A Bayesian Petrophysical Decision Support System for Estimation of Reservoir Compositions*

Willem Burgers,Wim Wiegerinck,Bert Kappen †
Mirano Spalburg ‡

May 10, 2010

Abstract

The exploration for oil and gas requires real-time petrophysical expertise to interpret measurement data acquired in boreholes and to recommend further steps. High time pressure and the far reaching nature of these decisions, as well as the limited opportunity to gain in depth petrophysical experience suggests that a decision support system that can aid the petrophysicist will be very useful.

In this paper we describe a Bayesian approach for obtaining compositional estimates that combines expert knowledge with information obtained from measurements. We define a prior model for the compositional volume fractions and observation models for each of the measurement tools. Both prior and observation models are based on domain expertise. These models are combined in a joint probability model. To deal with the nonlinearities in the model, Bayesian inference is implemented by using the hybrid Monte Carlo algorithm.

In the system, tool measurement values can entered and the posterior probability distribution of the compositional fractions can be obtained by applying Bayes' rule. Bayesian inference is also used for optimal tool selection, using conditional entropy to select the most informative tool to obtain better estimates of the reservoir.

Reliability and consistency of the method is demonstrated by inference on synthetically generated data.

1 Introduction

Oil and gas reservoirs are located in the earth's crust at depths of several kilometers, and when located offshore, in water depths of a few meters to a few kilometers. Consequently, the gathering of critical information such as the presence and type of hydrocarbons, size of the reservoir and the physical properties of the reservoir such as the porosity of the rock and the permeability is a key activity in the oil and gas industry.

*Expert Systems with Applications (2010). DOI:10.1016/j.eswa.2010.04.092 (In press)

†SNN Adaptive Intelligence, Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, Geert Grooteplein 21, 6525 EZ Nijmegen, The Netherlands

‡Shell International Exploration and Production B.V. Kessler Park 1, 2288 GS Rijswijk, The Netherlands

Pre-development methods to gather information on the nature of the reservoirs range from gravimetric, 2D and 3D seismic to the drilling of exploration and appraisal boreholes. Additional information is obtained while a field is developed through data acquisition in new development wells drilled to produce hydrocarbons, time-lapse seismic surveys and inwell monitoring of how the actual production of hydrocarbons affects physical properties such as the pressure and temperature. The purpose of information gathering is to decide which reservoirs can be developed economically, and how to adapt the means of development best to the particular nature of a reservoir.

The early measurements acquired in exploration, appraisal and development boreholes are a crucial component of the information gathering process. These measurements are typically obtained from tools that can be included in the borehole drilling equipment or from open/cased hole logging. The range of possible measurements varies depending on the type of logging. Some options, such as coring, are very expensive and may even risk other data acquisition options. In general acquiring all possible data imposes too great an economic burden on the exploration, appraisal and development. Hence data acquisition options must be exercised carefully bearing in mind the learnings of already acquired data and general hydrocarbon field knowledge. Also important is a clear understanding of what data can and cannot be acquired later and the consequences of having an incorrect understanding of the nature of a reservoir on the effectiveness of its development.

Making the right data acquisition decisions, as well as the best interpretation of information obtained in boreholes forms one of the principle tasks of petrophysicists. The efficiency of a petrophysicsist executing her/his task is substantially influenced by the ability to gauge her/his experience to the issues at hand. Efficiency is hampered when a petrophysicists experience level is not yet fully sufficient and by the rather common circumstance that decisions to acquire particular types of information or not must be made in a rush, at high costs and shortly after receiving other information that impact on that very same decision. Mistakes are not entirely uncommon and almost always painful. In cases, non essential data is obtained at the expense of extremely high cost, or essential data is not obtained at all; causing development mistakes that can jeopardize the amount of hydrocarbon recoverable from a reservoir and induce significant cost increases.

The overall effectiveness of petrophysicists is expected to improve in case a Bayesian net [Pearl, 1988] constructed and populated with petrophysical relationships and knowledge is available and can be used as a decision support system (DSS). In practice a DSS can increase the petrophysicists awareness of low probability but high impact cases and alleviate some of the operational decision pressure. In the longer run regularly updated DSSs may serve to capture and disseminate petrophysical experience and knowledge while also other petroleum engineering experts such as geologists, and reservoir engineers may start to use a Bayesian DSS for their purposes.

In this paper we describe a Bayesian net for the estimation of compositional volume fractions in a reservoir on the basis of logging data. It is an extension of the work described in Spalburg [2004].
The paper is organized as follows. In section 2 we describe the ideas of probabilistic modeling and Bayesian inference for the estimation of compositional volume fractions on the basis of measurements. In its subsections we will de-

scribe the model and the inference steps via a hybrid Monte Carlo sampler in more detail. Some experiments using synthetic data illustrating the Bayesian inference method are described in section 3. In section 4, we describe how the sampling method can be used to decide on the most informative next measurements. In section 3, we assess the reliability and consistency of the method by inference on synthetically generated data, and we end with a discussion and conclusions in section 5.

2 Probabilistic modeling

The primary aim of the model is to estimate the compositional volume fractions of a reservoir on the basis of borehole measurements. Due to incomplete knowledge, limited amount of measurements and noise in the measurements, there will be uncertainty in the volume fractions. We will use Bayesian probability theory to deal with this uncertainty.

The starting point is a model for the probability distribution $\mathrm{P}(\vec{v}, \vec{m})$ of the compositional volume fractions $\vec{v}$ and borehole measurements $\vec{m}$. A causal argument *"The composition is given by the (unknown) volume fractions, and the volume fractions determine the distribution measurement outcomes of each of the tools"* leads us to a Bayesian net formulation of the probabilistic model,

$$\mathrm{P}(\vec{v}, \vec{m}) = \prod_{i=1}^{Z} \mathrm{P}(m_i|\vec{v})\, \mathrm{P}(\vec{v})\,.$$

