

Uncertainty Reduction in Reservoir Modeling

Sam Subbey, Mike Christie, and Malcolm Sambridge

ABSTRACT. This paper presents a method by which uncertainty in reservoir modeling can be reduced by integrating observed data in the estimation of model parameters. The problem is inverse and ill-posed. Further, since the observed data and the reservoir properties are non-linearly related, reducing uncertainty becomes a non-trivial task.

The methodology involves generating multiple reservoir descriptions, conditioned on the observed data. This is facilitated by sampling in a multidimensional parameter space defined by Voronoi cells. In geophysics, this method has been successfully applied in solving the inverse and non-linear problem of seismogram inversion. Here, we present its application in petroleum reservoir engineering.

The strength of the method in performance prediction is validated using fine scale data from the Tenth SPE Comparative Solution Project. We use an upscaled model to history match the fine scale data and forecast the fine grid performance using the maximum likelihood model.

1. Introduction

In petroleum engineering, mathematical models of the reservoir are routinely used in the prediction of reservoir performance, and in making economic and management decisions. These models consist of a complex system of time dependent, non-linear partial differential equations, describing fluid flow (oil/gas/water) through the porous medium. The models are ascribed physical descriptive properties of the reservoir such as geometry and fluid flow properties. The predictive accuracy of the models therefore, depends on the degree of precision with which the physical reservoir properties are known.

Direct measurements of porous media properties of hydrocarbon reservoirs are available at only a limited number of points, where core samples are taken. Typically, the cores are of the order of 10^{-16} of the reservoir volume. Other measurements are inferred from logs, well tests or reservoir data, and all involve averaged responses over larger length scales. Further, since the laboratory experiments are often conducted under different flow scenarios than in the reservoir, [Dak94], there is a large degree of uncertainty in the reservoir properties. Uncertainties in the detailed description of porous media properties such as porosity, absolute permeability, relative permeability and capillary are large contributors to uncertainty in

1991 *Mathematics Subject Classification.* Primary 93A30; Secondary 74G75.

©0000 (copyright holder)

reservoir performance forecasting. Of all the porous media properties, probably the primary factor that limits the reliability of predictions of reservoir behavior by numerical simulation is the accuracy of reservoir rock/fluid properties, [YW91]. The only way to reduce the uncertainty in a reservoir model is to incorporate dynamic information derived from knowledge of aspects of the porous media into the model description. This information is normally available in the form of several time series (oil/gas/water production and pressure), throughout the production time of the reservoir. The process of incorporating such historical data in the generation of reservoir models is commonly known as history matching, [PP99]. Thus in history matching, by adjusting model parameters, one aims at obtaining a model output which is as close to the history data as possible.

Mathematically, history matching may be viewed as the task of finding a zero of a multivariate time-varying function, namely the reservoir simulator, [CGSS96]. As stated above, history data and reservoir petrophysical properties are related to each other through flow equations, which are highly non-linear. Consequently, incorporating dynamic engineering data into reservoir modeling is a high-dimensional, ill-posed, non-linear inverse problem, [WDC99]. Like most inverse problems, this problem is characterized by non-uniqueness of the solution. Thus there exists a family of model parameters that satisfy the forward problem and defining a ‘good’ history match is non-trivial. This is just one of the challenges of history matching. See [CGSS96] for a more general overview.

The traditional industrial approach to history matching can be summarized as following, [LBA⁺99]:

1. construct an initial reservoir model;
2. by adjusting the model parameters, perform the history matching;
3. use the adjusted model to forecast future reservoir performance.

We note however that this traditional approach does not address the fact that the solution of the problem is non-unique, and consequently, uncertainty in the forecast cannot be adequately quantified. Since the problem is characterized by non-uniqueness, an obvious method of estimating the uncertainty would be to generate multiple reservoir realizations, i.e., an ensemble of models which all honor the history data, and forecast future production using this ensemble. The spread in the predictions of the multiple reservoir realizations will in general, be indicative of the degree of uncertainty in the forecasts. Recently, this realization has led to a new approach to the problem of history matching.

