Lithology and Fluid Prediction from Prestack Seismic Data using a Bayesian model with Markov process prior

Hugo Hammer

Formerly: Norwegian Computing Center, Oslo, Norway,

presently: Oslo University college, Oslo, Norway.

E-mail: hugo.hammer@iu.hio.no

Odd Kolbjørnsen

Norwegian Computing Center, Oslo, Norway.

E-mail: Odd.Kolbjornsen@nr.no

Håkon Tjelmeland

Norwegian University of Science and Technology, Trondheim, Norway.

E-mail: haakont@stat.ntnu.no

Arild Buland

Statoil, Stavanger, Norway.

E-mail: abu@statoil.com

Abstract

We invert prestack seismic amplitude data to find rock properties of a vertical profile of the earth. In particular we focus on lithology, porosity and fluid. Our model includes vertical dependencies of the rock properties. This allows us to compute quantities valid for the full profile such as the probability that the vertical profile contains hydrocarbons and volume distributions of hydrocarbons. In a standard point wise approach, these quantities can not be assessed. We formulate the problem in a Bayesian framework, and model the vertical dependency using spatial statistics.

The relation between rock properties and elastic parameters is established through a stochastic rock model, and a convolutional model links the reflectivity to the seismic. A Markov chain Monte Carlo (MCMC) algorithm is used to generate multiple realizations that honours both the seismic data and the prior beliefs and respects the additional constraints imposed by the vertical dependencies. Convergence plots are used to provide quality check of the algorithm and to compare it with a similar method. The implementation has been tested on three different data sets offshore Norway, among these one profile has well control. For all test cases the MCMC algorithm provides reliable estimates with uncertainty quantification within three hours. The inversion result is consistent with the observed well data. In the case example we show that the seismic amplitudes make a significant impact on the inversion result even if the data have a moderate well tie, and that this is due to the vertical dependency imposed on the lithology fluid classes in our model.The vertical correlation in elastic parameters mainly influences the upside potential of the volume distribution.

The approach is best suited to evaluate a few selected vertical profiles since the MCMC algorithm is computer demanding.

Keywords: inversion, noise, numerical study, rock physics, seismics

1 Introduction

Seismic data is a key factor for identifying and risking prospects. Structural images from seismic data are the main source of information to map prospects and leads in the oil companies. The structure alone is however not sufficient for detection of hydrocarbon (HC) presence. Seismic amplitudes and amplitude variations with offset provide additional information about lithology-fluid (LF) type and reservoir quality. Standard AVO methods, which only consider a single contrast, neglect multilayered interference limiting their applicability.

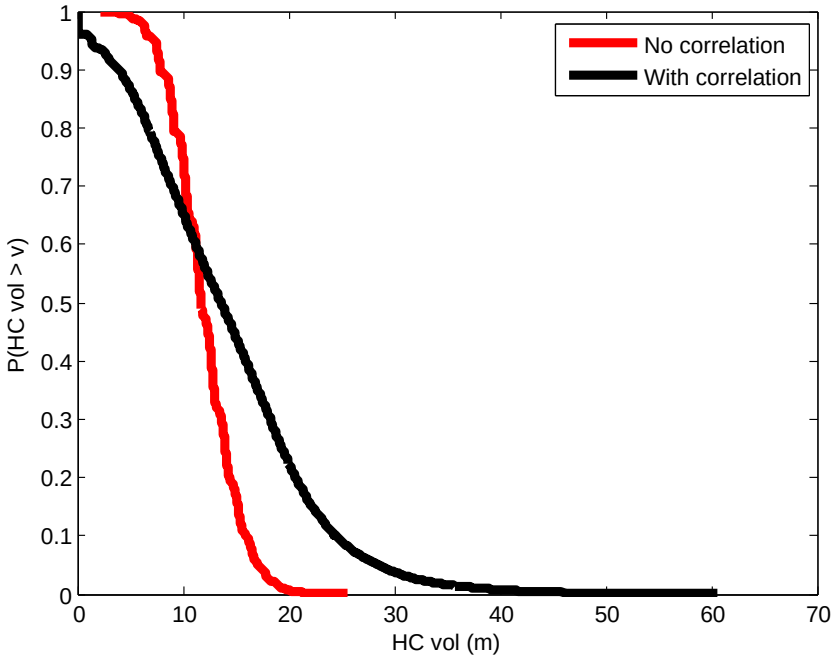

Figure 1: Cumulative volume curves $P(\text{HC volume} > v)$ for different v. Red curve: No spatial correlation is modeled, Black curve: Spatial correlation is modeled both in LF classes and elastic parameters.

Other methods which link seismic amplitude data to rock-physical properties are point-wise methods. These models compute the probabilities for the different LF classes in each cell, but not the spatial dependence between LF classes in neighboring cell. The point-wise methods are therefore not suitable to separate the probability of HC presence and HC volumes. We demonstrate this in Figure 1. Figure 1 contains volume curves for two models that differs in the spatial dependency in facies model. The volume curves displays the probability for observing a volume larger than the volume given at the first axis. For comparison both models relate to the same pointwise probabilities such that the probability of a LF-class is identical in all cells. The only difference between the models

is the spatial dependence. To highlight the importance of the difference, assume that a thickness of 20 m is required for a success case. In the independent model the probability of success is thus virtually zero, whereas the spatially dependent model has a chance of success of about 20% and an upside as large as 40 m thickness. This demonstrate the limitations of point-wise methods and the importance of modeling the spatial dependence.

In the exploration phase it is important to utilize all available knowledge to improve predictions and reduce risk prior to drilling. Knowledge is however limited and uncertain. To integrate different types of information, we adopt the Bayesian approach, see Duijndam (1988a,b); Ulrych et al. (2001); Scales and Tenorio (2001). The Bayesian solution is the posterior distribution of the LF classes and elastic properties gived the seismic data. For some inversion problems it is possible to characterize the posterior distribution analytically, but most often it is given by a Monte Carlo representation which is a large number of models drawn at random from the posterior distribution. The computational cost required to draw models from the posterior distribution will then be the limiting factor. In a practical setting, any Bayesian approach will therefore be a trade of between complexity in the prior and likelihood and our ability to produce realizations from the posterior distribution within a limited time span. The presented method focuses on the exploration setting, and is best suited for detection of new hydrocarbon reservoirs in the vicinity of existing wells.

A vertical profile of the earth is discretized into cells of constant properties and the LF class and rock-physical properties in these cells are the parameters of interest. In Buland and Omre (2003) vertical dependencies in elastic parameters are considered, in our work this is extended to include vertical dependencies for all rock properties elastic and categorical. In particular we use a prior model for the lithology and fluid class based on a Markov process, see Krumbein and

Dacey (1969). Previous work that has considered this type of LF-model, e.g. Larsen et al. (2006) and Hammer and Tjelmeland (2011), does not account for the correlation of elastic parameters within the same LF-class. The use of independence assumption is unrealistic since rock physical properties in the same sand interval will tend to be more similar than if different sand intervals are considered. The volume within a sand interval will always have larger uncertainty in a model where the porosity is correlated than in an independent model. The Markov process gives a larger set of plausible lithology combinations in comparison to the approach of Gunning and Glinsky (2004) which specify the relative position of lithologies prior to inversion.

A contribution of the current paper is that the more realistic Markov prior model is formulated such that the computational complexity drawing from the posterior distribution is not increased. The current work is along the lines of Kjønsberg et al. (2010). In particular in terms of the goal of the analysis. From a modeling perspective both approaches include vertical dependency in rock properties within one lithology, but differ with respect to the prior distribution for LF-classes. The Markov process used in the current approach gives a simple parametrization of the LF prior model and does not require that the target is positioned in the central part of the inversion region.

We consider pre-stack amplitude data and focus on LF classes and rock-physical properties. We use a stochastic rock physics model to create a link between the LF-class, porosity and elastic parameters, see Castagna et al. (1993); Mavko and Mukerji (1998); Avseth et al. (2005). To link the elastic parameters to the seismic, we use the likelihood model introduced in Buland and Omre (2003), but as explained above we have a more complex prior distribution for our parameters.

Related methods which consider the problem of reservoir characterization are

Haas and Dubrule (1994); Torres-Verdin et al. (1999); Contreras et al. (2005); Merletti and Torres-Verdin (2006). In these models the presence of lithologies are known and the spatial fractions of the facies are often fixed. The problem is to distribute a known set of lithologies in the reservoir region, in correspondence with the seismic data. The work of Bosch et al. (2007, 2009) apply Monte Carlo methods to do seismic amplitude inversion coupled with a petrophysical and geostatistical model. The problem faced in the exploration setting is that a major part of the uncertainty in the inversion is to determine which LF combinations are present in a prospect.

In the current paper we draw models from the posterior distribution by using a Markov chain Monte Carlo algorithm (Liu, 2001). The algorithm in this paper is a modification of the algorithm in Hammer and Tjelmeland (2011). Larsen et al. (2006) presents an alternative algorithm which is much faster, but only draw from an approximation to the posterior distribution. Buland et al. (2008) and Ulvmoen and Hammer (2009) show that the approximative algorithm loose a significant amount of the information content in the posterior distribution and in particular for properties involving the whole profile. This motivates for simulating from the posterior distribution without approximations. The papers Hammer and Tjelmeland (2011) and Ulvmoen and Hammer (2009) discuss statistical aspects in the simulation algorithm. We apply the methodology on three earth profiles offshore Norway and discuss the impact of the correlation structure have on the volume distribution of real data. We also compare the results obtained by Kjønsberg et al. (2010) both in terms of risk updates and performance of the Markov chain Monte Carlo algorithm.

2 Methodology

To make general assertions regarding the full vertical profile, we need to model the full set of dependencies between rock properties. A vertical profile of the earth is discretized into n cells where the rock properties of a cell represent the average property of this cell. We refer to the vertical positions along the profile with the index $i \in (1, \ldots, n)$. In our work we use a Markov process to model the vertical dependencies for lithology–fluid (LF) classes and elastic parameters.

2.1 Lithology-fluid distribution

The LF-class is divided into discrete categories e.g. shale, brine-sand and hydrocarbon-sand. Each cell in a vertical profile will belong to one LF-class, thus the cell thickness defines the scale of the LF-class. Characteristics of a vertical profile are the relative occurrence of each LF-class, thickness of intervals with identical LF class, and preferences in the ordering of LF-class. Examples of the latter property are that oil-sand should not be present directly below brine-sand, or that a shale with bad ceiling properties is more common above an interval of brine-sand than above an interval of hydrocarbon-sand.

In a Markov process it is possible to define parameters that preserve the probability of a LF-class, the mean thickness of a LF interval, and preferences in the relative ordering of LF-classes, see case study below for example. The model is defined such that the LF class in a position only depends on the LF classes in the neighboring cells. This reduces the number of parameters needed to specify a model. Since the Markov process is flexible with a low number of parameters it is ideal to use as a model for the vertical dependencies in the LF class. The parameters in the model can be estimated from analog well profiles, if this is available.

