25$ cells was used. Gravity observations (in milliGals) were forward simulated onto an array of 50×50 observations centered above each column of voxel cells. These observations are computed as anomalies (difference between the gravity observed and the gravity from a grid filled with the mean density of 2.67). IID Gaussian noise with a standard deviation of 1% of the mean anomaly observations was applied to these observations, producing the observations shown in Figure 3b. Note that the sensitivity of gravity readings drops off quickly with depth, so the deep dipping body produces only a subtle smear-like signature to the right of the body's anomaly.

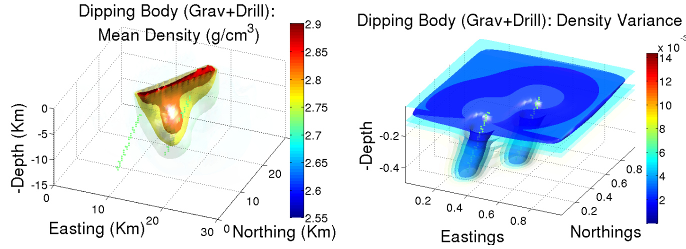
4.2 Real Dataset: Cooper Basin

The real dataset contains Bouger anomaly gravity measurements and core-sample densities from the Cooper Basin formation in South Australia. This data has been provided by the South Australian Department of Manufacturing, Innovation, Trade, Resources and Energy (DMITRE). The Cooper Basin formation is comprised of a basin-shaped basement, overlain with multiple sedimentary layers. The basement rocks are metamorphic, ranging from high grade gneisses to lower grade schist, and contain numerous intruded granite bodies. The large granite structures induce moderate to strong lows in the regional gravity datasets as as due to their relatively lower densities. The sedimentary layers overlaying the basement, particularly the Eromanga basin, contain significant hydrocarbon accumulations, so significant drilling and geophysical data collection has been conducted in the region.



(a) (b)

Figure 4: Outputs of the Gaussian process inversion algorithm using only gravity observations of the dipping body.



(a) (b)

Figure 5: Outputs of the Gaussian process inversion algorithm after fusing gravity and drill observations of the dipping body.

Measure	Gravity Only	Gravity & Drill Holes
UIQ	0.480	0.589
Corr.	0.538	0.620

(a)

Method	RMSE (g/cm^3)
GP	1.06
UBC	2.06

(b)

Table 1: Quantitative evaluation of inversion results on the dipping body. (a) UIQ index [Wang and Bovik, 2002] and correlation between the predicted density and the true density for GP inversion. (b) Mean square error of GP vs `UBC`.

Data Preparation

For the Cooper Basin data we used the Bouguer anomaly, which accounts for the Earth's total mass, variations in radius with latitude, the effect of elevation on the measurements and the mass of the local topography such as mountains and valleys. In addition, raw drillhole information provides the depth of samples and a corresponding density (drawn from analysis of the target formation). The samples tend to occur on the formation boundaries, although there are many exceptions. To mitigate the influence of this clustering around boundaries, density measurements over the extent of each target formation were averaged - averages over lines can be naturally incorporated into the inversion algorithm.

5 Empirical Results and Analysis

This section presents empirical results of our inversion algorithms on the simulated dataset (dipping body) and the real dataset (Cooper basin).

5.1 Dipping Body

Gravity observations were inverted in isolation by applying the GP inversion algorithm to learn the hyperparameters of an anisotropic polynomial covariance function. Because the accuracy of gravity surveys is well understood, appropriate small noise variances were included in the model.

The inversion produced the mean shown in Figure 4a. The GP formulation also provides uncertainty, which is shown in Figure 4b. This inversion is ill posed; we cannot infer whether a gravity signature corresponds to a large mass deep down, or a small mass on the surface, so there is little inherent depth resolution. The uncertainty in this case appears featureless because of the uniform sampling pattern. The mean is more interesting, as it shows the body extending down from the surface, although the inclination of the body cannot be resolved. Forward simulation of gravity observations on the predictive mean confirms that this is indeed one of the infinite valid solutions to the problem, selected through the prior over density structure.

Fusing drill hole measurements with the gravity observations can further constrain particular features, because the drill samples the density directly at specific locations. In our scenario one of the simulated drill-holes passes through the body. Using the multi-task model, updated mean and variance volumes are obtained: Figure 5a, and Figure 5b. The problem remains ill posed, but (as the variance shows), the model is now confident at depth where drill observations are provided, adding two columns of low-variance. The drill holes can only improve the model in the vicinity where they are correlated to the volume, so the overall quantitative improvement in prediction quality is relatively weak. In the regions the model is most incorrect, it is also assigning a high variance.