In this model, $\mathrm{P}(\vec{v})$ is the so-called *prior*, the prior probability distribution of volume fractions before having seen any data. In principle, the prior encodes the generic geological and petrophysical knowledge and beliefs [Spalburg, 2004]. The factor $\prod_{i=1}^{Z} \mathrm{P}(m_i|\vec{v})$ is the *observation model*. The observation model relates volume fractions $\vec{v}$ to measurement outcomes m_i of each of the Z tools i. The observation model assumes that *given* the underlying volume fractions, measurement outcomes of the different tools are independent. Each term in the observation model gives the probability density of observing outcome m_i for tool i given that the composition is $\vec{v}$. Now given a set of measurement outcomes $\vec{m}^o$ of a subset Obs of tools, the probability distribution of the volume fractions can be updated in a principled way by applying *Bayes' rule*,

$$\mathrm{P}(\vec{v}|\vec{m}^o) = \frac{\prod_{i \in Obs} \mathrm{P}(m_i^o|\vec{v})\, \mathrm{P}(\vec{v})}{\mathrm{P}(\vec{m}^o)}\,. \tag{1}$$

The updated distribution is called the *posterior* distribution. The constant in the denominator $\mathrm{P}(\vec{m}^o) = \int_{\vec{v}} \prod_{i \in Obs} \mathrm{P}(m_i^o|\vec{v})\, \mathrm{P}(\vec{v})\, \mathrm{d}\vec{v}$ is called the *evidence*.

The remainder of this section describes the prior and observation model, as well as the sampling method that we used to obtain the posterior.

2.1 Prior

The model assumes that the reservoir at the given depth is composed of K given minerals and fluids. The volume fraction of mineral j is assumed to have a definite but unknown value v_j which is between 0 and 1. The model assumes that the K given minerals and fluids occupy the whole volume of the composite,

$\sum_{j=1}^{K} v_k = 1$. In other words, the vector of volume fractions $\vec{v}$ is constrained to the K-part simplex, defined by $\mathbb{S}^K \equiv \left\{ \vec{v} \in \mathbb{R}_+^K \middle| \sum_{j=1}^{K} v_j = 1 \right\}$, where $\mathbb{R}_+$ is the space of reals > 0 (for the *closed* simplex $0 \leq v_j \leq 1$, but to ensure that the division operator is defined everywhere we use the *open* simplex $0 < v_j < 1$). As a shorthand notation, we will use the symbol φ to denote a summation over a set of compositional parts; $\varphi : \mathbb{S}^K \mapsto [0,1] \subset \mathbb{R}$, defined by $\varphi_\chi(\vec{v}) = \sum_{\forall j \in \chi} v_j$, where χ denotes the set of minerals and fluids to include in the summation.
The fluids and minerals (see table 1) are clustered into three basic geological groups: the non-reservoir minerals $\mathcal{N}$, matrix minerals $\mathcal{M}$ and fluid minerals $\mathcal{F}$. The number of minerals in each group is indicated by the symbol #. The number of distinct groups in the simplex is $\#\Omega = \Theta(\#\mathcal{N}) + \Theta(\#\mathcal{M}) + \Theta(\#\mathcal{F})$, where $\Theta : \mathbb{Z} \mapsto \{0,1\}; \Theta(n) = 1 \iff n > 0$. The model allows to decrease the number of minerals and fluids K by excluding these from the prior. In such case, a mineral group can be completely absent. Most commonly applied prior

	Name	Groups
1	Shale	$\mathcal{N}$
2	Coal	$\mathcal{N}$
3	Quartz	$\mathcal{M}$, $\mathcal{R}$
4	Dolomite	$\mathcal{M}$, $\mathcal{R}$
5	Calcite	$\mathcal{M}$, $\mathcal{R}$
6	Clay (wet)†	$\mathcal{M}$, $\mathcal{R}$
7	Halite	$\mathcal{M}$, $\mathcal{R}$
8	Pyrite	$\mathcal{M}$, $\mathcal{R}$
9	Siderite	$\mathcal{M}$, $\mathcal{R}$
10	Bound-water	$\mathcal{F}$, $\mathcal{R}$
11	Free-water	$\mathcal{F}$, $\mathcal{R}$
12	Oil	$\mathcal{F}$, $\mathcal{R}$
13	Gas	$\mathcal{F}$, $\mathcal{R}$

Table 1: Minerals and fluids included in the simplex. † wet clay is composed of dry clay and claybound water. $\mathcal{N}$: non reservoir minerals, $\mathcal{M}$: matrix minerals, $\mathcal{F}$: fluids, $\mathcal{R}$: reservoir minerals.

distributions (notably Gaussian) have support $\mathbb{R}$ or $\mathbb{R}_+$ and require truncation techniques to be fitted to the simplex. A more natural alternative is using a distribution that is by itself bound to this space. The Dirichlet distribution [MacKay, 2003]

$$\mathrm{P}(\vec{v}|\alpha,\vec{\mu}) \propto \prod_{j}^{K} v_j^{\alpha\mu_j - 1} \delta\left(1 - \sum_{i}^{K} v_i\right), \tag{2}$$

is a convenient candidate. The two parameters $\alpha \in \mathbb{R}_+$ (shape) and $\vec{\mu} \in \mathbb{S}^K$ (vector of means) can be used to fine-tune the prior to our liking. (The delta function — which ensures that the simplex constraint holds — is put here for clarity, but is in fact redundant since $\vec{v} \in \mathbb{S}^K$ and will be omitted in the remainder of this paper.)

An example of information we have included in the prior is the porosity. A generic assumption is that the porosity

$$\phi : \mathbb{S}^K \mapsto [0,1] \subset \mathbb{R}; \quad \phi\left(\vec{v}\right) = \frac{\varphi_{\text{fluids}}}{\varphi_{\text{reservoir}}}$$

has a uniform a-priori distribution up to the percolation limit (ϕ_{P}) after which the probability decays rapidly, as shown in figure 1.