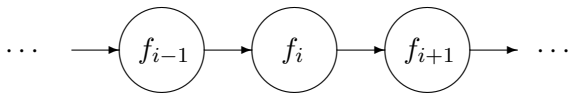

Figure 2: Directed acyclic graph (DAG) representation for the relation between the variables in $\mathbf{f}$.

The LF classes along the profile are denoted $\mathbf{f} = (f_1, \ldots, f_n)$, where the LF class in position i is denoted f_i. In each position there are in general L possible LF classes $1, \ldots, L$. The Markov process is defined sequentially, either top-down or bottom-up. We will use the top-down approach, but bottom-up will produce identical results. When defining the probability of a LF class in a new position given all LF classes above, the Markov property dictates that this should only depend on the LF class directly above. In Figure 2 this Markov property is represented by only having an arrow from the previous position and not all the previous positions. Based on the Markov property, the distribution for the Markov process can be written as

$$p(\mathbf{f}) = p(f_1) \prod_{i=2}^{n} p(f_i | f_{i-1}), \tag{1}$$

where $p(f_1)$ denotes the distribution for the LF class in the top position, and $p(f_i|f_{i-1})$ is the probability for the LF class in the current position when we know the LF class directly above, this is denoted transition probability.

Transition probabilities are specific to the resolution defined by the cell thickness, thus if the cell thickness is changed, a different transition probability should be used. It is however possible to preserve the relative occurrence of each LF-

class, the average thickness of intervals with identical LF class, and preferences in the ordering of LF-class. The average thickness $T(f)$ of a continuous interval with LF-class f is given by the formula,

$$T(f) = \frac{\Delta t}{1 - p(f|f)}, \tag{2}$$

where Δt is the cell thickness, and $p(f|f)$ is the probability of not changing the LF-class, i.e. a diagonal element on the transition matrix. The transition probabilities are are used to define the prior distribution of LF-classes in a vertical profile, when conditioning to seismic amplitude data the posterior distribution of lithology fluid classes will no longer retain the simple Markov structure of one step transitions. This is the main challenge when sampling the posterior distribution

Elastic parameters are denoted $\mathbf{m} = (\mathbf{m}_1, \ldots, \mathbf{m}_n)^T$. The link between a LF class and the elastic parameters in the same cell, are defined by a rock physical model. We select the model parameters to be the logarithm of P-wave velocity (V_P), S-wave velocity (V_S) and density (ρ), thus in each vertical position the elastic parameters are denoted by a three dimensional vector $\mathbf{m}_i = (m_{i,1}, m_{i,2}, m_{i,3})$. In a vertical position of the earth with LF class f_i, a stochastic rock physics model defines the distribution of the elastic parameters, $p(\mathbf{m}_i|f_i)$. For simplicity we will approximate this distribution with a Gaussian model which is defined through the mean, $\mathrm{E}\{\mathbf{m}_i|f_i\} = \boldsymbol{\mu}_{f_i}$ and covariance $\mathrm{Cov}\{m_{i,k}, m_{i,l}|f_i\} = \sigma_{kl,f_i}$. Note that each LF class will be approximated with a different distribution, thus the resulting distribution for the elastic parameters will be multi modal, having one mode for each LF class.

The vertical dependency in the elastic parameters is not defined through the rock physics model, but requires additional modeling. If we consider two different sand cells, then the elastic properties of these two cells tend to be more similar

if they come from the same sand interval than otherwise. This property can be modeled with the use of spatial correlation. The LF classes breaks the vertical profile into intervals where all consecutive cells have the same LF class. We assume that the elastic parameters in different intervals are independent, but model dependency within each of these intervals with a spatial correlation. The result of the vertical correlation is that we have small changes in the elastic parameters from the present to the next position if the LF class is the same for the two positions. In particular we select an exponential correlation function, i.e. if all positions between i and j have the LF class f, we have the correlation

$$\text{Cov}\{m_{i,k}, m_{j,l}|f\} = \sigma_{kl,f} \exp\left(-3\frac{|i-j|\Delta t}{R}\right), \tag{3}$$

where the first term is the covariance of the elastic parameters in same position, and the latter term is the component adjusting for the vertical relation, Δt is the size of the cells, i.e. the sampling interval, and R is commonly denoted correlation range. The factor R determines how fast the correlation decay. The factor 3 in the vertical term is set such that two cells being separated with distance R have negligible correlation. A small value for the correlation range results in independence between elastic parameters also within the same facies interval. A large value of the correlation range gives high dependency between elastic parameters.

Our particular choice of vertical dependency implies that when the LF class of the whole profile is fixed, the distribution of the elastic parameters, $p(\mathbf{m}|\mathbf{f})$, also have Markov properties. That is, the elastic parameters at the present position only depend on the elastic parameters at the position above, in addition to the LF classes in the profile. The distribution can then be written as

$$p(\mathbf{m}|\mathbf{f}) = p(\mathbf{m}_1|f_1)\prod_{i=2}^{n} p(\mathbf{m}_i|\mathbf{m}_{i-1}, f_{i-1}, f_i). \tag{4}$$

Figure 3 illustrates the relation between the variables in $\mathbf{f}$ and $\mathbf{m}$. Note in

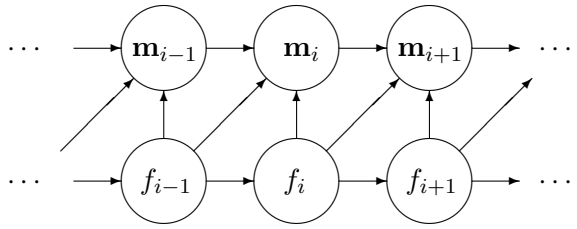

Figure 3: Directed acyclic graph (DAG) representation for the relation between the variables in $\mathbf{f}$ and $\mathbf{m}$.

particular that arrows into elastic parameters in one cell is only from the pair in the cell above and arrows out of this cell only goes to the pair below, thus giving the Markov structure.

2.2 Likelihood

The Likelihood model gives the statistical relation between the elastic parameters and the seismic amplitudes. This is modeled in two steps by first linking elastic parameters to the reflection coefficients and then linking these to seismic data. We denote the reflection coefficients and seismic data for all vertical positions along the profile and offset angle $\theta_j,\ j \in (1, \ldots, n_\theta)$ with the variables $\mathbf{c}_j$ and $\mathbf{d}_j$, respectively. Further let $\mathbf{c} = (\mathbf{c}_1, \ldots, \mathbf{c}_{n_\theta})$ and $\mathbf{d} = (\mathbf{d}_1, \ldots, \mathbf{d}_{n_\theta})$, representing all the reflection coefficients and seismic data related to the profile. The likelihood model we rely on, was developed in Buland and Omre (2003). We go through the details here to see how it appears in the current notation. We let the reflection coefficients, $\mathbf{c}_j$, be related to the elastic parameters $\mathbf{m}$ through a three term weak contrast approximation of the Zoeppritz equations (Aki and

Richards, 1980; Buland and Omre, 2003)

$$\mathbf{c}_j = \mathbf{A}_j \mathbf{D}\mathbf{m} + \boldsymbol{\varepsilon}_{1,j}, \tag{5}$$

where $\mathbf{A}_j$ represents the Aki and Richards equations and $\mathbf{D}$ gives the contrasts in the elastic parameters. The term $\boldsymbol{\varepsilon}_{1,j}$ is a multivariate Gaussian stochastic variable denoting noise and is assumed to be independent in each vertical position and for each angle. The standard deviation of the noise varies only with the angle. Thus given the elastic parameters and these assumptions of independence, we find that $p(\mathbf{c}|\mathbf{m}) = \prod_{j=1}^{n_\theta} p(\mathbf{c}_j|\mathbf{m})$, where $p(\mathbf{c}_j|\mathbf{m})$ is a Gaussian distribution given by expression (5).

The seismic data, $\mathbf{d}_j$, are related to the reflection coefficients through a convolutional model,

$$\mathbf{d}_j = \mathbf{W}_j \mathbf{c}_j + \boldsymbol{\varepsilon}_{2,j}, \tag{6}$$

where $\mathbf{W}_j$ is a wavelet matrix and $\boldsymbol{\varepsilon}_{2,j}$ is white noise also with different variances for the different angles. We can then write $p(\mathbf{d}|\mathbf{c}) = \prod_{j=1}^{n_\theta} p(\mathbf{d}_j|\mathbf{c}_j)$.

Substituting equation 5 into equation 6 gives

$$\begin{aligned} \mathbf{d}_j &= \mathbf{W}_j \mathbf{A}_j \mathbf{D}\mathbf{m} + \mathbf{W}_j \boldsymbol{\varepsilon}_{1,j} + \boldsymbol{\varepsilon}_{2,j} \\ &= \mathbf{W}_j \mathbf{A}_j \mathbf{D}\mathbf{m} + \boldsymbol{\varepsilon}_j. \end{aligned} \tag{7}$$

We have two sources of noise, wavelet coloured and white noise. The term $\mathbf{W}_j \boldsymbol{\varepsilon}_{1,j}$ is the wavelet coloured noise, while $\boldsymbol{\varepsilon}_{2,j}$ is white. The variable $\boldsymbol{\varepsilon}_j$ represents the total noise level.

This shows that the error term in equation 5 is not only due to errors in the linearization of the Zopperitz equation, but also due to the colored components generated by imperfections in the seismic processing. This means that $\mathbf{c}$ should not be interpreted directly as reflection coefficients, but reflection coefficients with errors. Since we really are interested in facies and elastic parameters this

interpretation is not important for our results. The quantity $\mathbf{c}$ is however essential in the algorithm that generate realizations from the posterior.

2.3 Posterior distribution

Of interest now, is to evaluate the posterior distribution

$$p(\mathbf{f}, \mathbf{m}, \mathbf{c}|\mathbf{d}) \propto p(\mathbf{f})p(\mathbf{m}|\mathbf{f})p(\mathbf{c}|\mathbf{m})p(\mathbf{d}|\mathbf{c}), \tag{8}$$

and in particular the distribution of LF classes $\mathbf{f}$ and elastic parameters $\mathbf{m}$. We evaluate the posterior distribution using a Monte Carlo algorithm. Due to the specific model formulation, it is possible to adjust the algorithm presented in Hammer and Tjelmeland (2011) to obtain realizations from the posterior distribution in expression (8). Below and in appendix A we describe algorithm we use. This accounts for the dependency in the elastic parameters and also provides simplifications in comparison to Hammer and Tjelmeland (2011).