In the literature, two major techniques of history-matching are identifiable namely, deterministic and stochastic methods, [PP99]. The most efficient deterministic methods are the gradient methods. In this approach one defines an objective function which defines a measure of the discrepancy between the history data and the model generated data. The mathematical model is then differentiated with respect to the model parameters in order to minimize the objective function. In the literature, these methods have been reported as having fast convergence rates for an optimal set of parameters. These methods however suffer a drawback in that they may fail to converge, or in some cases, converge to local minima of the mathematical model, [PP99].

Prominent among the stochastic methods are those based on simulated annealing and genetic algorithms. Apart from being computationally less cumbersome to

implement compared with deterministic models, stochastic methods avoid calculation of gradients. Their convergence rates however are, in general, much slower than deterministic methods. For instance, simulated annealing and genetic algorithms are theoretically considered as global optimization algorithms. In the practical application of these algorithms however, the global minimum of the objective function may not always be attained, see [PP99]. Rather than seek a single global optimum, a better approach will be to find an ensemble of models that preferentially sample the good data-fitting regions of the model parameter-space, bearing in mind the fact that the inverse problem has a non-unique solution.

This paper presents a new derivative-free approach for reducing uncertainty in reservoir modelling by generating multiple reservoir realizations conditioned on history data. The algorithm falls in the same category as simulated annealing and genetic algorithms but differs from these methods in that the objective here is to find an ensemble of models that selectively sample parameter space, concentrating on the best history-matching regions of the parameter space, [Sam99]. The new search algorithm makes use of geometrical constructs known as Voronoi cells to derive the search in parameter space.

The strength of the new method in performance prediction is validated using fine scale data from the Tenth SPE Comparative Solution Project, [CB01]. The approach in this paper is to pick one fine grid solution and regard it as a stand in for production data. We use an upscaled model, [Chr96] and [Chr01], to history match the fine scale data and forecast production. Realizations in the ensemble are evaluated on the basis of the quality of their match to the fine grid data.

This paper is organized in the following way. In section 1.1, we discuss tessellation using Voronoi cells in general and present a description of the neighbourhood algorithm. Section 1.2 presents the problem, while section 1.3 gives a description of the models. Sections 1.4 and 1.5 present the model parameters to be determined and the algorithm control parameters, respectively. We present our results and discussions in section 1.6 and section 1.7 summarizes our main conclusions.

1.1. Voronoi cells and the Neighbourhood Algorithm. We begin with an introductory discussion on Voronoi cells, see [Aur91] and [Sam99].

DEFINITION 1.1. Let ψ denote a set of n points in the plane. For two distinct points $x_1, x_2 \in \psi$, the *dominance* of x_1 over x_2 , $dom(x_1, x_2)$ is defined as

$$(1.1) \quad dom(x_1, x_2) = \{x \in R^2 | \delta(x, x_1) \leq \delta(x, x_2)\},$$

for δ denoting the Euclidean distance function.

From Definition 1.1 above, $dom(x_1, x_2)$ is a closed half plane bounded by the perpendicular bisector (*separator*) of x_1 and x_2 . Further, the *region* of a point $x_1 \in \psi$ is the portion of the plane lying in all of the dominances of x_1 over the remaining points in ψ , i.e.,

$$(1.2) \quad reg(x_1) = \bigcap_{x_2 \in \psi - \{x_1\}} dom(x_1, x_2).$$

The regions are thus convex polygons constructed from $n - 1$ half planes. Each point on an edge of a region is equidistant from exactly two points, and each vertex is equidistant from at least three. Thus the regions are joined edge to edge and vertex to vertex, i.e., a polygonal partition of the plane. This partition is called the *Voronoi diagram*, $V(\psi)$, of the finite point-set ψ .

