MULTI-OBJECTIVE OPTIMIZATION AND JOINT INVERSION FOR ACTIVE SENSOR FUSION - IN A NUTSHELL

Sebastian Haan[1], *Fabio Ramos*[2], *Dietmar Müller*[3]

[1] Sydney Informatics Hub, [2] School of Computer Science, [3] School of Geosciences
The University of Sydney
2006 NSW, Australia

ABSTRACT

A critical decision process in data collection is how to efficiently combine a variety of sensor types and to minimize total cost. Here we propose a probabilistic framework for multi-objective optimization and sensor fusion given an expensive cost function for allocating new measurements. This new method is devised to jointly solve multi-linear forward models of 2D remote sensor data to 3D-geophysical properties using sparse Gaussian Process kernels while taking into account the cross-variances between different geophysical measurements. Multiple optimization strategies are tested and evaluated on a set of geophysical data. We demonstrate the advantages on a specific example of a joint inverse problem, recommending where to place new ground drill-core measurements given gravity and magnetic remote sensor data, the same approach can be applied to a variety of remote sensing problems - ranging from constraints limiting surface access for data collection to adaptive multi-sensor positioning.

***Index Terms*—** Bayesian optimisation, Sensor Fusion, Joint Inversion, Gaussian Process

1. INTRODUCTION

One of the most important aspects when dealing with incomplete information is where to collect new observations. This may include the layout of a particular geophysical survey, drill-site placements or new aerial data, in particular if measurements are very costly or resources are limited. Bayesian Optimisation (BO) solves this decision making problem by finding global optimal solutions using a probabilistic framework given multiple measurements, prior knowledge, and model uncertainties. This offers solutions for wide range of sensor fusion problems and allocation problems such as: What is the optimal mixture and placement of different sensors? Where to sample if the cost function is incomplete or if there are large uncertainties in future total budget allocations? How to position sensor grids over time if the model state is dynamic or has a moving target?

Typically the function(s) involved in BO are unknown or very expensive to evaluate, and are therefore surrogated by prior models, such as Gaussian process (GP) models for tractability (see [1]). Once new data are available to evaluate the objective function, the prior is updated to form the posterior distribution over the function space.

While BO research began by focusing on single-objective bound-constrained optimisation (see for earliest work [2], [3]), it has rapidly developed to address multi-objective problems (see e.g. [4]). Today multi Objective BO is recognised as a robust statistical method in handling multiple constraints and to find the optimal solution for competing objectives and multi-task problems. However, the main challenge still lies in constructing a prior model that can be computationally efficiently evaluated and that transforms the combined data and uncertainties from multiple sensor measurements coherently into a posterior function approximating the true model, such as a 3D model of a geophysical properties; this is also know as an inversion problem.

One promising way to infer the geophysical properties is by simultaneously interpreting multiple sensor measurements using a single model. The motivation behind this joint inversion is to provide a better constrained joint solution of multiple, and often distinct sensor types, rather than taking individual solutions that only satisfy their aspect of data on their own. In case these multiple sensors are sensitive to different aspects of the geology, e.g. density and magnetic susceptibility, we can take advantage of the statistical properties of a model that simultaneously need to satisfy two or more independent measurements. The resulting 'boost' in information is dependent on the strength of the correlation between the different physical properties that are measured, but if statistically properly taken into account, this approach can offer a much greater value than the 'sum of their parts'.

In this paper we propose a probabilistic framework that combines multi-objective optimisation and joint inverse problems given an expensive cost function. We demonstrate the capabilities of this method by jointly solving multi-linear forward models of distinct 2D remote sensor data to 3D-geophysical properties using sparse Gaussian Process kernels while taking into account the correlations between distinct geophysical aspects.

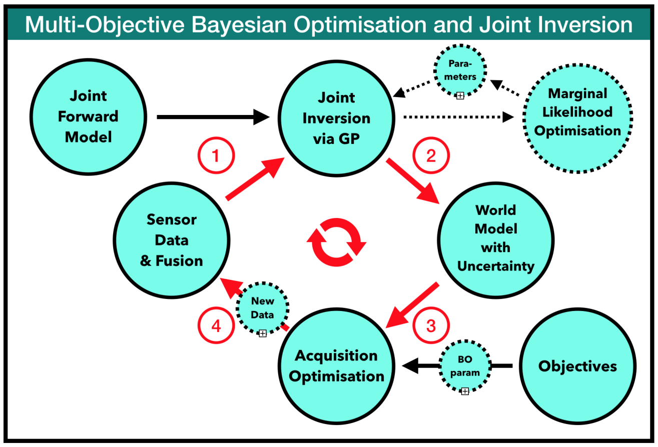

Fig. 1. Graphical model of the probabilistic framework for multi-objective optimisation and joint inversion; the process stages are: 1) Joint inversion via Gaussian Processes based on multiple sensor data fusion and forward model, 2) Posterior model generation of 3D multi-geophysical properties. 3) Maximisation of acquisition function to allocate optimal new sensor location and type. 4) New acquired data is combined with existing data; process repeats until maximum number of iterations is achieved.