By comparing the side-profiles of the reconstructed body with the gravity only case, and the synthetic truth, there is a valuable qualitative difference — the incline of the body is resolved. These side-profiles are compared in Figure 6. Table 1 shows a quantitative evaluation of our GP inversion method (a) having additional drill hole data and (b) compared to the UBC solution.

5.2 Fusing Magnetic Observations

Magnetic field measurements on the Earth's surface vary significantly with respect to location, and are weakly influenced by induced magnetism from dense, iron-rich minerals below the surface. While drill-holes and gravity provide indirect measurements of a common attribute (density), additional observations of Total Magnetic Anomaly (TMA) (which is the anomaly of the Earth's magnetic field magnitude) can only be fused with gravity and drill-holes if a relationship can be captured between the physical attributes of magnetic susceptibility and density. The current GP model was extended to fuse TMA by learning a simple covariance relationship between these attributes, producing a multi-sensor multi-task covariance model consistent with the physical sensors.

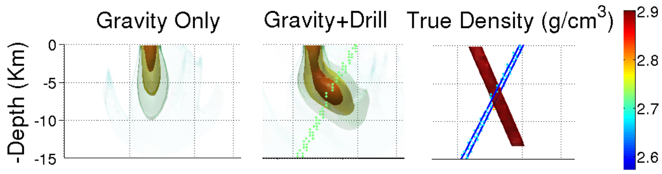
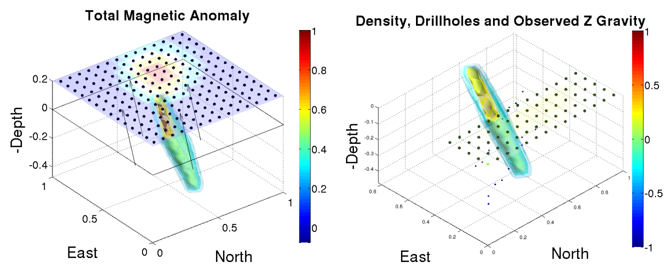


Figure 6: Side-profiles of the dipping body in the mean predictive density before (left) and after (middle) fusing information from the drill holes. (right) Ground truth.



(a) (b)

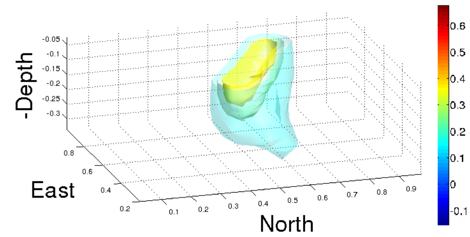
Figure 7: Simulated magnetic data along with gravity and drillhole data. (a) Forward magnetism simulation for the dipping body (no drill-holes were simulated; glyphs indicate magnetic field direction). (b) Incomplete gravity coverage was simulated over the dipping body, with two full drillholes.

The benefit of fusion is most clearly seen when the coverage pattern differs between sensors. The dipping body has been assigned a positive susceptibility, and gravity observations were simulated over only half of the test region, accompanied by drill holes (Fig. 7b). TMA was available over the full region, but without any direct measurements of susceptibility. The magnetic anomaly observations were forward simulated above ground level for the Earth field vector at the Cooper Basin's location, yielding the observations in Fig. 7a.

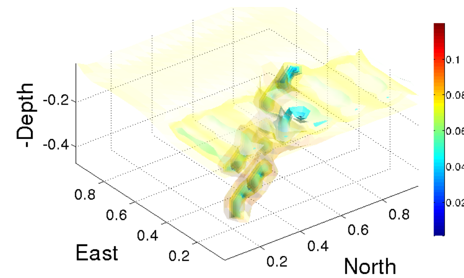
The inversions were initially run independently, leading to poor predictions of both density and susceptibility from lack of information. By then allowing the GP to model a stochastic coupling between the density and susceptibility, the joint inversion was able to use complementary information from each dataset. This has led to inversion results that closely resemble the full coverage inversions. In addition, the predictive variance assigns maximum confidence to locations where all sensor modalities were available together. The fused density (Fig. 8a and Fig. 8b) and susceptibility (Fig. 9a and Fig. 9b) are provided for this case.

5.3 Cooper Basin Inversion

The inversion algorithm was applied to the real Cooper Basin data to obtain the two predictive means shown in Figure 11. The location of this data is shown in Figure 10 (left) marked as a black rectangle, along with the processed observations (centre) that show the incomplete and irregular gravity coverage available. Interestingly, the figure on the right shows the forward simulated gravity from the GP prediction, which both agrees with the observations, and fills in the gaps.