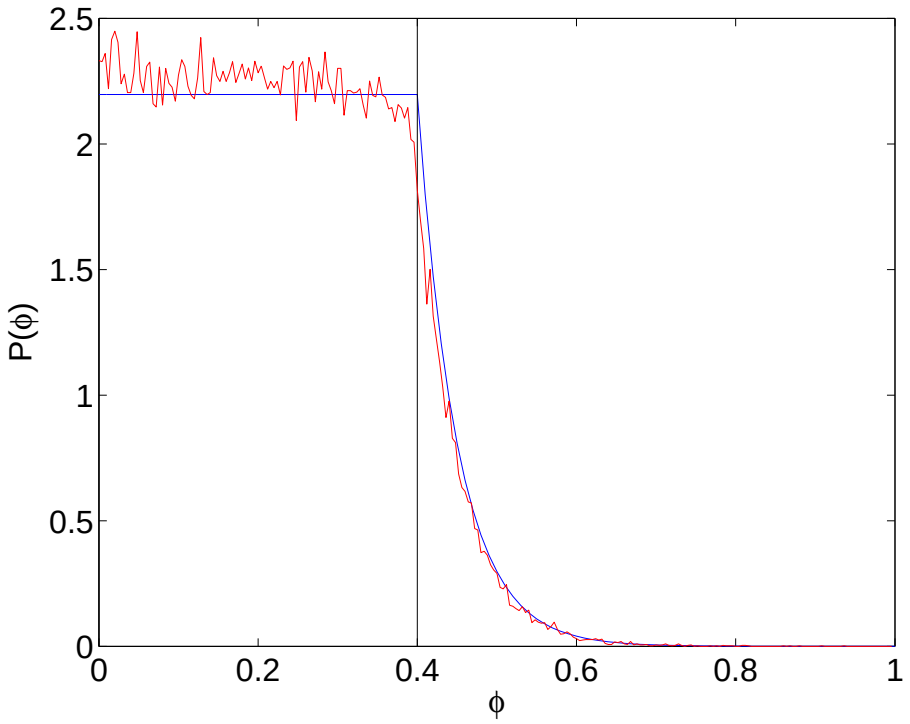

Figure 1: Porosity prior for $\phi_{\text{P}} = 0.4$ and $\tau = 0.05$; theoretical curve and one obtained from sampling.

We can model this as follows. We choose the Dirichlet parameters $\alpha = \#\Omega$ and $\mu_\kappa = 1/\left(\#\Omega\#\kappa\right)$ where κ denotes the mineral group. This leads to the following Dirichlet distribution

$$\mathrm{P}_0\left(\vec{v}\right) \propto \prod_{\forall k \in \mathcal{N}} v_k^{\frac{1-\#\mathcal{N}}{\#\mathcal{N}}} \prod_{\forall j \in \mathcal{F}} v_j^{\frac{1-\#\mathcal{F}}{\#\mathcal{F}}} \prod_{\forall i \in \mathcal{M}} v_i^{\frac{1-\#\mathcal{M}}{\#\mathcal{M}}} \tag{3}$$

which induces $\mathrm{P}\left(\phi\right) = \mathcal{U}\left(0,1\right)$ by construction (see appendix A). The percolation limit is included by multiplying (3) by

$$P(\phi;\phi_{\text{P}},\tau) \propto \begin{cases} 1 & \iff \quad \phi \leq \phi_{\text{P}} \\ \exp\left(-\frac{\phi-\phi_{\text{P}}}{\tau}\right) & \iff \quad \phi > \phi_{\text{P}} \end{cases}$$

where parameter τ controls the decay strength. The resulting prior is

$$P_{\phi_{\text{P}}}(\vec{v}) \propto P_0(\vec{v}) P(\phi(\vec{v});\phi_{\text{P}},\tau)\,. \tag{4}$$

2.2 Observation model

The other important term in the Bayesian net is the observation model. This term models our belief in the outcome of the measurements $\vec{m} \in \mathbb{R}^{\mathrm{Z}}$ given the actual composition $\vec{v}$ of the reservoir. In other words, given a certain composition this defines the probability distribution over the measurement values. The vector $\vec{m}$ consists of measurement outcomes m_j, one for each tool. As stated earlier, the model assumes that the outcomes of these different tool measurements are independent given the compositional volume fractions $\vec{v}$, so

$$P(\vec{m}|\vec{v}) = \prod_{j=1}^{\mathrm{Z}} P(m_j|\vec{v}) \tag{5}$$

For each of the measurement tools, we assume additive Gaussian distributed measurement noise, i.e. we assume

$$m_j = f_j(\vec{v}) + \xi_j \; . \tag{6}$$

The functions $f_j : \mathbb{S}^K \mapsto \mathbb{R}$ are the deterministic tool values [Spalburg, 2004]. These are the idealized noiseless measurement outcomes. They are modeled by tool-specific mathematical functions, which are based on the physics of the measurement tools. These functions are provided by domain experts. A more detailed description of these functions, however, is beyond the scope of this paper. The noise ξ_j is additive and Gaussian distributed with a tool specific variance σ_j^2. These variances are also provided by domain experts. Where necessary, a log transform was applied to turn measurements with typical log-normal distributed multiplicative noise into quantities with additive Gaussian noise. So, the observational probability model can be written as

$$\mathrm{P}\left(\vec{m}|\vec{v},\vec{\sigma}\right) \propto \prod_{j=1}^{\mathrm{Z}} \exp\left(-\frac{(m_j - f_j(\vec{v}))^2}{2\sigma_j^2}\right) \; . \tag{7}$$

2.3 Bayesian inference

The next step is given a set of observations $\{m_i^o\}$, $i \in \mathrm{Obs}$, to compute the posterior distribution. If we were able to find an expression for the evidence term, i.e. for the marginal distribution of the observations $\mathrm{P}\left(\vec{m}^o\right) = \int_{\vec{v}} \prod_{i \in \mathrm{Obs}} \mathrm{P}\left(m_i^o|\vec{v}\right) \mathrm{P}\left(\vec{v}\right) \mathrm{d}\vec{v}$ then the posterior distribution (1) could be written in closed form and readily evaluated. Unfortunately $\mathrm{P}\left(\vec{m}^o\right)$ is intractable and a closed-form expression does not exist. In order to obtain the desired compositional estimates we therefore have to resort to sampling methods.

The goal of any sampling procedure is to obtain a set of N samples $\{x_i\}$ that come from a given (but maybe intractable) distribution π. Using these samples we can approximate expectation values $\langle A \rangle$ of a functions $A(x)$ according to

$$\langle A \rangle = \int_{\mathcal{A}_x} A(x)\pi(x)\mathrm{d}x \approx \frac{1}{N}\sum_{i=1}^{N} A(x_i) \tag{8}$$

For instance, if we take $A(x) = x$, the approximation of the mean $\langle x \rangle$ is the sample mean $\frac{1}{N}\sum_{i=1}^{N} x_i$.