2.4 Simulation algorithm

We use a Markov chain Monte Carlo (MCMC) algorithm (Liu, 2001) to generate a large number of realizations of $\mathbf{f}$, $\mathbf{m}$ and $\mathbf{c}$ from the posterior distribution (8). The realizations are then a representation of the posterior distribution. The updates in each iteration use values from the previous iteration. Therefore we need some initial values to run the first iteration of the algorithm. We generate the initial values by a random draw of LF classes, elastic parameters and reflection coefficients from the prior model, $p(\mathbf{f}, \mathbf{m}, \mathbf{c}) = p(\mathbf{f})p(\mathbf{m}|\mathbf{f})p(\mathbf{c}|\mathbf{m})$. This is a natural choice since this is the information we have about the variables before the seismic data is considered. The algorithm then iteratively generates new realizations of $\mathbf{f}$, $\mathbf{m}$ and $\mathbf{c}$ using two steps in each iteration. In the first step we simultaneously update $\mathbf{m}$ and $\mathbf{c}$ by drawing from $p(\mathbf{m}, \mathbf{c}|\mathbf{f}, \mathbf{d})$, representing

the distribution for elastic parameters and reflection coefficients given that the seismic data and LF classes are known. It is simple to generate realizations from this distribution, since it is multivariate Gaussian, i.e. the model in Buland and Omre (2003) apply. This update is what is called a Gibbs step in the statistical literature (Liu, 2001) and the updates of $\mathbf{m}$ and $\mathbf{c}$ will be realizations from the posterior distribution for the given LF classes.

The second step in each iteration is more complex. We start by proposing a potential new set of LF classes and elastic parameters $(\mathbf{f}^*, \mathbf{m}^*)$ by drawing randomly from a proposal distribution, $(\mathbf{f}^*, \mathbf{m}^*) \sim q(\mathbf{f}, \mathbf{m}|\mathbf{c})$ where the proposed values condition on the value of c generated in the previous Gibbs step of the algorithm. The realization of LF classes and elastic parameters $(\mathbf{f}^*, \mathbf{m}^*)$ is then accepted with a certain probability

$$\alpha = \min\left\{1, \frac{p(\mathbf{f}^*, \mathbf{m}^*, \mathbf{c}|\mathbf{d})q(\mathbf{f}, \mathbf{m}|\mathbf{c})}{p(\mathbf{f}, \mathbf{m}, \mathbf{c}|\mathbf{d})q(\mathbf{f}^*, \mathbf{m}^*|\mathbf{c})}\right\}, \tag{9}$$

where $\mathbf{f}, \mathbf{m}$, and $\mathbf{c}$ are the values from the previous iteration. If the proposal is rejected, then the LF classes and elastic parameters are not altered by this step. This type of algorithm is standard in the statistical literature and falls in the class of Metropolis–Hastings algorithms. The key part in the construction of an efficient algorithm, is the proposal distribution $q(\mathbf{f}, \mathbf{m}|\mathbf{c})$. In our approach we use a mixture of Gaussian distributions as proposal mechanism, see Appendix A for details.

We do not need the updated value of $\mathbf{m}$ from the first step of the algorithm in the second step. The only reason we also update $\mathbf{m}$ in the first step is for computational reasons. Simultaneously updating $\mathbf{m}$ and $\mathbf{c}$ is more effective than doing a Gibbs step where only $\mathbf{c}$ is updated, i.e. updating $\mathbf{c}$ from the distribution $p(\mathbf{c}|\mathbf{f}, \mathbf{m}, \mathbf{d})$.

Generating proposals from $q(\mathbf{f}, \mathbf{m}|\mathbf{c})$ is memory demanding on a computer. The

result is that we are not always able to generate good proposals for the LF classes and elastic parameters along the whole profile. This is resolved by generating proposals only for an interval of the profile in each iteration. In each iteration we have the interval that we update. This is commonly known as block updates in the statistical literature.

2.5 Computing aggregated properties from the posterior realization

Having obtained N realizations from the posterior distribution, we want to compute statistics for the properties of our interest. Below we present how to compute the probabilities for different LF classes in each position, the probability that the profile contains HC and the volume distributions of HC if HC is present. Other quantities can also be evaluated.

2.5.1 LF probabilities

The standard result in point wise methods is the probability of the LF class in each of the cells, i.e. $P(f_i = l|\mathbf{d})$ for positions $i = 1, \ldots, n$ and $l = 1, \ldots, L$. We estimate this by counting the amount of realizations where $f_i = l$ and divide by the total amount of realizations N,

$$\widehat{P}(f_i = l|\mathbf{d}) = \frac{1}{N}\sum_{k=1}^{N} I(f_{i,k} = l), \tag{10}$$

where $\widehat{P}(f_i = l|\mathbf{d})$ is the estimate of the LF probability, $f_{i,k}$ is the LF class at position i for realization k, and $I(f_{i,k} = l)$ the indicator function returning 1 if $f_{i,k} = l$ and 0 else.

2.5.2 Probability of discovery

Further we can also compute the posterior probabilities of discovering gas, oil or both gas and oil in the profile. This can not be computed from point wise methods. In practical terms an accumulation of hydrocarbons is not regarded a discovery if the thickness is too small. The definition of discovery will therefore wary from case to case. In our presentation we define a discovery as a presence of commercial hydrocarbon in a gross rock thickness of more than 10 meter.

To estimate the posterior probability of finding only gas, we count the portion of the realizations with hydrocarbon discovery and where the volume of gas is larger then zero and the volume of oil is equal to zero. Similarly we can compute the probabilities of discovering only oil and discovering both oil and gas in the profile.

2.5.3 Volume distributions of gas and oil

The previous Section divides the realizations from the posterior into four groups, those with no discovery, pure gas cases, pure oil cases and mixed cases. For all three success groups we want the distribution of the hydrocarbon volume, as this is essential information in the prospect evaluation.

The porosity ϕ of sand stone is important in volume computations. In a model with fixed fluid density this is directly linked to the density of the saturated rock through the relation,

$$\rho = \phi\rho_F + (1-\phi)\rho_M. \tag{11}$$

where ρ, ρ_F and ρ_M are the bulk density of the saturated rock, the fluid and the host-rock, respectively. The density of the different hydrocarbon phases, brine and the host-rock are fixed in the rock physical model we use. In addition we

also fix the saturation in each fluid class, such that the fluid density is fixed by the fluid class. Variability in saturation is obtained using multiple fluid classes. Similarly the variability in the density of the host-rock can be obtained using multiple lithology classes. When both the density of the saturated rock and the fluid class is known, expression (11) is inverted to compute the porosity:

$$\phi = \frac{\rho_M - \rho}{\rho_M - \rho_F}. \tag{12}$$

The LF-class and the elastic parameters are given in each realization, thus the density of the saturated stone is given along with the fluid class. This is used to compute the porosity of hydrocarbon filled sandstone using relation 12. The saturation is also defined by the fluid class. The cell thickness in meters is given as the cell thickness in terms the one-way travel time multiplied with the interval velocity in the cell.

The volume of gas in a realization from the pure gas group is the sum of the hydrocarbon volume in all cells containing gas. The volume in one cell is computed as the product of cell volume, porosity and saturation. The total gas volume V_{gas} in a realization from the pure gas group is hence computed by the formula,

$$V_{\text{gas}} = \sum_{i \in \text{gas cells}} S_i \phi_i V_{P,i} \Delta t/2, \tag{13}$$

were S_i is the hydrocarbon saturation, ϕ_i the porosity, $V_{P,i}$ is the P-wave velocity, and $\Delta t/2 = 2$ ms is the one-way travel time of the cell thickness. The index i refers to cell i along the profile. The volume of oil is found similarly. Since we only consider 1D profiles in this paper it is really a volume per unit area which is computed.

This computation is done for all realizations in each of the three groups only gas, only oil and both gas and oil, and are used to compute the volume distribution of each group. In the single phase cases we find the volume densities of the

given phase, whereas in the two phase case we find the joint distribution.

2.6 Evaluation of the simulation algorithm

Since the Markov chain Monte Carlo algorithm is an iterative algorithm that produces realizations from the correct distribution in the long run, it is important to control how it behaves with a finite number of realizations. Firstly it is important to monitor that the realizations are from the correct distribution, i.e. that the algorithm does not get trapped in a local mode. Further we want consecutive realizations in our updating scheme to be as dissimilar as possible. In statistical terms the first property is denoted convergence, the latter is denoted mixing.

2.6.1 Convergence

Evaluation of convergence for MCMC algoritms is a challeging topic. In this paper we apply a simple, effective and much used method based on running the MCMC algorithm from several different initial values and see that all the chains give the same results. Note that convergence does not mean convergence to a unique optimal value, but that the algorithm produces realizations from the posterior distribution. For more on MCMC convergence, see for example Liu (2001).

2.6.2 Mixing properties

In the experiment we focus on evaluating how fast we can estimate the LF probabilities $P(f_i = l|\mathbf{d})$ using the estimate in equation 10 in the previous section. To investigate this, we run K independent Markov chains with a large amount of realizations N (after burn in) in each. We are then able to get a very

precise estimate using all the realizations from all the K chains. We simply denote the estimate with the true probability $P(f_i = l|\mathbf{d})$. Next we estimate the LF probabilities using only the first part of one of the chains. Using the first ν realization, denote the resulting estimate $\widehat{P}_\nu(f_i = l|\mathbf{d})$. We are now interested in evaluating how fast the estimate $\widehat{P}_\nu(f_i = l|\mathbf{d})$ converge to the true probability $P(f_i = l|\mathbf{d})$. We therefore take the difference in absolute value and average over all the possible LF classes $(1, \ldots, L)$

$$\delta_{i,\nu} = \frac{1}{L} \sum_{l=1}^{L} \left| \widehat{P}_\nu(f_i = l|\mathbf{d}) - P(f_i = l|\mathbf{d}) \right|. \tag{14}$$

The decay of the discrepancy with an increasing number of realizations ν illustrates the efficiency of the algorithm.

3 Real data example

In this case study we invert seismic data from three vertical profiles offshore Norway. We denote the profiles A, B and C. These are the same profiles and data as considered in Kjønsberg et al. (2010). From the inversion we find facies probabilities and pore volume distributions. In profile A we have well log observations in addition to the seismic data, shown in Figure 4. From left to right we have well log elastic parameters, LF-class, porosity, seismic data and synthetic seismic. The LF-class is determined from the well logs for shale content and brine saturation. The synthetic seismic is calculated using the convolutional model. Comparing the seismic data with the synthetic seismic data, we see that there is a moderate fit. In Figure 5 we see the seismic data for profile B and C. Based on structural information, we expect to find hydrocarbons (HC) in profile B, while C is expected to be outside the reservoir. The length of the profiles are 0.4 seconds and the sampling interval is 4 ms which also is the resolution used for the definition of LF-class.