Summarily, given some number of points in a plane, their Voronoi diagram divides the plane according to the *nearest-neighbor rule* namely, *each point is associated with the region of the plane closest to it.* Each Voronoi polygon can thus be regarded as a closed set where each cell is simply the nearest neighbor region about the generator point. By nature of their construction, Voronoi cells are unique, space filling and have size (volume, area) inversely proportional to the density of the generating points. For a detailed literature on Voronoi cells, see [OBS92].

The neighbourhood approximation property of Voronoi diagrams provides a simple way of performing non-smooth interpolation of an irregular distribution of points in multidimensional space. Thus if one represents sought model parameters by a Voronoi diagram, the problem of generating multiple model realizations translates into one of sampling a multidimensional parameter space defined by the Voronoi cells. By exploiting the interpolatory properties of Voronoi cells, new models can be generated and concentrated in specific regions of parameter space, using a prescribed rule. This is the principle behind the neighbourhood sampling algorithm, [Sam99], see also [Sam98] for a more general discussion.

The neighbourhood algorithm (NA) generates multiple models iteratively in parameter space in the following way. Firstly, an initial set of n_s models are randomly generated. In the second step, the n_r models with the lowest misfit models among the most recently generated n_s models are determined. Finally, new n_s models are generated by uniform random walk in the Voronoi cell of each of the n_r chosen models. The algorithm returns to the second step and the process is repeated, see Figure 1. Thus at each iteration, n_r cells are resampled and in each Voronoi cell, n_s/n_r models are generated.

The philosophy behind the algorithm is that the misfit of each of the previous models is representative of the region of its neighbourhood, defined by its Voronoi cell. Therefore at each iteration, new samples are concentrated in the neighbourhoods surrounding the better data-fitting models [Sam99]. Thus the algorithm exploits information obtained in previous models to selectively sample parameter space regions which give better fit to the observed data. Thus the refinement in the model space is achieved through refinement in the parameter space. The degree of this refinement is, by the nature of the algorithm, based on the rank, rather than the actual value, of the misfit between the history and the simulated data.

From the discussion above, it is intuitive that the two parameters which control the algorithm are n_s and n_r . Indeed, these are the only tuning parameters which control the algorithm. The amount of exploration and exploitation performed by the algorithm is dependent on these parameters. See [Sam99] for a full discourse. In summary, as n_s increases, more weight is given to previous samples because the Voronoi cells are updated less frequently. As n_r increases, the algorithm is expected to be more explorative of parameter space and less exploitative, i.e., less local. An attraction of the algorithm is that it evolves naturally in sampling parameter space for better data-fit regions. Thus no external directionality is imposed on the search method.

Using the NA sampling algorithm, we are able to generate multiple reservoir realizations conditioned on history data. In the next section, we describe a specific application of the NA sampling algorithm to data derived from the 10th SPE Comparative Project, [CB01].

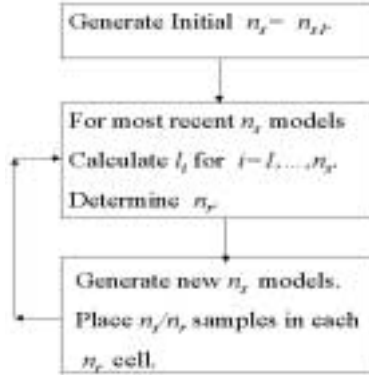


FIGURE 1. NA Sampling Algorithm

1.2. Problem Description. The history matching problem could be stated as follows.

1. Run a fine grid model to obtain fine grid average pressure $P(t)$, produced oil and water volume, $Q_o(t)$ and $Q_w(t)$, respectively, for $0 \leq t \leq t_0$.
2. Use $\Gamma = (P(t), Q_o(t), Q_w(t)), 0 \leq t \leq t_0$ as a stand in for history data.
3. Choose a coarse grid model.
4. Use the coarse grid model to history match the data in 2, by determining relative permeability functions $K_{ro}(S_w)$ and $K_{rw}(S_w)$.
5. Using the maximum likelihood model, forecast $P(t)$, $Q_o(t)$ and $Q_w(t)$ for $t > t_0$.