	Name
1	GR
2	GR-K
3	GR-U
4	GR-Th
5	Density
6	Photo-Electric
7	Neutron
8	Pulsed Neutron Capture
9	NMR-Bound
10	NMR-Claybound-water
11	NMR-Free
12	Flushed Zone Resistivity
13	Deep zone Resistivity
14	Sonic

Table 2: Petrophysical Tools Modeled.

An important class of sampling methods are the so-called Markov Chain Monte Carlo (MCMC) methods [Neil, 1993]. In MCMC sampling a Markov chain is defined that has an equilibrium distribution π, in such a way that (8) gives an good approximation when applied to a sufficiently long chain $x_1, x_2, \ldots, x_N$. To make the chain independent of the initial state x_0, a burn-in period is often taken into account. This means that one ignores the first $M \ll N$ samples that come from intermediate distributions and begins storing the samples once the system has reached the equilibrium distribution π.

In our application we use the hybrid Monte Carlo (HMC) sampling algorithm [Duane et al., 1987]. HMC is a powerful class of MCMC methods that are designed for problems with continuous state spaces, such as we consider in this paper. HMC can in principle be applied to any noise model with a continuous probability density, so there is no restriction to Gaussian noise models. HMC uses Hamiltonian dynamics in combination with a Metropolis [Metropolis et al., 1953] acceptance procedure to find regions of higher probability. This leads to a more efficient sampler than a sampler that relies on random walk for phase space exploration. HMC also tends to mix more rapidly than the standard Metropolis Hastings algorithm.

3 Simulations

The performance of the method relies heavily on the quality of the sampler. Therefore we looked at the ability of the system to estimate the composition of a (synthetic) reservoir and the ability to reproduce the results. For this purpose, we set the composition to a certain value $\vec{v}^*$. We apply the observation model to generate measurements $\vec{m}^o$. Then we run HMC to obtain samples from the posterior $P(\vec{v}|\vec{m}^o)$. Consistency is assessed by comparing results of different runs to each other and by comparing them with the "ground truth" $\vec{v}^*$. Here, we take

$v^*_{\text{quartz}} = 0.3$, $v^*_{\text{wet-clay}} = 0.5$, $v^*_{\text{freewater}} = 0.12$ and $v^*_{\text{oil}} = 0.08$ from which we generated a set of observations $\vec{m}^o = \{m^o_j\}$. With $\vec{m}^o$ as input, ten simulations were ran. From the resulting samples the means (μ) and (asymmetric) error bars ($\mu - \sigma_1, \mu + \sigma_2$) are computed for each mineral. These values are drawn in figure 2 (the composition $\vec{v}^*$ is indicated by asterisks).

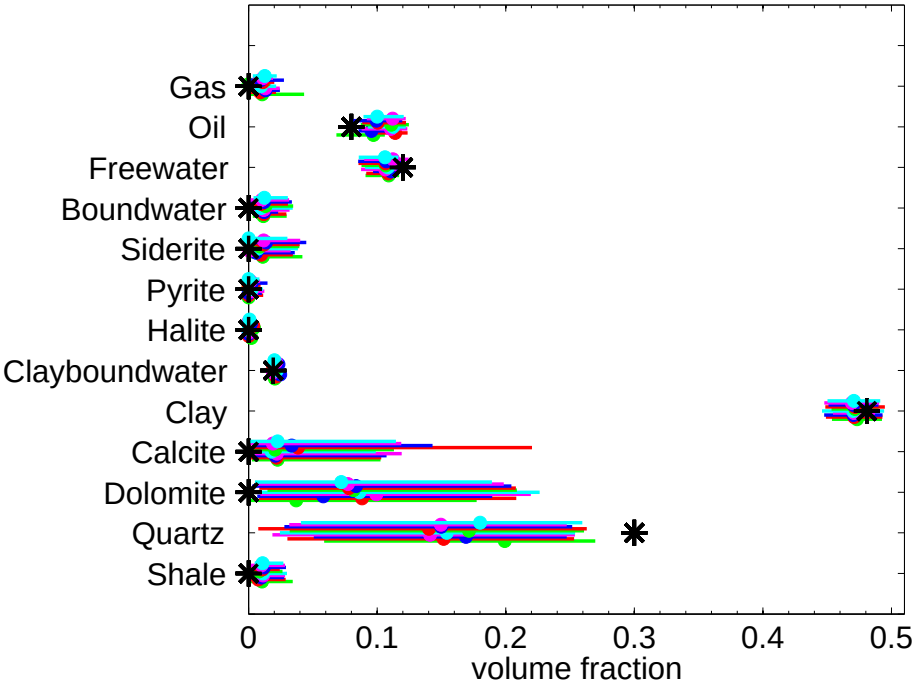

Figure 2: Consistency: 10 runs of 131.072 samples each. Burnin 1000 samples. Flat porosity prior. For each run, means (bold dots) and (asymmetric) error bars of each mineral are plotted. $*$: values used to generate synthetic measurements. Note the large deviation for Quartz.

The estimates (figure 2) based on synthetic measurements are within one error bar from the actual composition, with the exception of quartz. This is caused by the fact that the tools employed are incapable of distinguishing between quartz and dolomite (and calcite); during sampling states with quartz and dolomite are visited alternatively (figure 3 top). The resulting distribution is multimodal and is therefore described poorly by a mean value and error bars, but better by a distribution. Multimodality suggests that there are other reservoir compositions that lead to $\vec{m}^o$ under the current error model.

Figure 2 also shows that the ten chains converge to the same result, (all chains start at a different random position on the simplex).

The addition of observations reduce the uncertainty about the composition of a reservoir. To illustrate this three scenarios have been sampled. Figure 4 shows likely compositions when only the prior knowledge (as given in equation 4) is present. This effectively limits possible solutions to $\varphi_{\text{matrix}} > 0.6$. When an observation for an acoustic measurement is added, the space of likely configurations is reduced to include this new information, as shown in figure 5.