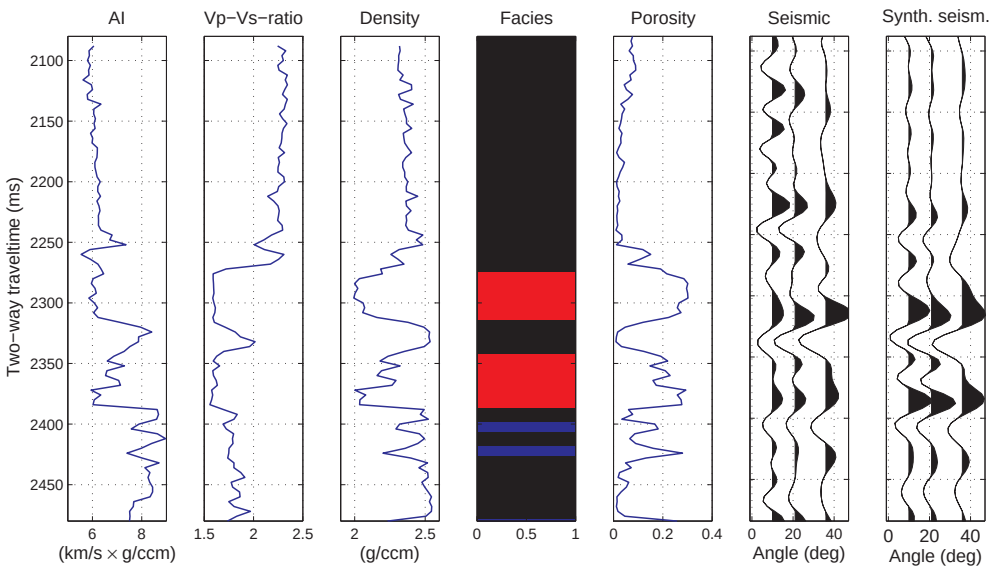

Figure 4: Location A. From left to right in the figure: Well log elastic parameters, LF-class, porosity, seismic data and synthetic seismic. Colour code for facies categories: shale – black, brine sand – blue, gas sand – red.

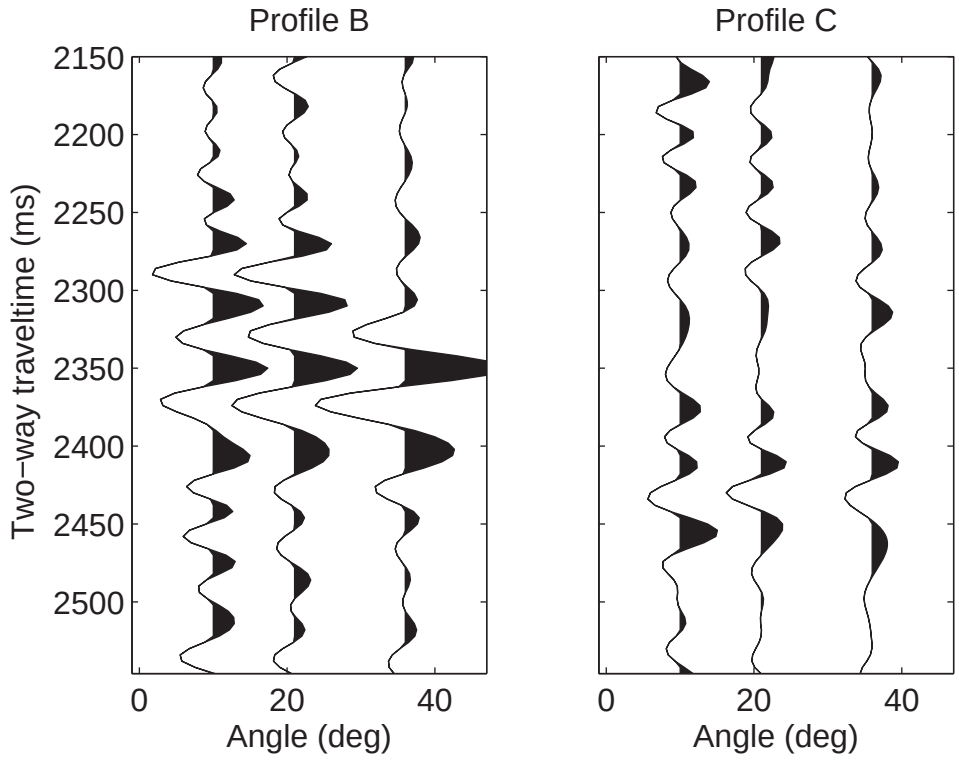

Figure 5: Seismic data for the profiles B and C.

Table 1: Transition probabilities. Each column show how the probability for the LF class in the current position varies with the LF class in the cell above.

	HS	Shale	Brine	Oil	Gas	Fizz
HS	0.510	0.415	0.018	0.019	0.019	0.019
Shale	0.028	0.928	0.011	0.011	0.011	0.011
Brine	0.025	0.236	0.739	0	0	0
Oil	0.013	0.099	0.165	0.724	0	0
Gas	0.014	0.113	0.116	0.033	0.721	0
Fizz	0.013	0.106	0.1448	0	0	0.737

3.1 Parametrization of the model

The well log in location A is used to guide the prior model formulation. The locations B and C are from the vicinity of A, thus the same prior parameters are used for all three locations. The lithologies considered are hot shale, shale and sand. The sand contains one of four different fluids being brine, oil, high saturated gas, and low-saturated gas, the latter is denoted fizz. Hot shale is shale that contains organic material and has lower acoustic impedance than standard shale. The model gives six distinct LF classes.

It is possible to estimate transition probabilities from near by wells or analogies. In the present case we do not have a well with all LF-classes present, and thus this is not possible. In our approach we use the transition probabilities as a mean to impose additional information in the model, and thus there will be a certain degree of subjectivity in the choice transition probabilities, corresponding to the subjectivity in selecting LF-classes an prior probabilities for these. The transition probabilities, $p(f_i|f_{i-1})$, in the Markov process prior, see expression (1), is given in Table 1. Here row one to six denote the probabilities going from hot shale, shale and brine-, oil-, gas- and fizz-saturated sandstone,

Table 2: Properties of the prior model. First row: Probabilities $p(f_i)$. Second row: Average thickness of the intervals in ms.

	Hot shale	Shale	Brine	Oil	Gas	Fizz
P	0.05	0.75	0.10	0.033	0.033	0.033
T	8.1	55.6	15.3	14.5	14.3	15.2

respectively. The transition probabilities contain information about the vertical properties of the lithology. The Table also gives the probability of transitions between different lithologies, e.g. it is about five times more likely that there is a shale below a hot shale than that there is either of the sands. Some of the transition probabilities are zero, this indicates that the transition is illegal. The Table shows that it is impossible to go downward from brine to oil. This is consistent with the fluid ordering in equilibrium. Table 2 summarize the effect the transition probabilities have on the prior probabilities for LF classes in a vertical position $p(f_i)$ and the average length of continuous LF-class interval, when the lithology definition is given at 4 ms. We can use well log information to find information about average thickness of intervals and what is common transitions. This information in addition to stating what should be illegal transition, is to construct prior transition probabilities.

The distribution of the elastic parameters for each LF class $p(\mathbf{m}_i|f_i)$ is illustrated in Figures 6 and 7. Figure 4 indicates that there is a trend in the elastic parameters. This trend is not included in the current case example. It is possible to include a trend in our model as well, but in the exploration setting this is often hard to define the trend with sufficient accuracy. The important feature in the inversion, is the relative positions of the distributions of elastic parameters for the LF classes. This is often a more stable than the absolute positions of the distributions. In Figures 6 and 7 we see that the distribution for some of the

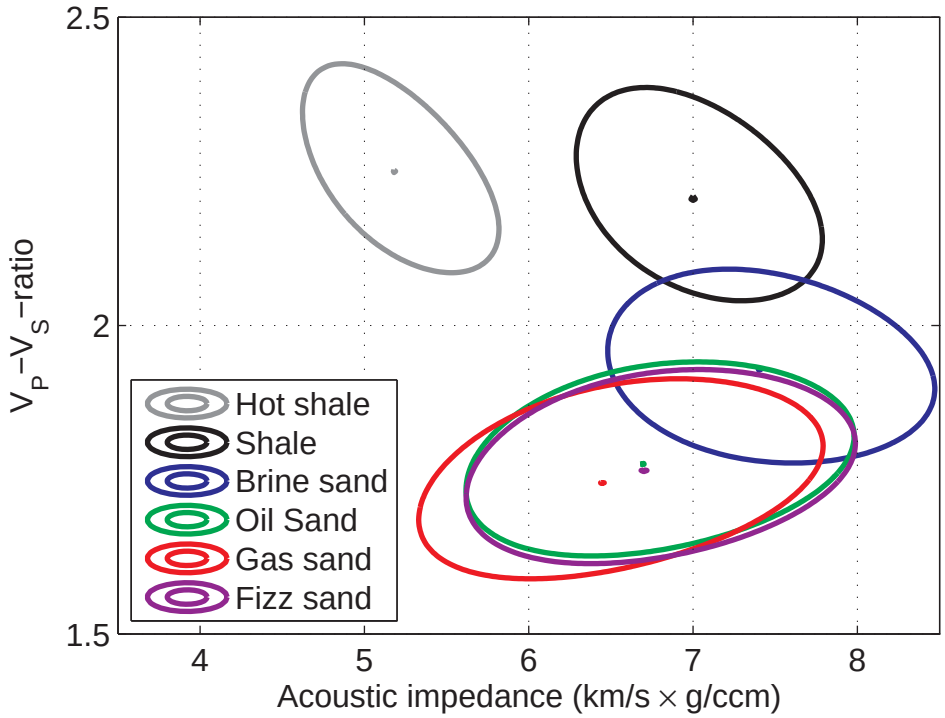

Figure 6: Plots of the prior distribution for AI and V_P/V_S. The colours gray, black, blue, green, red and purple refer to hot shale, shale and brine-, oil-, gas- and fizz-saturated sandstone, respectively. The center is the most likely value, the outer limit shows the region containing 90% of data.

LF-classes have a substantial overlap. For example is it almost impossible to distinguish between oil-sand and fizz-sand. In the modeling approach it is however still important to include both since fizz-sand is a failure case and oil-sand is a success case. Even though data will not alter the relative occurrence of the two LF-classes, the possibility of having fizz-sand will influence the probability of success as well as the volume distribution. The last parameter is the correlation range of the elastic parameters. In our initial runs we use a correlation range of 50 ms for the elastic parameters.