2. METHODOLOGY

Our approach for joint inversion and optimisation builds upon a fully probabilistic model, including likelihood objectives for model selection and complete uncertainty propagation through all stages of data processing: from prior selection and input, forward model inversion, to the final sensor optimisation and real world model output. Fig. 1 shows a graphical model of the developed probabilistic framework.

To solve the inverse problem for linear systems, we deploy a Bayesian framework with Gaussian process priors over the physical properties Φ of interest. Gaussian processes (GP) are a flexible, probabilistic approach using kernel machines with non-parametric priors, and are successfully used in a large range of machine learning problems (see [1]). An important advantage of the Bayesian method is that it generates a predictive distribution with a mean $\mu(\Phi)$ and variance $\sigma^2(\Phi)$, which are prerequisites for Bayesian optimisation with respect to, e.g., information gain from a new measurement. Another advantage is that the GP marginal likelihood function is well defined by the values of their hyper-parameters, which allows it to optimise them exactly. This reasoning about functions under uncertainty and their well-tuned interpolation character allows GPs to work extremely well for sparse data as well as for big data (see [5] for solving large covariance matrix).

To take fully into account cross-covariances between multiple model parameters (e.g., rock density and magnetic susceptibility), we construct sparse cross-covariance terms between all kernel pairs following [6]. The full covariance matrix is then constructed by setting sparse covariance functions on all diagonal block elements and sparse-sparse cross-covariance functions on all non-diagonal block elements. Cross-covariance amplitude terms are given by the linear correlation coefficients between the corresponding geophysical properties.

2.1. Multi-Objective Bayesian Optimisation

Bayesian Optimisation (BO) is a powerful framework for finding the extrema of objective functions that are noisy, expensive to evaluate, do not have a closed-form (e.g. black-box functions), or have no accessible derivatives. The set of objectives is defined in an acquisition function g, which guides the search for a user-defined optimum. The posterior solution of the GP is used to query for the next most promising point $\mathbf{x}^\star$ based on objectives that are defined in g, equivalent of finding the $\mathbf{x}^\star = \arg\max_{\mathbf{x}} g(\mathbf{x})$. The BO algorithm maximises g, that quantifies the benefit of choosing a specific location to be sampled. Once the maximum of g is found and a new measurements is taken at the proposed position, the model is updated with the new data, and the BO algorithm proposes a new point at each iteration step until the objectives are achieved or a maximum of iterations is reached.

The key of BO is the acquisition function g, which typically has to balance between a) exploration, i.e., querying points that maximise the information gain and minimize the uncertainty of a model, b) exploitation, i.e. querying points that maximise the reward (e.g. concentrating search in the vicinity locations with high value such as minerals), and c) minimize the number of samples given an expensive cost function for any new measurement. Multi-objective optimisation is particular important for joint inversion problems, which generate posteriors via a GP prior with multiple outputs such as different geophysical properties. We have applied and compared several acquisition functions that make use of the mean and variance of the posterior, such as the probability of improvement, the Expected Hyper Volume Improvement (EHVI) and the Upper Confidence Bound (UCB; see for an overview [7]), which we define by maximising

$$UCB_\Phi(\mathbf{x}) = \mu_\Phi(\mathbf{x}) + \kappa_\Phi \cdot \sigma_\Phi(\mathbf{x}) - \gamma_\Phi c(x) \qquad (1)$$

given the mean value for the prediction $\mu_\Phi(\mathbf{x})$, the variance $\sigma_\Phi(\mathbf{x})$, and a cost function $c(\mathbf{x})$, which is defined by the cost of obtaining a measurement at the sample points $\mathbf{x}$. The parameter κ_Φ and γ_Φ define the trade-off in exploration-to-exploitation and gain-to-cost, respectively. While κ is typically set to a constant value (here best results achieved with a κ of 1-3), in principle it is possible to employ an adaptive kappa that changes with the number of BO iterations.