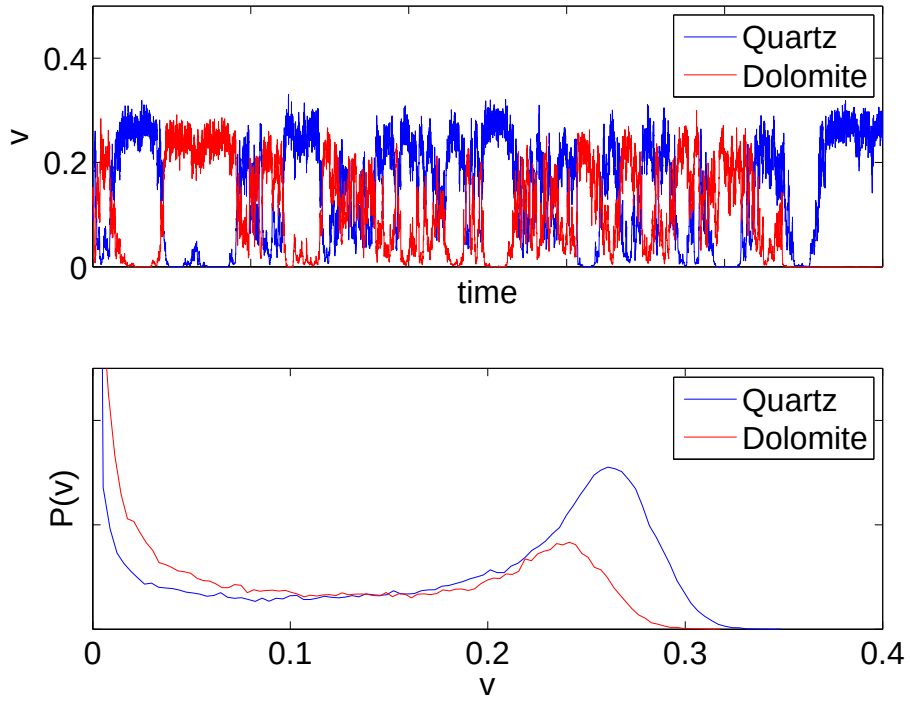

Figure 3: Diagrams for quartz and dolomite. Top: time traces (10 000 time steps) showing the mutually exclusive behavior, bottom: resulting multimodal probability distribution. The two peaks indicate the two main states, the valley corresponds to transient behavior between those two states.

Figure 6 shows that the inclusion of a resistivity tool reduces the space even further.

Results of simulations with other values of $\vec{v}^*$ (not reported here) confirm that the sampler generates reproducible results, consistent with the underlying compositional vector. In these simulations, we assumed that the observations model to generate measurement data (the generating model) is equal to the observation model used to apply Bayes' rule (the inference model). We also performed simulations where they are different, in particular in their assumed variance. We found that the sampler is robust to cases where the variance of the generating model is smaller than the variance of the inference model. In the cases where the variance of the generating model is bigger, we found that the method is robust up to differences of a factor 10. After that we found that the sampler suffered severely from local minima, leading to irreproducible results.

4 Decision Support

Suppose that we have obtained a subset of measurement outcomes $\vec{m}^o$, yielding a distribution $\mathrm{P}(\vec{v}|\vec{m}^o)$. One may subsequently ask the question which tool t should be deployed next in order to gain as much information as possible?

When asking this question, one is often interested in a specific subset of

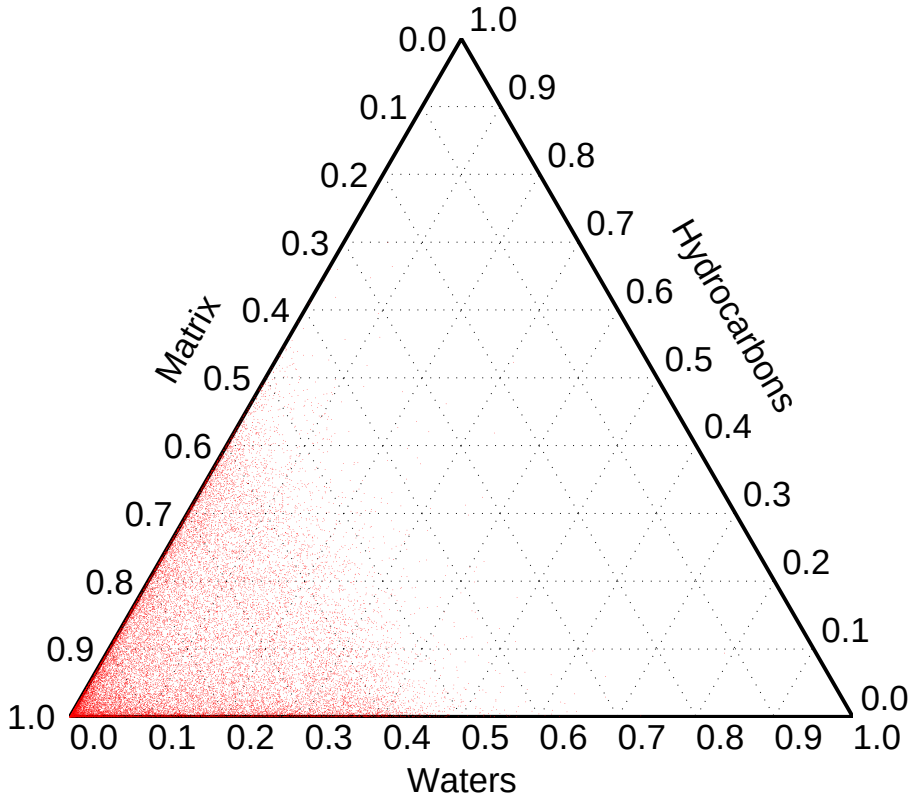

Figure 4: Likely compositions when only the prior is known.

minerals and fluids. Here we assume this interest is actually in one specific component u. The question then reduces to selecting the most informative tool(s) t for a given mineral u.

We define the informativeness of a tool as the expected decrease of uncertainty in the distribution of v_u after obtaining a measurement with that tool. Usually, entropy is taken as a measure for uncertainty [MacKay, 2003], so a measure of informativeness is the expected entropy of the distribution of v_u after measurement with tool t,

$$\begin{aligned}\langle \mathrm{H}_{u,t}|\vec{m}^o\rangle \equiv -\int \mathrm{P}\left(m_t|\vec{m}^o\right)\int \mathrm{P}\left(v_u|m_t,\vec{m}^o\right)\\ \times \log_2\left(\mathrm{P}\left(v_u|m_t,\vec{m}^o\right)\right)\mathrm{d}v_u\mathrm{d}m_t\end{aligned} \tag{9}$$

Note that the information of a tool depends on the earlier measurement results since the probabilities in (9) are conditioned on $\vec{m}^o$.