For all three profiles we have seismic data for $n_\theta = 3$ angles, $10°, 21°$ and $36°$. The wavelet and the noise level are estimated from the well log at profile A. The

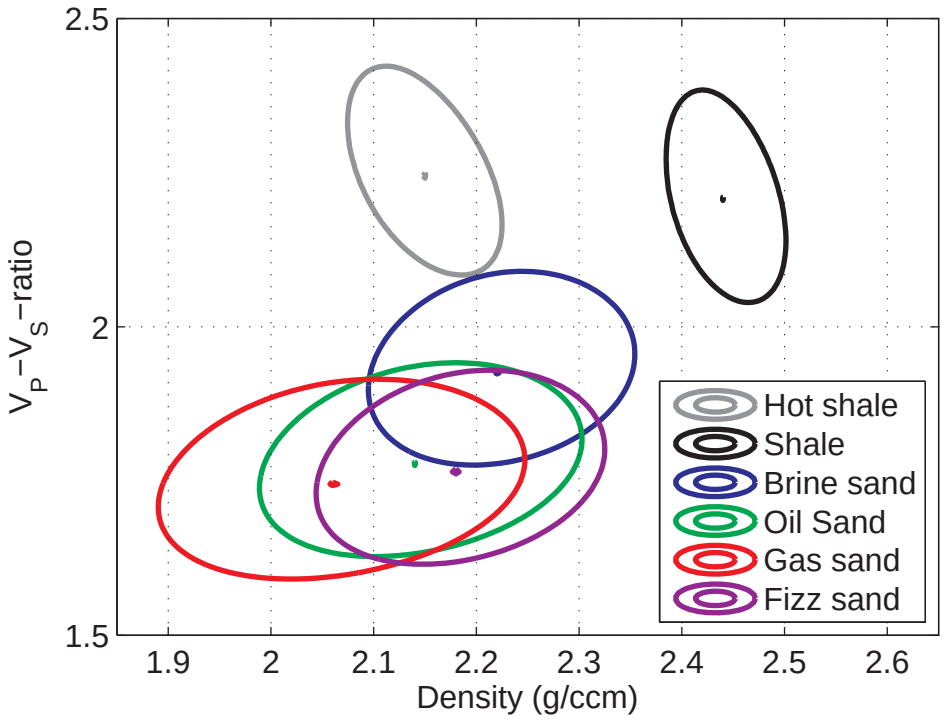

Figure 7: Plots of the prior distribution for density and V_P/V_S. The colours gray, black, blue, green, red and purple refer to hot shale, shale and brine-, oil-, gas- and fizz-saturated sandstone, respectively. The cent-re is the most likely value, the outer limit shows the region containing 90% of data.

resulting signal-to-noise levels are 1.5, 1.7 and 1.6 for the three angles $10°, 21°$ and $36°$, respectively. The noise levels for $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ are defined such that 10% of the energy in the noise is white and 90% of the energy is wavelet colored.

As explained in the last paragraph in Section 2.4 we are not able to generate good proposals for the whole profile in each iteration. In the simulation, we therefore update an interval of 41 nodes in each iteration.

3.2 Results

In Figure 8 we show 10 randomly picked LF-realizations from the posterior distribution together with the point wise LF prediction in location A. There is clearly large uncertainty in LF-classes, at the same time all realizations have certain similarities. In particular two layers of non standard shale in the position of the two reservoir layers are clearly visible. These two layers are clearly seen in the point wise probabilities as well.

Figure 9 shows the realizations of the elastic parameters which correspond to the five first LF-class realizations in Figure 8. We see that the acoustic impedance has the least uncertainty, and that the V_p-V_s-ratio and density has less uncertainty in the interval 2300 ms-2350 ms.

Figure 10, shows the uncertainty in the contrasts of the elastic parameters compared with the well log. All major contrasts are obtainable within the model. There are also possibilities for additional contrasts in the data, these correspond to layers that are not in the well, but are plausible based on the seismic amplitude data. The comparison of contrasts is done to avoid the effect of the low frequent trend. There virtually no influence of the vertical trend in the plot of the well logs, thus the relative positions of lithology classes is reasonably stationary.

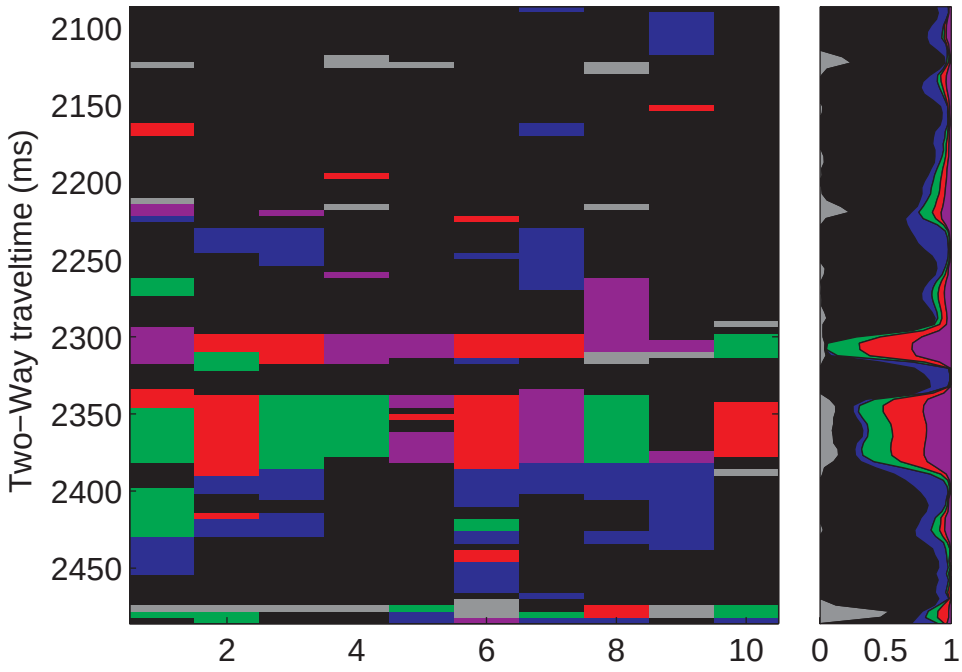

Figure 8: LF-prediction in location A. To the left ten different realizations of LF-class from the posterior distribution. To the right the point wise probability of LF-class. The colours gray, black, blue, green, red and purple refer to hot shale, shale and brine-, oil-, gas- and fizz-saturated sandstone, respectively.

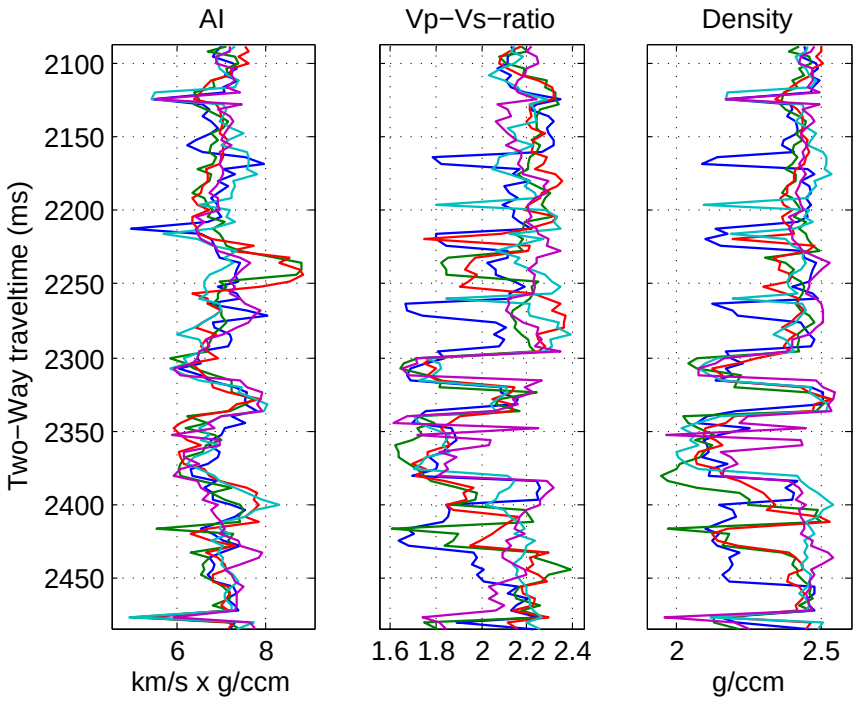

Figure 9: Realizations of elastic parameters. The figure show the realizations of the elastic parameters corresponding to the five first facies realizations in Figure 8.

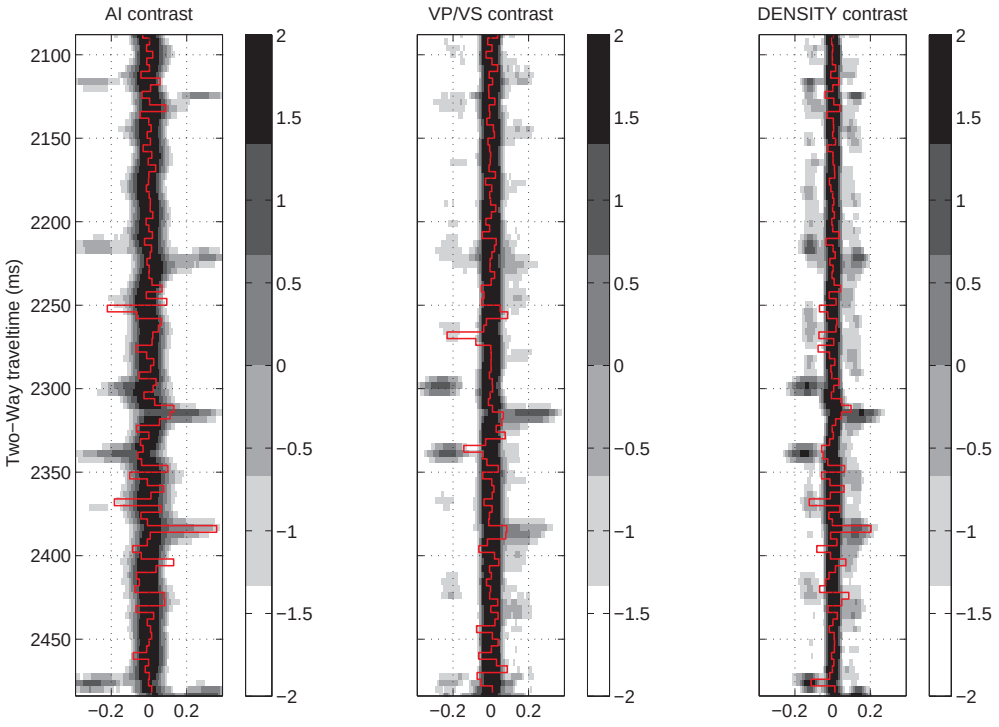

Figure 10: Contrasts in elastic parameters. From left to right is the contrasts in acoustic impedance, Vp-Vs-ratio and density. The red curve is the well log, the colours are the logarithm of the histogram in each position.