1.3. Model Description.

1.3.1. *The Fine Grid.* The fine grid is a geostatistical model represented in a three-dimensional domain $[(x, y), z] \in R^2 \times R$, being $[1200 \times 2200 \times 170]$ feet³, and discretized into $[60 \times 220 \times 85]$ cells. The top 70 feet (35 layers) represents the Tarbert formation, and is a representation of a prograding near shore environment. The bottom 100 feet (50 layers) is fluvial, and represents the Upper Ness. The model, which consists of part of a Brent sequence, was originally generated for use in the PUNQ project [FBC⁺99]. Whereas the original model in [FBC⁺99] had a uniform k_v/k_h , the model used here, as in [CB01], has a k_v/k_h of 0.3 in the channels, and 10^{-3} in the background. There are 4 producers in the four corners of the model, each producing at 4000psi bottom hole pressure and a central injector with an injection rate of 5000barrels/day. All wells are vertical, and completed throughout the formation. The fine grid relative permeability functions were given by

$$(1.3) \quad K_{rw}(S_w) = \left(\frac{S_w - S_{wc}}{1 - S_{wc} - S_{or}} \right)^2,$$

$$(1.4) \quad K_{ro}(S_w) = \left(\frac{1 - S_w - S_{or}}{1 - S_{wc} - S_{or}} \right)^2,$$

$$S_{wc} = S_{or} = 0.2.$$

For details of the fluid properties, see [CB01].

1.3.2. *The Coarse Grid.* The coarse grid is an upscaling of the coarse grid and consists of $[5 \times 11 \times 10]$ cells. Here, single scale upscaling is used, [Chr96]. The well positions and completions are the same as in the fine scale model.

1.4. Parameterization. To model oil and water relative permeabilities, relative permeability functions in [Chi81] were adopted. Since the model consists of two formations, different pairs of relative permeability curves were used to represent respectively, the Tarbert and Upper Ness formations. For each of the formations, the relative permeability curves for oil, K_{ro} , and for water, K_{rw} , as functions of water saturation, S_w , are defined by

$$(1.5) \quad K_{ro}(S_w) = K_{ro}(S_{iw})e^{-\alpha R^\zeta},$$

$$(1.6) \quad K_{rw}(S_w) = K_{rw}(S_{or})e^{-\beta R^{-\eta}},$$

$$(1.7) \quad R = \frac{S_w - S_{iw}}{1 - S_{or} - S_w}.$$

This implies that a total of 12 parameters, 6 for each formation, i.e., four empirical parameters $(\alpha, \beta, \zeta, \eta)$ and two rock/fluid parameters $(K_{ro}(S_{iw}), K_{rw}(S_{or}))$, had to be determined.

1.5. Control Parameters. We represent the model parameters by $\gamma = (\alpha, \beta, \zeta, \eta, K_{ro}(S_{iw}), K_{rw}(S_{or}))$, and the simulated values of $P(t)$, $Q_o(t)$ and $Q_w(t)$ respectively, by $F(t)$, $G_o(t)$ and $G_w(t)$. We note that F , G_o and G_w depend non-linearly on γ .

Next, define the following vectors

$$(1.8) \quad \Delta \vec{P} = \vec{P} - \vec{F}(\gamma),$$

$$(1.9) \quad \Delta \vec{Q}_j = \vec{Q}_j - \vec{G}_j(\gamma), \quad j = o, w,$$

then in the least square sense, the measure of misfit, E_T , can be expressed as

$$(1.10) \quad E_T = \left\langle \Delta \vec{P} | C_p^{-1} | \Delta \vec{P} \right\rangle + \sum_{j=o,w} \left\langle \Delta \vec{Q}_j | C_j^{-1} | \Delta \vec{Q}_j \right\rangle.$$

The covariance matrices C_p and C_j contain information about the model and observation errors. In this paper, we make two simplifying assumptions namely,

1. we assume, as is customary, [Siv96], that the measurement errors are independently Gaussian distributed with mean zero and constant variance;
2. that the model errors are negligible compared to the measurement errors.