The most informative tool for mineral u is now indentified as that tool t^* which yields in expectation the lowest entropy in the posterior distribution of v_u:

$$t^*_{u|\vec{m}^o} = \operatorname*{argmin}_t \langle \mathrm{H}_{u,t}|\vec{m}^o\rangle$$

In order to compute the expected conditional entropy using HMC sampling

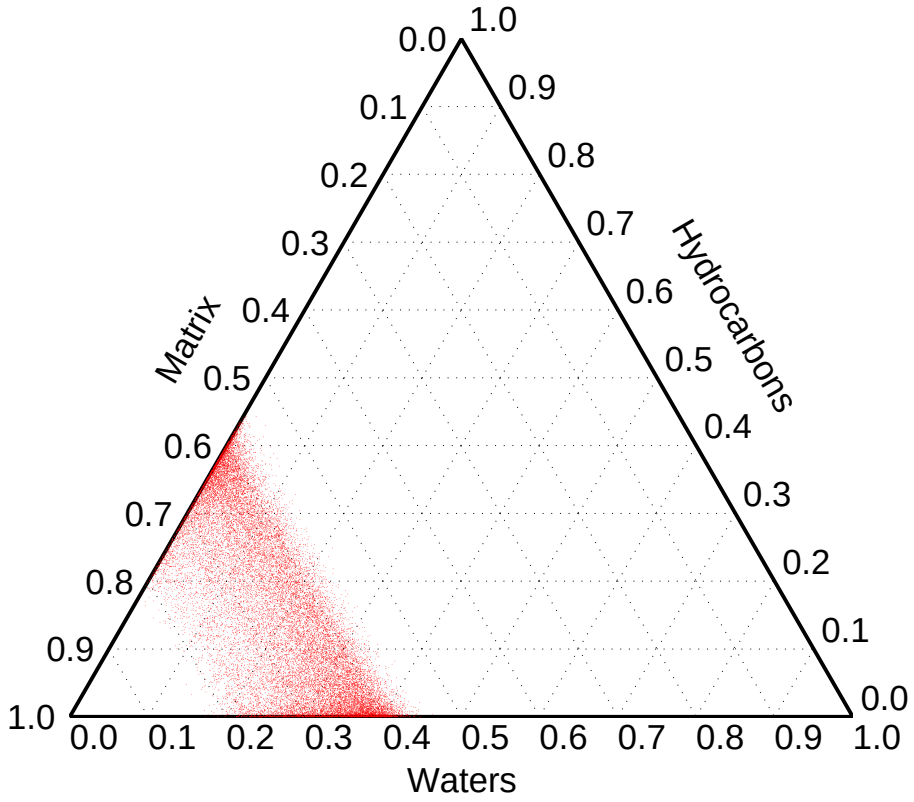

Figure 5: Likely compositions when the observation of an acoustic tool is included.

methods, we first rewrite the expected conditional entropy (9) in terms of quantities that are conditioned only on the measurement outcomes $\vec{m}^o$,

$$\begin{aligned}\langle \mathrm{H}_{u,t}|\vec{m}^o\rangle = &-\int\int \mathrm{P}\left(v_u, m_t|\vec{m}^o\right) \\ &\times \log_2\left(\mathrm{P}\left(v_u, m_t|\vec{m}^o\right)\right) \mathrm{d}v_u \mathrm{d}m_t \\ &+ \int \mathrm{P}\left(m_t|\vec{m}^o\right) \int \log_2\left(\mathrm{P}\left(m_t|\vec{m}^o\right)\right) \mathrm{d}m_t\end{aligned} \tag{10}$$

Now the HMC run yields a set $V = \left\{v_1^j, v_2^j, \ldots, v_K^j\right\}$ of compositional samples (conditioned on $\vec{m}^o$). We augment these by a set $M = \left\{m_1^j = f_1(\vec{v}^j) + \xi_1^j, \ldots, m_{\mathrm{Z}}^j = f_{\mathrm{Z}}(\vec{v}^j) + \xi_{\mathrm{Z}}^j\right\}$ of synthetic tool values generated from these samples (which are indexed by j) by applying equation 6. Subsequently, discretized joint probabilities $\mathrm{P}\left(v_u, m_t|\vec{m}^o\right)$ are obtained via a two dimensional binning procedure over v_u and m_t for each of the potential tools t. The binned versions of $\mathrm{P}\left(v_u, m_t|\vec{m}^o\right)$ (and $\mathrm{P}\left(m_t|\vec{m}^o\right)$) can be directly used to approximate the expected conditional entropy using a discretized version of equation 10.

We illustrate the idea of the decision support with the following simulated example. In this example, we are interested in the most informative tool for the