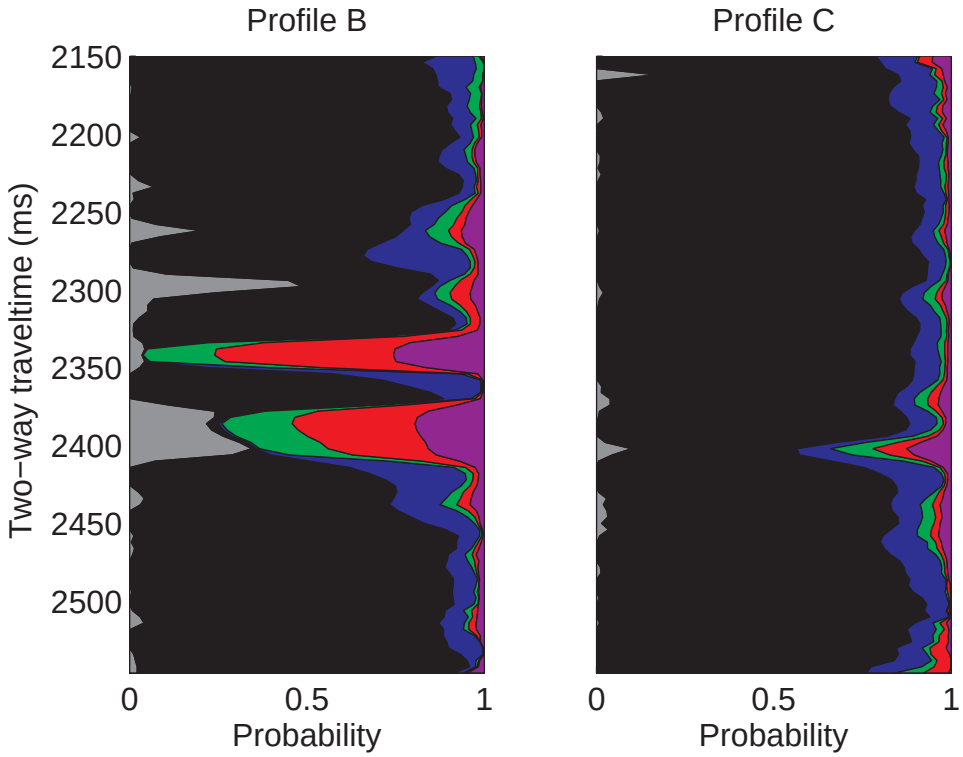

Figure 11: Results for profiles B and C. Marginal probabilities from the simulations with range R=50 ms.

In Figure 11 we see marginal probabilities for profile B and C. In profile B we have high probability for hydrocarbon sand in two separate layers and a third layer above these is most likely hot-shale. The top reservoir layer in profile B is more likely to be gas than oil or fizz, but it is generally hard to distinguish between the three hydrocarbon cases, as we expected from the rock physics models. In profile C the amplitudes give less room for hydrocarbon presence than in profile B.

The target for the inversion is however not the point wise LF-class probabilities, but the success probability when drilling a well in the location, and the volume distribution for the success cases. We summarize the probabilities for discovering oil and/or gas in all three profiles in Table 3. The prior probabilities of discovery

Table 3: Prior and posterior probabilities for discovering oil and/or gas, range 50 m. A success is defined as the discovery of gross rock thickness containing commercial hydrocarbon sand larger than 10 meters.

	no HC	oil	gas	gas and oil
Prior	0.22	0.19	0.15	0.43
Profile A	0.07	0.16	0.18	0.59
Profile B	0.07	0.14	0.27	0.52
Profile C	0.40	0.21	0.15	0.24

are included in the table for comparison. We find that the seismic amplitudes provide evidence for hydrocarbon presence in profile A and B, and against it in profile C.

The volume distribution is of interest in the success scenarios. Figures 12, 13 and 14, contain the volume distributions for the three profiles. Again we see that the volume distribution in profile C is substantially less than in profiles A and B. The volume distributions in profile A has two modes in the single fluid case, and three modes in the mixed fluid case.

In the well which coincides with profile A it was found gas in a column of 28.7 m. This value is about the same as the 95% quantile of the posterior volume distribution. This means that in the posterior distribution it is about a 5% chance to observe a volume larger then the volume extracted from the well.

To investigate the effect of the vertical correlation of elastic parameters within the same continuous LF-class, we also tested a case with correlation range of 10 m. The discovery probabilities for these runs are listed in Table 4. There are no strong systematic effects. The largest observed difference is the reduction in the discovery probability in profile C.

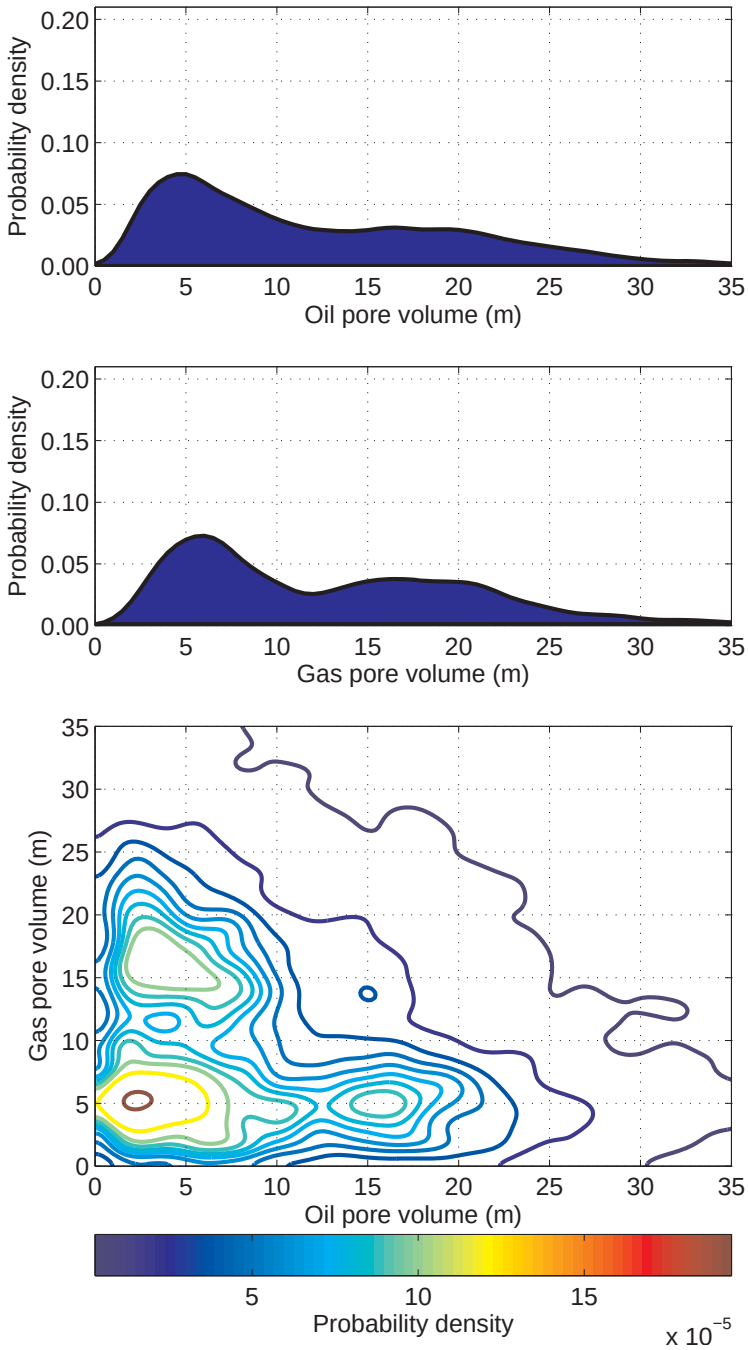

Figure 12: HC pore volume distributions for profile A. Top row: Volumes of oil given that the profile contains only oil; Middle row: Volumes of gas given that the profile contains only gas. Bottom row: Volumes of oil and gas given that the profile contains both oil and gas.

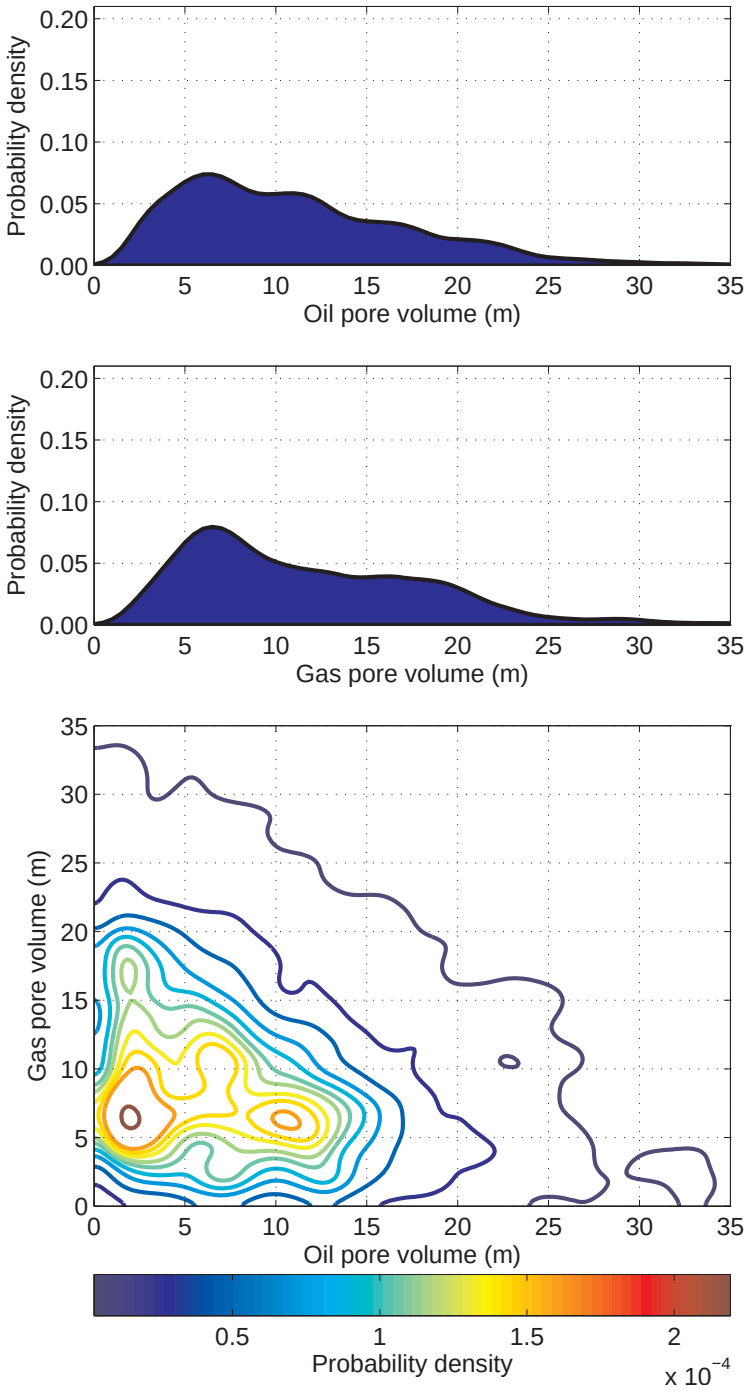

Figure 13: HC pore volume distributions for profile B. Top row: Volumes of oil given that the profile contains only oil; Middle row: Volumes of gas given that the profile contains only gas. Bottom row: Volumes of oil and gas given that the profile contains both oil and gas.