These assumptions reduce the covariance matrices to constant value, diagonal matrices, whose elements are the variances of the measurement errors. For each given realization $(F(t), G_o(t)$ and $G_w(t))$ of the model, the likelihood L of the observations, given the model parameters γ , is given by

$$(1.11) \quad L(\Gamma|\gamma) \sim e^{-\frac{1}{2}E_T}.$$

L expresses the degree of the likelihood that the history data can be explained by the reservoir model for which the likelihood holds. The refinement criterion for the NA sampling algorithm is based on the rank of the misfit ℓ , defined by

$$(1.12) \quad \ell = -\log L(\Gamma|\gamma) = \frac{1}{2}E_T$$

For the results reported in this paper, based on practical experience, the authors adopted $n_s = 24$ and $n_r = 12$. Thus at each iteration we generate 24 new models

in parameter space by placing 2 new models in the Voronoi cells of the 12 best models.

1.6. Results and Discussions. We history matched on total oil/water volumes produced, and on pressure data, for $t_0 = 300$ days. We then used our maximum likelihood model to forecast future production and pressure profiles for $t = 2000$ days.

Figures 2 and 3 show the maximum likelihood model predictions of the oil and water production rates, compared to the fine grid data. These figures show that our predictions highly reconcile the history data. Figure 4 shows that our average field pressure prediction is almost spot on. Figure 5 shows plots of the relative permeability functions for the fine grid model and for the Tarbert formation.

As mentioned above, during the sampling process, the algorithm concentrates new models in better fitting regions of parameter space. This leads to a refinement in the model space, i.e., with time, we should expect to obtain better fitting models as the sampling evolves. Figure 6 shows a plot of the model indices versus the misfit values. One observes that the misfit values decrease as we move from lower to higher model indices.

Factors that may affect the accurate prediction of our maximum likelihood model include, among other, the time t_0 for the history matching period. If the history matching is conditioned on insufficient information, even though the maximum likelihood model will be accurate for $t \leq t_0$, the essential fine grid flow properties may not be adequately captured. Hence the fine grid performance predictions will be highly inaccurate. As we introduce more dynamic data into the history matching period, we expect to achieve better predictive ability of the maximum likelihood model. It is known that an upper bound for t_0 exists, above which the introduction of more data into the history matching period will give no significant effect. Thus the additional computational work and CPU time involved in incorporating the excess data is uncalled for.

Further, the two major sources of uncertainty affecting our predictions are model and observation errors. In this paper, we have assumed that the observation errors are independently Gaussian distributed, and that the model errors are negligible compared with the observational errors. In reality however, the rates and pressure data are time series, which co-vary with time. A more accurate approach will be to introduce the covariance relationships through the covariance matrices. The assumption that the model errors are negligible may also be an oversimplification, [GhL⁺01]. These issues will be addressed elsewhere.