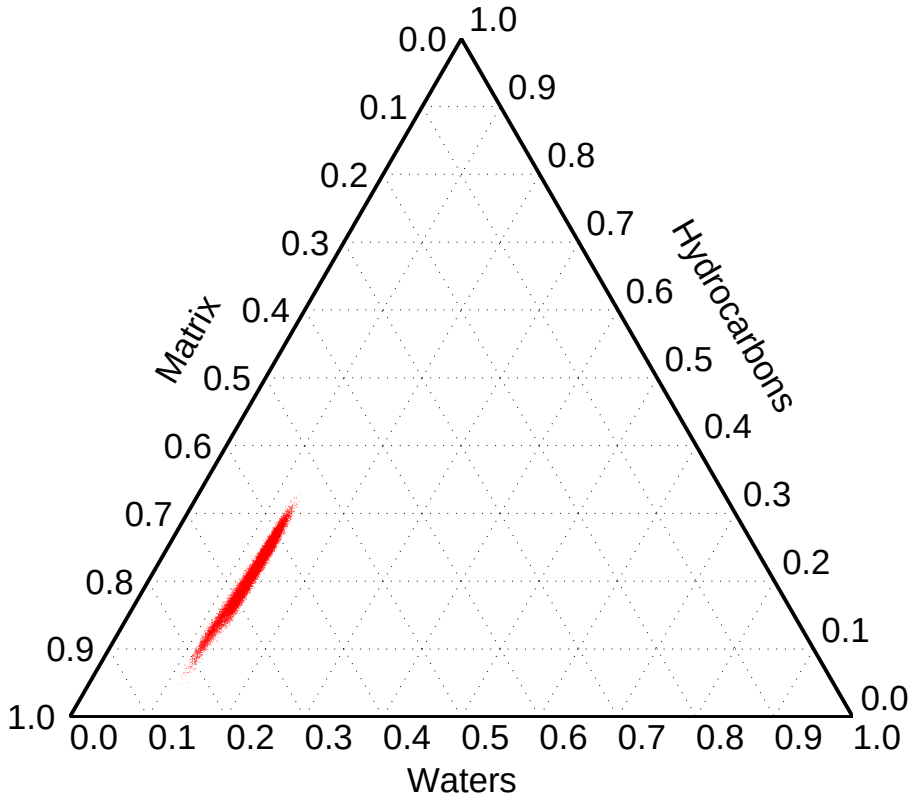

Figure 6: Likely compositions when the observation of a sonic tool and resistivity tool are included.

mineral u = 'oil'. We assume that we have no previous measurements, i.e. we use the prior distribution. By applying the procedure described above, we find that the most informative tool for oil is 'NMR-free' given prior information only.

Now the question may be in what sense a typical measurement with the 'NMR-free' tool differs from other candidate tools. In the following we therefore simulate measurements with different tools that are responsive to oil, and compare the resulting posteriors, illustrating—with hind-sight—the information content of the different tools. To simulate measurements we first draw at random a mineral composition $\vec{v}$ from the prior. This composition is assumed to be the ground truth. Next, this composition is used to synthesize observations $m_t = f_t(\vec{v}) + \xi_t$ for the different tools t , being 'Sonic', 'NMR-free', 'Density' and 'Neutron'; the only tools that are directly responsive to oil, and therefore potentially good candidates. Each of these values m_t was subsequently used as observation in a HMC run, resulting in four sample sets. From these sets four conditional marginals P $(v_{\text{oil}}|m_t)$ were obtained. These are plotted together with the marginal of the prior P (v_{oil}) in figure 7. In this figure, we see clearly that the posterior with 'NMR-free' differs the most from the prior. This suggests that the 'NMR-free' measurement was (again in hind-sight) indeed the most informative measurement. Indeed, it can be shown that minus the expected

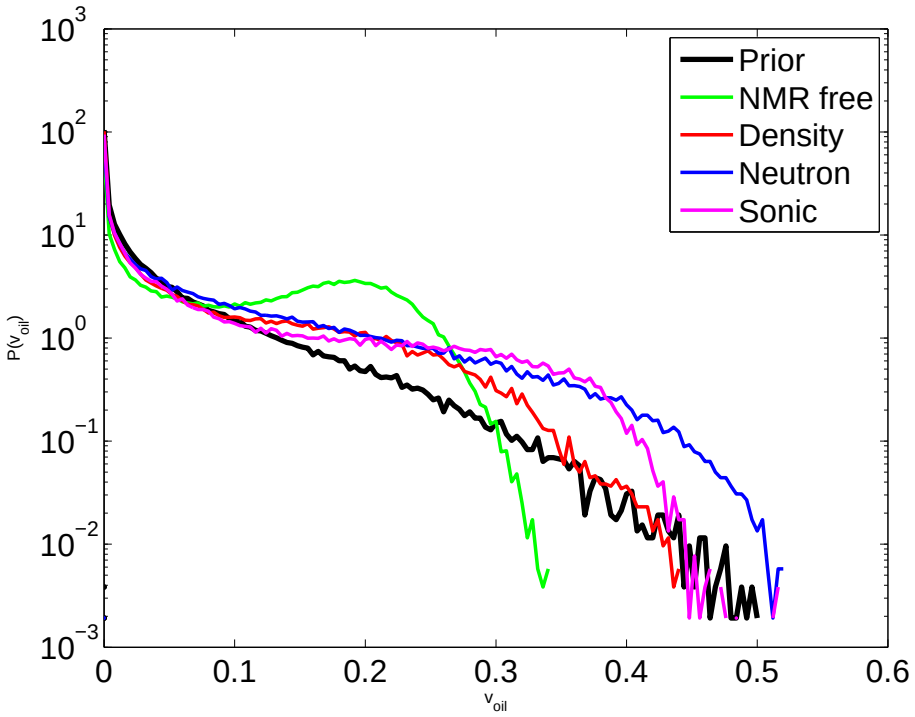

Figure 7: Histograms for v_{oil}; without observations (prior), or with one of NMR-free, Sonic, Density or Neutron.

entropy after the observation is up to a constant equivalent to the expected Kullback-Leibler divergence between the marginal after the observation and the marginal prior to the observation (see appendix B). In other words, informative tools lead to the posteriors that differ from the distribution prior to the measurement.

The outcome of our implementation of the decision support tool is a ranking of tools according to the expected entropies of their posterior distributions. In this way, the user can select a tool based on a trade-off between expected information and other factors, such as deployment costs and feasibility.

5 Discussion

This research has demonstrated a model and methodology for obtaining compositional estimates given some (or none) observations combined with expert knowledge, and presented a way of selecting the most informative tools (a method of quantifying information to be gained by performing measurements). The ability of the system to estimate compositions is tested using synthetic data. The estimates are within one error bar (uncertainty bound) from the actual value for unimodal distributions. For multimodal distributions the mean and error bar are poor statistics, and more information has to be obtained from histograms. Tests also confirmed that the method is consistent, different simula-

tions result in the approximately the same estimates with only minor differences. Bayesian models are modular and therefore easily extended. Obvious extensions are other tools (and/or other noise models), and an extension of the number of minerals and fluids. A more elaborate extensions is to include observations from multiple (adjacent) depths. The model then estimates a composition for each depth, and these compositions are stochastically linked via a lithology model. This changes the reservoir from the current homogeneous (aside from invasive effects in radial direction) description to an inhomogeneous one, with the advantage of including more information, and thereby generating more reliable results, but modeling lithology (transitions) might prove to be a challenge.