3. EXPERIMENT SETUP

3.1. Linear Forward Models

To test the proposed BO algorithm, we first define the set of forward models that transform the localized measurement of a remote sensor grid into a 3D representation of geophysical properties of a region. The most common geophysical linear forward model are gravity and magnetic forward models: The gravity forward model is defined by using Li's tractable approximation for a 3-D field of constant density prisms ([8]) and can be determined analytically. The induced magnetic field calculation uses Li's tractable approximation for a 3-D field of prisms of constant magnetic susceptibility, which depends on the magnetic mineral content such as iron-rich minerals including magnetite and haematite below the surface and is measured by the response of their magnetic dipoles induced by the Earth's magnetic field. The joint GP inversion takes into account a covariance that exists between density and magnetic susceptibility.

3.2. Synthetic Model Setup

To test the joint GP inversion and the proposed BO algorithm, a set of multiple distinct synthetic geophysical models are created as ground truth, including two-dipping body, layered, and clumpy models as shown in Fig. 2. For each model we generate a 3D voxel cube with geological structures given by their density and magnetic susceptibility and the corresponding 2D gravity and magnetic remote sensor measurements. The advantage of using simulated environments is that this allows us to evaluate precisely the accuracy of the reconstructed 3D cube in comparison to the ground truth for a range of different conditions.

We define the goal of the BO experiment to find the optimal location where to place a new sample measurement as given by a set of acquisition functions. The experiments are carried out in the following steps (see Fig. 1):

1. Calculate response $\mathbf{y}$ of gravity and magnetic sensors above surface, and of any pre-existing drill-core samples.

2. Reconstruct cube by solving joint GP inversion and calculate geophysical properties in 3D as defined by their mean and variance for each voxel.

3. Find the x, y location for new drill-core that maximises the BO acquisition function(s).

4. Add selected drill-core data and repeat until the maximum of iterations (here 25) or final objective is achieved.

4. RESULTS

To evaluate the efficiency of the proposed BO optimisation methods, we show in Fig. 3 the results of our simulated experiments for a sampling series of 25 drill-core locations. The Root Mean Square Error (RMSE) is calculated for the three different synthetic models and the performance of multiple BO optimisation functions are compared to two mock traditional non-BO approaches: The first one samples from a random selection of 25 uniform evenly spaced drill-core locations across the entire model space. The second non-BO method selects drill-core locations from a random distribution that is probabilistically weighted with the gravity and magnetic field sensor measurement. These two methods are compared to our BO optimisation using a) Pareto EHVI, b) the sum of the two UCB function over density and magnetic properties, c) the Expected Improvement (EI, [9]), and d) optimising only the variance to minimise the total uncertainty of the reconstructed cube.

The results of these experiments are presented in Fig. 3, which shows that the two BO optimisation strategies Pareto and UCB consistently outperform the non-BO methods by roughly a factor of two in final accuracy and speed of convergence towards the true model. Only in case that the model consists of continuous smooth layers, the Variance-only BO method has the fasted convergence speed, but eventually converges for larger drill-core samples towards the same RMSE values as Pareto and UCB. Furthermore, our experiment shows that the BO method with cost weighting significantly outperforms the standard non-BO method while simultaneously reducing the cumulative costs of drilling

5. CONCLUSIONS

A probabilistic framework for joint inversion and multi-objective BO to select new sensor measurements is developed and its capabilities are demonstrated on a test-set of synthetic geophysical models by solving the example problem of allocating iteratively new expensive drill-core samples based on sparse 2D remote sensor data. The results show for a wide range of geological models that this method can a) accurately reconstruct a 3D model of multiple geophysical properties and their associated uncertainties; and b) significantly outperform traditional non-BO sampling schemes, which may ultimately lead to more efficient and cost-saving sensor selection.

A key feature of the applied method is to consider all cross-variances between geophysical properties by solving simultaneously multiple forward models and to fuse jointly multiple and distinct sensors. Thus, this approach can take full advantage of additional sensor information to reveal properties of the reconstructed geophysical model at higher accuracy. By using non-parametric GP priors, the solution of the inverse problem provides a complete posterior distribu-