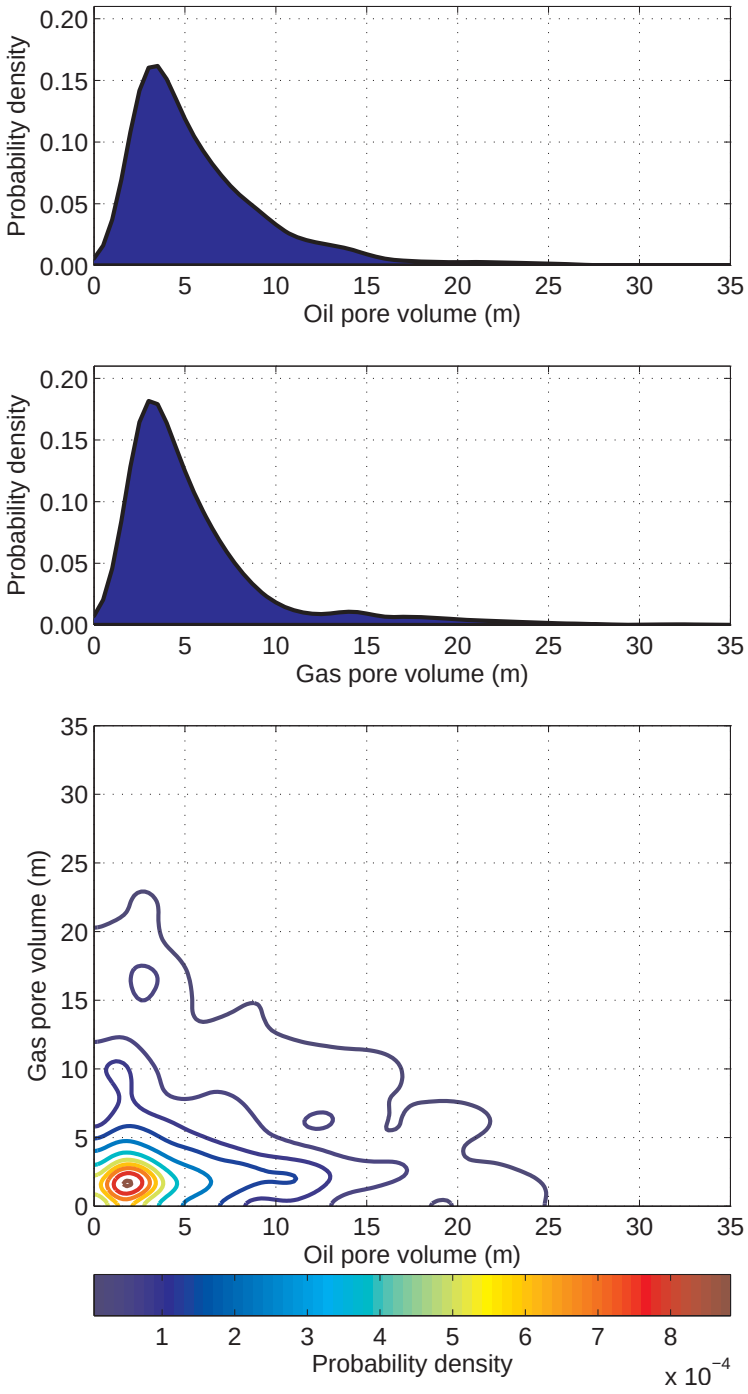

Figure 14: HC pore volume distributions for profile C. Top row: Volumes of oil given that the profile contains only oil; Middle row: Volumes of gas given that the profile contains only gas. Bottom row: Volumes of oil and gas given that the profile contains both oil and gas.

Table 4: Posterior probabilities for discovering oil and/or gas, range 10 m. A success is defined as the discovery of gross rock thickness containing commercial hydrocarbon sand larger than 10 meters.

	no HC	oil	gas	gas and oil
Profile A	0.05	0.14	0.21	0.60
Profile B	0.07	0.11	0.31	0.51
Profile C	0.48	0.23	0.12	0.17

In Tables 5 and 6 we evaluate how the modeling of spatial dependence in the prior model affect the volume distribution. We consider three different prior model choices. We denote the first choice 'No correlation'. In this case we have no spatial correlation in the prior distribution. For the LF classes, this means that the LF class in one position is independent of the LF classes in all other position. This is in contrast to the prior based on the transition matrix in Table 1, where it is essential that the LF class in a position is dependent on the LF class in the previous position. As prior probabilities in a position, $p(f_i)$, for the independence prior, we use the same probabilities that we have for the transition matrix in Table 1 (given in the first row in Table 2). For the independence prior, we further set the spatial correlation range in the elastic parameters equal to 0 m. We denote the second prior model we consider 'Range = 10 m'. In this model we use the transition matrix in Table 1 and use a spatial correlation range in the elastic parameters equal to 10 m. The third prior model choice, denoted 'Range = 50 m', is equal to the previous case except that the spatial correlation range in the elastic parameters is equal to 50 m. In Table 5 we present volume distributions for these three prior model choices. We list the mean volume and the 10% and 90% quantiles for each case. In the mixed case we add the volume of oil and gas to get one summarizing volume. The 10% and 90% quantiles are defined as values for the volume such that it is 10% and 90%

probabilities, respectively, to observe a volume less then the given quantile. In the first four rows we have results for the model with no prior correlation, in the next four rows for the prior model with range 10 m and the last four rows for range 50 meters. We see that for the prior models with vertical correlation, the hydrocarbon volumes is much larger for profile A and B then for the prior model with no spatial correlation. For the prior model with no correlation the volumes are almost the same for Profile A, B and C. This means that the modeling of spatial dependence in the prior model is essential to separate the success cases Profile A and B from the failure case C. Further we see that upper limit for the volume generally are larger for the model with correlation range 50 m compared to 10 m. On average the upper limit of the hydrocarbon volume is increased by 1.6 m. This corresponds to an average increase of 12% in the upper limit. Thus there is a larger upside potential for models with long spatial correlations. This is of particular importance in marginal developments.

In Table 6 we have computed risked volume of HC for the three prior model choices. This Table also contains the results obtained from Kjønsberg et al. (2010). Similar to Table 5, we see from the three first columns that the modeling of dependence in the prior model is essential to separate Profile A and B from C. The major difference in results in this paper and Kjønsberg et al. (2010) is in profile B, where the risked volume is about half of what is predicted in Kjønsberg et al. (2010). The cause of this difference between the two models is that we have included two more failure LF-classes in the current model, i.e. low-saturated gas and hot-shale. This results in a model that is more likely to contain two potential reservoir layers than three.

Case	only oil	only gas	gas and oil
No correlation			
Prior	2.9 (1.3, 4.8)	2.3 (0.9, 3.8)	4.7 (2.5, 7.1)
Profile A	6.1 (3.6, 9.5)	5.6 (2.6, 8.7)	10.5 (6.2, 15.0)
Profile B	6.1 (3.3, 9.2)	6.2 (3.0, 9.0)	10.5 (6.4, 15.0)
Profile C	5.4 (2.6, 8.6)	4.3 (2.5, 6.7)	8.6 (4.9, 12.6)
Range=10 m			
Prior	4.6 (1.4, 9.1)	3.8 (1.1, 7.5)	6.9 (2.4, 12.5)
Profile A	12.2 (3.6, 22.6)	13.7 (4.3, 24.1)	20.4 (8.6, 32.5)
Profile B	8.7 (3.1, 15.9)	11.0 (4.9, 19.3)	16.2 (7.7, 25.7)
Profile C	5.8 (2.5, 11.0)	5.0 (2.4, 9.2)	8.1 (3.3, 14.0)
Range=50 m			
Prior	4.6 (1.4, 9.3)	3.9 (1.2, 7.7)	7.0 (2.4, 12.7)
Profile A	12.3 (3.4, 24.0)	13.4 (4.2, 24.1)	20.1 (8.2, 32.1)
Profile B	11.2 (4.0, 20.4)	11.8 (4.6, 20.4)	17.5 (8.4, 26.8)
Profile C	6.2 (2.5, 12.0)	6.2 (2.5, 10.8)	9.6 (3.3, 18.3)

Table 5: Prior and posterior HC volume (m) distributions given HC discovery. The first four rows is for a case with no vertical correlation in the prior model. The next four row is for range 10 m. The last four rows are for range 50 m. The first value is the mean value, and the two values in the parentheses are the 10% and 90% quantiles of the distribution.

	No correlation	Range=10 m	Range=50 m	Kjønsberg et al. (2010)
Prior	4.6	4.6	4.6	7.9
Profile A	10.3	16.9	16.3	19.4
Profile B	10.4	12.8	13.9	25.2
Profile C	8.2	3.3	4.6	5.7

Table 6: Risked volumes (m) of hydrocarbon.

3.3 Performance of the simulation algorithm

To evaluate the performance of the MCMC algorithm we focus on profile B. We run ten independent chains with $N = 20000$ iteration. The mean time per iteration is 5.68 seconds and the initial burn–in is completed in approximately 10 minutes.

To quantify how efficiently the algorithm estimates properties of the posterior distribution we compute the average discrepancy over the positions $i = (21, \ldots, 80)$ by defining,

$$\delta_\nu = 1/60 \sum_{i=21}^{80} \delta_{i,\nu}, \tag{15}$$

where the local discrepancy $\delta_{i,\nu}$ is defined in expression (14).

In Figure 15 we have plotted the discrepancy δ_ν as a function of computing time. All ten chains experience the same type of convergence pattern and speed. On average the chains use 3 hours and 15 minutes to reduce the discrepancy below the threshold 0.02. The convergence of the methodology in Kjønsberg et al. (2010) is presented in the same figure, and use an average time of 35 hours to achieve the same accuracy. Thus the proposed method improves the speed by a factor larger than 10 even though two more LF-classes are considered.

4 Conclusions

We use a Bayesian framework to integrate seismic restack data in the risk assessment of prospects. We propose a model that is easy to parameterize and can model complex spatial dependencies. We do the risk assessment by generating multiple realizations from the posterior distribution. By using a spatial model, we are able to separate presence and volume. This is essential in the exploration phase.

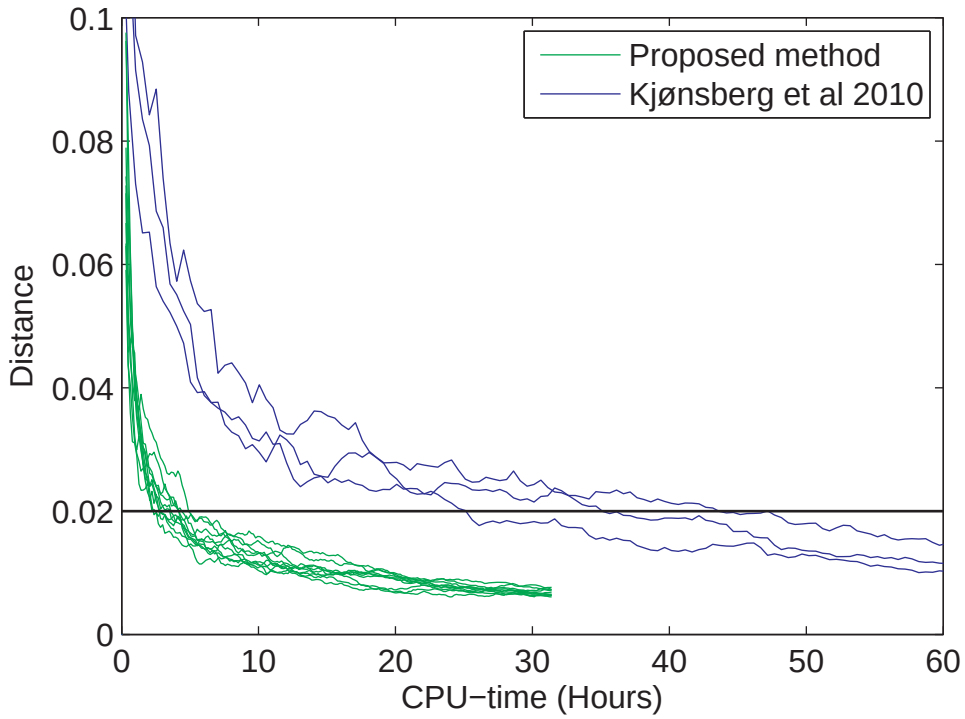

Figure 15: The panel show δ_ν as function of the CPU–time. The convergence for ten chains using the proposed methodology is displayed in green, the blue lines show the convergence for three chains using the methodology in Kjønsberg et al. (2010).