Another possible addition is the possibility to retrieve priors from a database. These might be sets of location specific shapes and means for the current Dirichlet prior, but might also be of completely different shape (even multimodal to account for mixtures). The current Dirichlet does not model mutually exclusive states (layered structures such as shales). A solution might be to model the prior as having multiple modes where each mode has a different set of fluids or minerals. For example the non-reservoir mode consisting of shale and coal (which behave in an exclusive fashion) and a reservoir mode consisting of Dirichlet distributed reservoir minerals and fluids. This could be sampled using reversible jump MCMC [Green, 1995]. It remains to be seen if this increased explanatory power of the model outweighs the modeling effort.

Regarding the representation of the results, an improvement might be to employ some form of automatic clustering algorithm to identify the modes of the distribution in order to report better descriptive statistics (as the variance and mean are poor characteristics for multimodal distributions).

Shell E&P are planning to use the developed methodology and software to increase uncertainty awareness among their petrophysicists. This uncertainty is inherent to log evaluation and must be dealt with in a consistent manner. Other uses are to assist in the evaluation of (ambiguous) reservoir logs so that evaluation uncertainties become visible and to enable petrophysicists to investigate options to reduce these uncertainties in order to reduce the number of faulty log evaluations.

Acknowledgements

The research reported here is part of the Interactive Collaborative Information Systems (ICIS) project, supported by the Dutch Ministry of Economic Affairs, grant BSIK03024.

A Porosity Prior

A usual requirement is that the porosity has a flat prior distribution. Porosity is a ratio of the form $\phi = \frac{X}{X+Y}$. Obviously $0 \leq X, Y \leq 1$ since both are summations over a subset of $\mathbb{S}^K$. If we choose parameters appropriately ($uw + \sqrt{uw^2} \ll 1$) then by approximation $X \sim \Gamma(u_1, w)$ and $Y \sim \Gamma(u_2, w)$. This ensures $\phi \sim \text{Beta}(u_1, u_2)$. The gamma distribution is given by

$$\Gamma(x|u, w) \equiv \frac{1}{Z(u, w)} x^{u-1} \exp\left(-\frac{x}{w}\right)$$

for $x, u, w \in \mathbb{R}^+$, and the beta distribution is given by

$$\text{Beta}\left(\phi|u_1, u_2\right) \equiv \frac{1}{Z\left(u_1, u_2\right)}\phi^{u_1-1}\left(1-\phi\right)^{u_2-1}$$

for $\phi \in [0,1] \quad u_1, u_2 \in \mathbb{R}^+$. If we require $\phi \sim \mathcal{U}\left(0,1\right)$ then it suffices to set $u_1 = u_2 = 1$, leading to $X, Y \sim \Gamma\left(\cdot|1, w\right)$. Both X and Y are of the form $\sum_j^J v_j$. If we assume $v_i \perp\!\!\!\perp v_j \iff i \neq j$, then $v_j \sim \Gamma\left(\frac{1}{J}, w\right)$ ensures the right distribution of X, Y. Parameter w is free to choose (within bounds described above) since it drops out in the normalization procedure; however using $w = 1/\#\text{groups}$ ensures that the means sum to unity. The prior over $\vec{v} \in \mathbb{S}^K$ yields

$$\text{P}\left(\vec{v}\right) \propto \prod_{n=1}^{\text{N}} v_n^{\frac{1-\text{N}}{\text{N}}} \prod_{j=1}^{\text{F}} v_{j+\text{N}}^{\frac{1-\text{F}}{\text{F}}} \prod_{i=1}^{\text{M}} v_{i+\text{N}+\text{F}}^{\frac{1-\text{M}}{\text{M}}} \tag{11}$$

where the products are over non-reservoir minerals (N), fluids (F) and matrix (M) minerals respectively. Although it is not necessary to model the non-reservoir minerals as Γ distributed (since these do not affect the value of the porosity), doing so ensures that the exp's drop out. The joint probability (11) is a Dirichlet distribution with parameters $\mu_j = 1/(\text{GE})$ and shape $\alpha = \text{G}$, where $\text{E} \in U = \{\text{N}, \text{F}, \text{M}\} \setminus \{0\}$ and G the number of mineral groups (e.g. the cardinality of U).

B KL Divergence

The Kullback-Leibler divergence (KL divergence) is a measure of difference between two probability distributions [MacKay, 2003].

$$\text{D}\left(\text{P}(x)||\text{Q}(x)\right) \equiv \int \text{P}(x) \log\left(\frac{\text{P}(x)}{\text{Q}(x)}\right) \text{d}x$$

For example, the KL divergence between conditional marginal $\text{P}\left(v_u|m_t\right)$ and prior $\text{P}\left(v_u\right)$ is given by

$$\text{D}_{ut}(m_t) = \text{D}\left(\text{P}\left(v_u|m_t\right)||\text{P}\left(v_u\right)\right)$$

The expected KL divergence is given by

$$\langle \text{D}_{ut} \rangle \equiv \int \text{P}\left(m_t\right) \text{D}_{ut}(m_t) \text{d}m_t$$

which is equal to $H_u - \langle H_{ut} \rangle$, where $H_u = -\int \text{P}\left(v_u\right) \log(\text{P}\left(v_u\right)) \text{d}v_u$ denotes the entropy of the prior, and $\langle H_{ut} \rangle$ is defined by equation 9. Because H_u is independent of t, the following identity must hold

$$\underset{t}{\text{argmin}} \langle H_{ut} \rangle = \underset{t}{\text{argmax}} \langle D_{ut} \rangle$$

References

S. Duane, A.D. Kennedy, B.J. Pendleton, and D. Roweth. Hybrid Monte Carlo. *Physics Letters B*, 195(2):216–222, September 1987.

P.J. Green. Reversible jump markov chain monte carlo computation and bayesian model determination. *Biometrika*, 82(4):711–732, December 1995.

D.J.C. MacKay. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, Cambridge, United Kingdom, 2003. http://www.inference.phy.cam.ac.uk/ mackay/itila/book.html.

N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equation of state calculation by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, June 1953.

R.M. Neil. Probabilistic inference using markov chain monte carlo methods. Technical Report CRG-TR-93-1, University of Toronto, September 1993.

J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers, Inc., San Francisco, CA, second edition, 1988.

M.R. Spalburg. Bayesian uncertainty reduction for log evaluation. *SPE International*, October 2004. SPE88685.