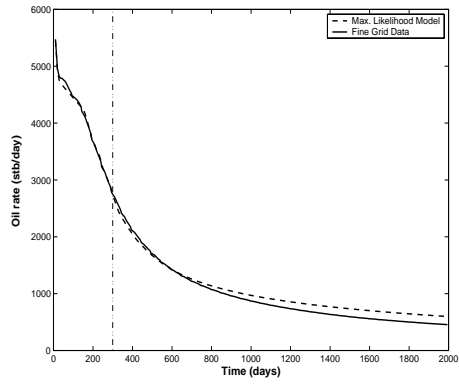


FIGURE 2. Oil rate

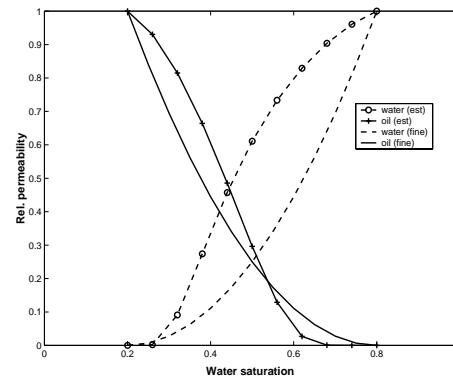


FIGURE 5. Rel. perm. functions

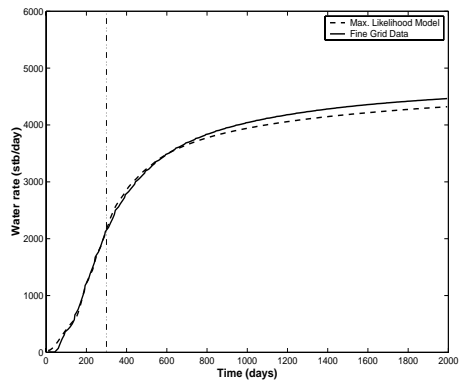


FIGURE 3. Water rate

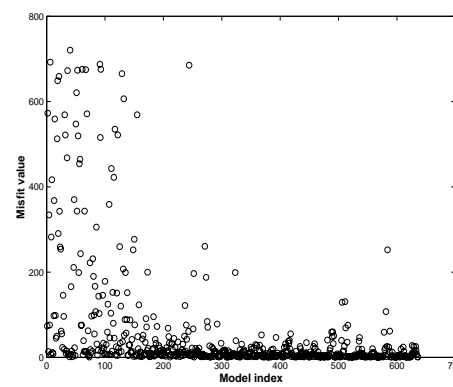


FIGURE 6. Model refinement

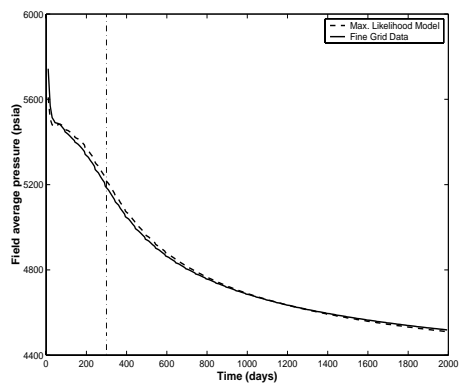


FIGURE 4. Pressure

1.7. Conclusions. We have demonstrated a new stochastic method for generating multiple history matching models. The method, though conceptually simple (depending on two parameters), exhibits a complex self-adaptive behavior in searching parameter space.

The method differs from other methods for history matching in the following way. Rather than seeking a single global optimum, the algorithm selectively samples parameter space to derive an ensemble of models that share the common property of fitting the observed data to some degree of accuracy. This approach is reasonable since the inverse problem is ill-posed. Specifically, unlike other stochastic methods, the algorithm performs a guided search in parameter space by using information derived from the complete ensemble of previously generated models. Hence externally imposed directionality of the search process is avoided.

Our results show that the maximum likelihood model is highly accurate in forecasting the fine grid performance.

Acknowledgement

We thank BP for funding this work as part of the “Prediction Under Uncertainty” project at Heriot-Watt University. Thanks are also due to Schlumberger-GeoQuest for use of the ECLIPSE reservoir simulator.

Nomenclature.

C	Covariance matrix for pressure
E	Error function
F	Simulated pressure
G	Simulated oil/water rate
K	Relative permeability
k	Absolute permeability
L	Likelihood function
ℓ	Misfit function
n_s	Number of models generated per iteration step
n_r	Number of models to resample
P	History pressure
Q	History oil/water rate
R	Correlation parameter
S	Saturation
t	time

Greek.