The algorithm we use to generate realizations from the posterior distribution, provides reliable results even in a complex case with many LF-classes, and obtain a good representation of the posterior distribution within a few hours. When compared to the method presented in Kjønsberg et al. (2010) the current approach is about ten times faster, even though more LF-classes are considered. Since the approach still takes hours to run it is best suited to evaluate a few selected vertical profiles.

In the Bayesian setting the inversion results are influenced by both the prior model and the seismic amplitudes. Comparing our risk updates to Kjønsberg et al. (2010) we highlight the importance of taking all LF-class scenarios into account in the prior model. Leaving out failure lithologies from the prior models result in too optimistic volume predictions. By comparing the inversion results in different profiles, we find that the seismic amplitudes make a significant impact on the inversion result even if the data has a moderate signal to noise ratio. The sensitivity to data is severely reduced for models that assumes independence, thus the vertical dependency in facies and elastic parameters gives an important contribution to volume estimates.

The use of vertical correlation in elastic parameters has no systematic effect on hydrocarbon presence, but influence the upside potential in the volumes. The upper limits for the volume is generally larger. In our cases we found that the upper limit is increased by an average of 12% when the correlation is used. The use of vertical dependency in elastic parameters does not ad computational cost over the independence assumption. It is therefore recommended to use this correlation in order to build as as much information as possible into the model.

5 ACKNOWLEDGMENTS

We would like to thank Statoil for the permission to publish the data included in this paper.

References

Aki, K., and Richards, P. G., 1980, Quantitative seismology: Theory and methods: W. H. Freeman and Company.

Avseth, P., Mukerji, T., and Mavko, G., 2005, Quantitative seismic interpretation -applying rock physics tools to reduce interpretation risk: Cambridge University Press.

Bosch, M., Cara, L., Rodrigues, J., Alonso Navarro, A., and Díaz, M., 2007, A Monte Carlo approach to the joint estimation of reservoir and elastic parameters from seismic amplitudes: Geophysics, **72**, no. 6, O29–O39.

Bosch, M., Carvajal, C., Rodrigues, J., Torres, A., Aldana, M., and Sierra, J., 2009, Petrophysical seismic inversion conditioned to well-log data: Methods and application to a gas reservoir: Geophysics, **74**, no. 2, O1–O15.

Buland, A., and Omre, H., 2003, Bayesian linearized AVO inversion: Geophysics, **68**, 185–198.

Buland, A., Kolbjørnsen, O., Hauge, R., Skjæveland, Ø., and Duffaut, K., 2008, Bayesian lithology and fluid prediction from seismic prestack data: Geophysics, **73**, no. 3, C13–C21.

Castagna, J. P., Batzle, M. L., and Kan, T. K., 1993, Rock physics - the link between rock properties and avo response, *in* Castagna, J. P., and Backus, M. M., Eds., Offset-dependent reflectivity - Theory and practice of AVO analysis: Soc. Expl. Geophys.

Contreras, A., Torres-Verdin, C., Chesters, W., Kvien, K., and Globe, M., 2005, Joint stochastic inversion of petrophysical logs and 3D pre-stack seismic data to assess the spatial continuity of fluid units away from wells: Application

to a Gulf-of-Mexico deepwater hydrocarbon reservoir: SPWLA 46th Annual Logging Symposium, **46**, 1–15.

Duijndam, A. J. W., 1988a, Bayesian estimation in seismic inversion. Part I: Principles: Geophysical Prospecting, **36**, 878–898.

——— 1988b, Bayesian estimation in seismic inversion. Part II: Uncertainty analysis: Geophysical Prospecting, **36**, 899–918.

Gunning, J., and Glinsky, M. E., 2004, Delivery: an open-source model-based bayesian seismic inversion program: Computers & Geosciences, **30**, no. 6, 619–636.

Haas, A., and Dubrule, O., 1994, Geostatistical inversion-a sequential method of stochastic reservoir modelling constrained by seismie data: First Break, **12**, no. 11, 561–569.

Hammer, H., and Tjelmeland, H., 2011, Approximate forward-backward algorithm for a switching linear Gaussian model – with application to seismic inversion: Journal Computational Statistics & Data Analysis, **55**, no. 1, 154–167.

Kjønsberg, H., Hauge, R., Kolbjørnsen, O., and Buland, A., 2010, Bayesian monte carlo method for seismic predrill prospect assessment: Geophysics, **75**, no. 2, O9–O19.

Krumbein, Y., and Dacey, M., 1969, Markov chains and embedded markov chains in geology: Mathematical Geology, **1**, no. 1, 79–96.

Larsen, A. L., Ulvmoen, M., Omre, H., and Buland, A., 2006, Bayesian lithology/fluid prediction and simulation on the basis of a Markov-chain prior model: Geophysics, **71**, no. 5, R69–R78.

Liu, J. S., 2001, Monte carlo strategies in scientific computing: Springer, Berlin.

Mavko, G., and Mukerji, T., 1998, A rock physics strategy for quantifying uncertainty in common hydrocarbon indicators: Geophysics, **63**, no. 1, 1997–2008.

Merletti, G., and Torres-Verdin, C., 2006, Accurate detection and spatial delineation of thin-sand sedimentary sequences via joint stochastic inversion of well logs and 3D pre-stack seismic amplitude data: SPE, , no. SPE 102444, 1–17.

Scales, J., and Tenorio, L., 2001, Prior information and uncertainty in inverse problems: Geophysics, **66**, no. 2, 389–397.

Scott, A. L., 2002, Bayesian methods for hidden Markov models: Recursive computation in the 21st century: Journal of the American Statistical Association, **97**, 337–351.

Torres-Verdin, C., Victoria, M., Merletti, G., and Pendrel, J., September 1999, Trace-based and geostatistical inversion of 3-D seismic data for thin-sand delineation: An application in San Jorge Basin, Argentina: The leading edge, pages 1070–1077.

Ulrych, J., Sacchi, M., and Woodbury, A., 2001, A bayes tour of inversion: A tutorial: Geophysics, **66**, no. 1, 55–69.

Ulvmoen, M., and Hammer, H., 2009, Bayesian lithology/fluid inversion – comparison of two algorithms: Computational Geosciences (accepted).

A The proposal distribution in the MCMC algorithm

Here we present the proposal distribution $q(\mathbf{f}, \mathbf{m}|\mathbf{c})$ in the second step of the MCMC algorithm. For further details, see Hammer and Tjelmeland (2011). Ideally we want to generate proposals from $p(\mathbf{f}, \mathbf{m}|\mathbf{c})$, because proposals from this distribution always will be accepted, see expression (9). We are only able to generate proposals from $p(\mathbf{f}, \mathbf{m}|\mathbf{c}, \mathbf{d})$ for very low dimensions of n. In higher dimensions we introduce an approximation to this distribution which we will use as our proposal distribution $q(\mathbf{f}, \mathbf{m}|\mathbf{c})$.

When the reflection coefficients are known, the seismic data does not bring any additional information about the elastic parameters this is formalized in the relation $p(\mathbf{f}, \mathbf{m}|\mathbf{c}, \mathbf{d}) = p(\mathbf{f}, \mathbf{m}|\mathbf{c})$. Thus this is our target distribution. It can be written as

$$\begin{aligned} p(\mathbf{f}, \mathbf{m}|\mathbf{c}) \propto & \left[p(f_1) \prod_{i=2}^{n} p(f_i|f_{i-1}) \right] \times \\ & \left[p(\mathbf{m}_1|f_1) \prod_{i=2}^{n} p(\mathbf{m}_i|f_{i-1}, f_i, \mathbf{m}_{i-1}) \right] \times \\ & \left[p(\mathbf{c}_{\cdot,1}) \prod_{i=2}^{n} p(\mathbf{c}_{\cdot,i}|\mathbf{m}_{i-1}, \mathbf{m}_i) \right] \end{aligned} \tag{16}$$

and is a first order hidden Markov model. Here $\mathbf{c}_{\cdot,i}$ represents the reflection coefficients for all the offset angels at position i. A directed acyclic graph (DAG) for the relation between the variables is given in Figure 16. In the statistical literature realizations from distributions that has the form of expression (16) is obtained by constructing a forward–backward algorithm, see Scott (2002). The forward part sequentially integrates out $\mathbf{m}_i$ and f_i for $i = 1, \ldots, n$. After $(f_i, \mathbf{m}_i), i = 1, \ldots, k$ are integrated out, the remaining distribution $p(f_{k+1}, \ldots, f_n, \mathbf{m}_{k+1}, \ldots, \mathbf{m}_n|\mathbf{c})$ is a mixture of L^k Gaussian densities. We see that the number of mixture terms grows exponentially with k. Thus, we are only able to handle this distributions for very low values of n because

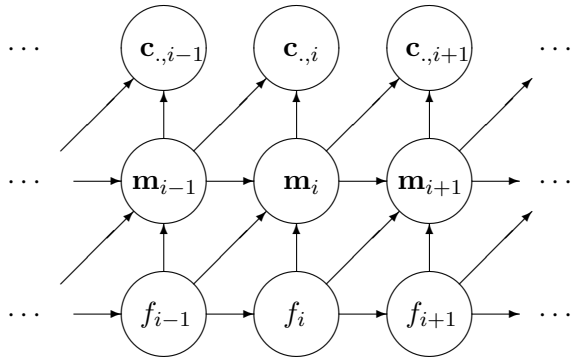

Figure 16: Directed acyclic graph (DAG) representation for the relation between the variables in $\mathbf{f}$, $\mathbf{m}$ and $\mathbf{c}$.

of memory limitations. For higher dimensions of n, we instead introduce an approximate forward integration procedure, where we ignore the less important Gaussian terms, keeping a number of mixture terms such that computer memory is not exceeded. Thereafter, the backward simulation is computationally straight forward. We use the probability distribution defined by this approximate forward-backward procedure as the proposal distribution $q(\mathbf{f}, \mathbf{m}|\mathbf{c})$.