(a) Fused density mean



(b) Fused density variance

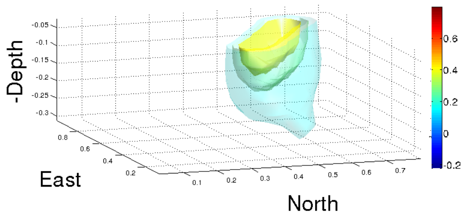
Figure 8: Joint inversion results for density with fusion model that includes gravity and magnetic data of the dipping body.

Inversion method	RMSE (g/cm^3)
Gaussian process inversion	0.029
UBC inversion	0.073

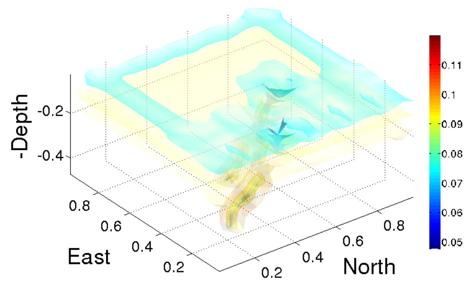
Table 2: Cross-validation Errors for our GP inversion method and `UBC` codes on the Cooper Basin dataset.

The results of the inversion are presented in Figure 11 with gravity only, and with drill-hole data in the same region. It appears that the features in the predictive mean are geologically reasonable and can be compared against existing geological knowledge of the area. For example, the low density zone in blue in Figure 11 is likely to be controlled by a combination of the base of the trough depo-centres described in Section 4.2 and the low density granitic material that intrudes the basement and deeper basin sediments. The NE strike of the northern margin of this anomaly is consistent with dominant NE structural trend in the region and the strike of the faults that are likely to bound the basement highs. The higher density anomalies that have been resolved at deeper levels visible in Fig. 11 (right) are likely to relate to basement rocks between intrusive bodies although this needs to be tested further against other available datasets and interpretations.

Furthermore, while the ground truth density in the Cooper Basin is not known, it is possible to obtain limited quantitative results by cross validating on the drill holes (Table 2). Crustal rock densities are close to $2.67g/cm^3$, so the error standard deviation is approximately 1% and 3% for our method and the UBC result respectively. Cross validation was conducted using the GP inference algorithm, and as a point of comparison the `UBC` software was run on the same data. The results in this case suggest that our approach outperforms our UBC setup. See section 6 below for a general comparison between



(a) Fused susceptibility mean



(b) Fused susceptibility variance

Figure 9: Joint inversion results for susceptibility with fusion model that includes gravity and magnetic data of the dipping body.

these techniques.

6 Relation between UBC and GP inversions

Both UBC inversion and GP inversion can be framed into a single Gaussian prior - Gaussian likelihood model. The UBC solution is the mode of the posterior (MAP), which can be obtained by minimizing the unnormalized negative log posterior. However, UBC considers the geophysical parameters, e.g. densities, uncorrelated (a priori) and in the best case having different but fixed variances (which depend on depth). Such variances are set beforehand in order to achieve a weighting effect that compensates the decay of the sensor sensitivity with respect to the depth. GP inversion, on the contrary, allows these parameters to be correlated. These correlations are learned from the data through maximization of the marginal likelihood (i.e. hyper-parameter learning).

7 Discussion

This project has applied Bayesian machine learning methods to the new application of joint geophysical inversions. Our approach allows for an assessment of uncertainty associated with inversions of geophysical datasets, enabling the geophysicists and geologists who will use the results in decision-making workflows to robustly assess the likelihood of predictions being accurate throughout the model volume. In the longer term this will allow interpretation workflows to include a quantifiable assessment of uncertainty, something geophysical surveyors as a community are only just beginning to tackle.

In future work we will investigate low-rank approximations to $\mathbf{K}_{yy}$ to address the cost of GP inference once the covariance matrix has been computed (see e.g. Quiñonero Candela and Rasmussen [2005] for an overview). We will also investigate likelihood approximations that exploit the structure of spatial problems, such as those presented in Stein *et al.* [2004]. Given that the inversion problem is ill posed, it is critical to form priors that capture the knowledge of geologists about the plausible structure of rocks. Their prior knowledge is also critical for characterising non-linear dependencies between different rock properties, which are poorly constrained by geophysical data alone.

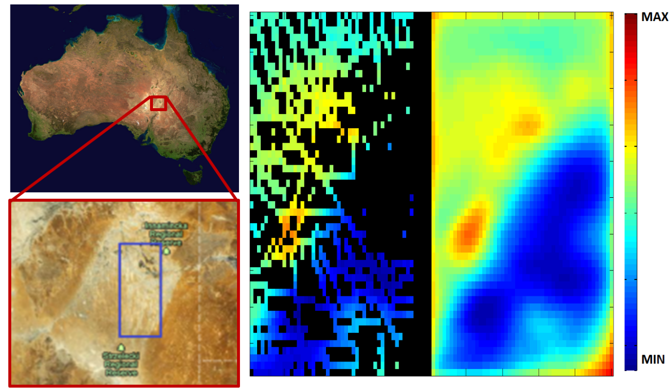


Figure 10: Satellite view of Australia with the Cooper Basin region of interest highlighted by the blue bounding box (below). The measured gravity anomalies (middle), and the predicted gravity anomalies (right) correspond to this blue box.