$\alpha, \beta, \zeta, \eta$	Empirical coefficients
Δ	Difference
δ	Euclidean distance
Γ	Set of observed/history data
γ	Set of model parameters

Subscript / Superscript.

h	Horizontal
iw	Irreducible water
j	Indicator index for oil($j = o$) and water($j = w$)
o	Oil
or	Residual oil
p	Pressure
ro	Relative to oil
rw	Relative to water
T	total
v	Vertical
w	Water
wc	Connate water

References

- [Aur91] F Aurenhammer, *Voronoi Diagrams - A Survey of a Fundamental Geometric Data Structure*, ACM Computing Surveys **23** (September 1991), no. 3, 345–405.
- [CB01] M A Christie and M J Blunt, *Tenth SPE Comparative Solution Project: A Comparison of Upscaling Techniques*, SPE (2001), no. SPE 66599, 1–9.
- [CGSS96] P S Craig, M Goldstein, A H Seheult, and J A Smith, *Bayes Linear Strategies for Matching Hydrocarbon Reservoir History*, BAYESIAN STATISTICS (1996), no. 5, 69–95.
- [Chi81] G L Chierici, *Novel Relations for Drainage and Imbibition Relative Permeabilities*, SPE (1981), no. SPE 10165, 1–10.
- [Chr96] M A Christie, *Upscaling for Reservoir Simulation*, JPT (November 1996), no. SPE 37324, 1004–1008.
- [Chr01] M A Christie, *Flow in Porous Media - Scale Up of Multiphase Flow*, Current Opinions in Colloid & Interface Science (2001), no. 6, 236–241.
- [Dak94] L P Dake, *The Practice of Reservoir Engineering*, ELSEVIER, The Netherlands, 1994.
- [FBC⁺99] F J T Floris, M D Bush, M Cuypers, F Roggero, and A R Syversveen, *Comparison of production forecast uncertainty quantification methods- an integrated study*, Conference paper- 1st Conference on Petroleum Geostatistics (1999), 1–20, Toulouse, France.
- [GHhL⁺01] J Glimm, S Hou, Y h Lee, D Sharp, and K Ye, *Prediction of Oil Production with Confidence Intervals*, SPE (2001), no. SPE 66350, 1–5.
- [LBA⁺99] O J Lépine, R C Bissell, S I Aanonsen, I C Pallister, and J W Barker, *Uncertainty Analysis in Predictive Reservoir Simulation Using Gradient Information*, SPE Journal (1999), no. SPE 57594, 251–259.
- [OBS92] A Okabe, B Boots, and K Sugihara, *Spatial Tessellations - Concepts and Applications of Voronoi diagrams*, JOHN WILEY & SONS, ENGLAND, 1992.
- [PP99] R C M Portella and F Prais, *Use of Automatic History Matching and Geostatistical Simulation to Improve Production Forecast*, SPE (1999), no. SPE 53976, 1–10.
- [Sam98] M Sambridge, *Exploring multidimensional landscapes without a map*, Inverse Problems (1998), no. 14, 427–440.
- [Sam99] M Sambridge, *Geophysical inversion with a neighbourhood algorithm -1. Searching a parameter space*, Geophysics J. Int. (1999), no. 138, 479–94.
- [Siv96] D S Sivia, *Data Analysis - A Bayesian Tutorial*, CLAREDON PRESS, Oxford, 1996.
- [WDC99] X-H Wen, C V Deutsch, and A S Cullick, *High Resolution Reservoir Models Integrating Multiple-Well Production Data*, SPE (1999), no. SPE 38728, 115–129.
- [YW91] P H Yang and A T Watson, *A Bayesian Methodology for Estimating Relative Permeability Curves*, SPE (1991), no. SPE 18531, 79–92.

DEPT. OF PETROLEUM ENGINEERING, HERIOT-WATT UNIVERSITY, EDINBURGH, EH14 4AS,
UK

E-mail address: sam.subbey@pet.hw.ac.uk

DEPT. OF PETROLEUM ENGINEERING, HERIOT-WATT UNIVERSITY, EDINBURGH, EH14 4AS,
UK

E-mail address: mike.christie@pet.hw.ac.uk

INSTITUTE OF ADVANCED STUDIES, AUSTRALIAN NATIONAL UNIVERSITY, CANBERRA, ACT
0200, AUSTRALIA

E-mail address: malcolm@rses.anu.edu.au