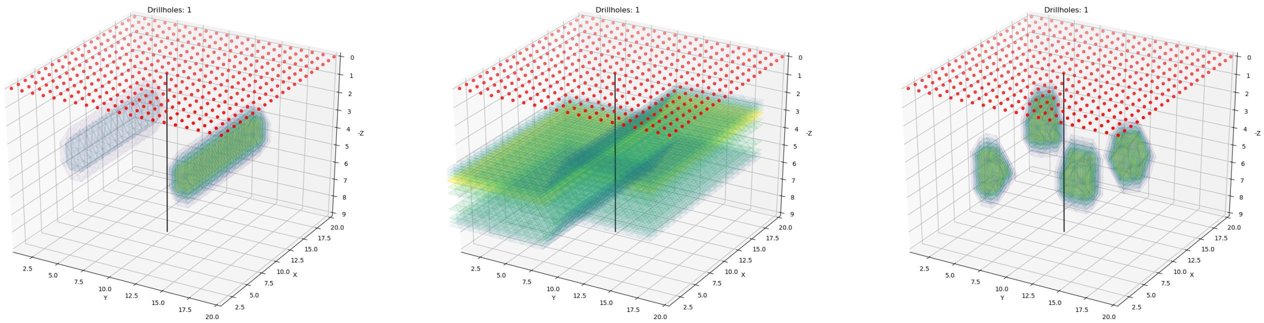

Fig. 2. The 3D synthetic models' density distribution for three different models, from left to right: 1) Two-dipping body model, 2) Three-layer model, and 3) Clumpy model. The magnetic susceptibility follows a similar distribution with varying correlation coefficients. The remote gravity/magnetic sensor grid is shown as points near the surface.

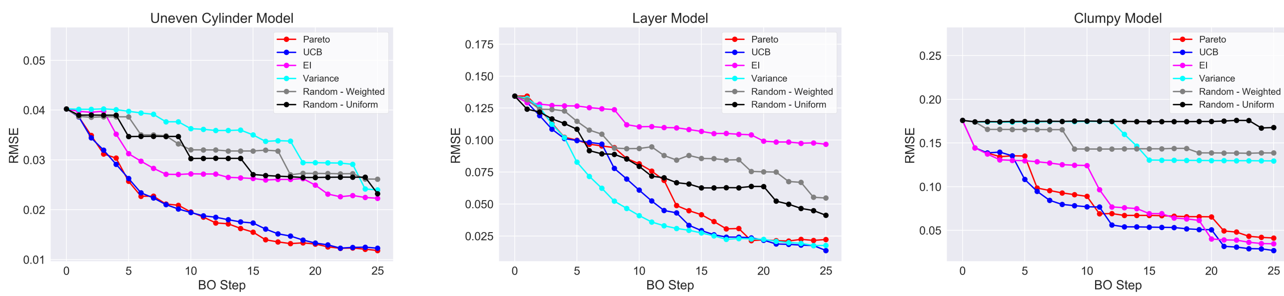

Fig. 3. RMSE results between reconstructed and true model for the three models. Multiple BO drill-core selection strategies are compared against each other in terms of their performance: 1) Pareto; 2) UCB; 3) EI; and 4) Variance only. Additionally, two non-BO standard methods are shown: 5) randomly sampling drill-cores from a uniform evenly spaced grid (Random - Uniform); 6) random sampling with probabilistic sensor data weighting (Random - Weighted).

tion for all geophysical properties under consideration at any location or grid resolution.

In general, the approach provides a very flexible approach for a large range of sensor fusion problems and is computationally efficient enough to rapidly update a model with new data while solving multiple measurement decision problems simultaneously.

6. REFERENCES

[1] Carl Edward Rasmussen and Christopher KI Williams, *Gaussian process for machine learning*, MIT press, 2006.

[2] Harold J Kushner, "A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise," *Journal of Basic Engineering*, vol. 86, no. 1, pp. 97–106, 1964.

[3] J Mockus, "On bayesian methods of optimization," *Towards Global Optimization*, pp. 166–181, 1975.

[4] Shinkyu Jeong and Shigeru Obayashi, "Efficient global optimization (ego) for multi-objective problem and data mining," in *Evolutionary Computation, 2005. The 2005 IEEE Congress on*. IEEE, 2005, vol. 3, pp. 2138–2145.

[5] Arman Melkumyan and Fabio Ramos, "A sparse covariance function for exact gaussian process inference in large datasets.," in *IJCAI*, 2009, vol. 9, pp. 1936–1942.

[6] Armon Melkuyman and Fabio Ramos, "Multi-kernel gaussian processes," in *IJCAI*, 2011, vol. 22, p. 1408.

[7] Eric Brochu, Vlad M Cora, and Nando De Freitas, "A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning," *arXiv preprint arXiv:1012.2599*, 2010.

[8] Li Yaoguo and Douglas W Oldenburg, "3d-inversion of gravity data," *Geophysics*, vol. 63, no. 1, pp. 109–119, 1998.

[9] Donald R Jones, Matthias Schonlau, and William J Welch, "Efficient global optimization of expensive black-box functions," *Journal of Global optimization*, vol. 13, no. 4, pp. 455–492, 1998.