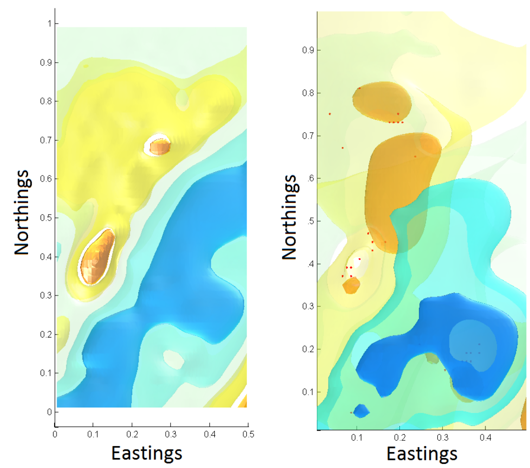


Figure 11: The predictive mean outputs of the Gaussian process inversion algorithm on the Cooper Basin data. Gravity only (left) and gravity with drill holes (right).

Acknowledgements

NICTA's standard thankyou blurb. In addition This work is supported by the Australian Renewable Energy Agency (ARENA) Emerging Renewables Program.

References

Edwin V. Bonilla, Kian Ming A. Chai, and Christopher K. I. Williams. Multi-task Gaussian process prediction. In *Advances in Neural Information Processing Systems 20*. MIT Press, 2008.

Miguel . Carreira-Perpiñan. *Continuous latent variable models for dimensionality reduction and sequential data reconstruction*. PhD thesis, Dept. of Computer Science, University of Sheffield, UK, 2001.

Noel. A. C. Cressie. *Statistics for Spatial Data*. Wiley, New York, 1993.

Peter. K. Fullagar and Glenn. A. Pears. Towards geologically realistic inversion. In *Fifth Decennial International Conference on Mineral Exploration*, 2007.

Antonio Guillen, Philippe Calcagno, Gabriel Courrioux, Aurore Joly, and Patrick Ledru. Geological modelling from field data and geological knowledge: Part II. Modelling validation using gravity and magnetic data inversion: Recent Advances in Computational Geodynamics: Theory, Numerics and Applications. *Physics of the Earth and Planetary Interiors*, 171(1-4):158–169, 2008.

Ernst Huenges and Patrick Ledru, editors. *Geothermal Energy Systems: Exploration, Development, and Utilization*. John Wiley & Sons, 2010.

Philip Kearey, Michael Brooks, and Ian Hill. *An Introduction to Geophysical Exploration*. Blackwell Science, 3rd edition, 2002.

Yaoguo Li and Douglas. W. Oldenburg. 3D inversion of magnetic data. *Geophysics*, 61(1):394–408, 1996.

Yaoguo Li and Douglas. W. Oldenburg. 3D inversion of gravity data. *Geophysics*, 63(1):109–119, 1998.

D. Nagy, Gabor Papp, and Judit Benedek. The gravitational potential and its derivatives for the prism. *Journal of Geodesy*, 74(7):552–560, 2000.

Douglas W. Oldenburg, Yaoguo Li, Colin G. Farquharson, Peter Kowalczyk, Theo Aravanis, Alan King, Ping Zhang, and Anthony Watts. Applications of geophysical inversions in mineral exploration problems. *The Leading Edge*, 17:461–465, 1998.

Joaquin Quiñonero Candela and Carl Edward Rasmussen. A unifying view of sparse approximate Gaussian process regression. *Journal of Machine Learning Research*, 6:1939–1959, 2005.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 2006.

Pejman Shamsipour, Denis Marcotte, and Michel Chouteau. 3d stochastic joint inversion of gravity and magnetic data. *Journal of Applied Geophysics*, 79:27–37, 2012.

Michael L. Stein, Zhiyi Chi, and Leah J. Welty. Approximating likelihoods for large spatial data sets. *Journal of the Royal Statistical Society Series B*, 66(2):275–296, 2004.

Michael L. Stein. *Interpolation of Spatial Data: Some Theory for Kriging*. Springer, 1999.

Albert Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, 2005.

Zhou Wang and A. C. Bovik. A universal image quality index. *IEEE Signal Processing Letters*, 9(3):81–84, March 2002.

Ulvi Yurtsever, Caren Marzban, and Marina Meilă. On the gravitational inverse problem. *Applied Mathematical Sciences*, 5(57):2839–2854, 2011.