Computational Geosciences manuscript No.
(will be inserted by the editor)

Bayesian lithology–fluid inversion - algorithm efficiency

Marit Ulvmoen, Hugo Hammer*

Norwegian University of Science and Technology
N-7491 Trondheim
Norway

Received: 02.04.09 / Revised version: date

Abstract Algorithms for inversion of seismic prestack AVO data into lithology-fluid classes in a vertical profile are evaluated. The inversion is defined in a Bayesian setting where the prior model for the lithology-fluid classes is a Markov chain, and the likelihood model relates seismic data and elastic material properties to these classes. The likelihood model is approximated such that the posterior model can be calculated recursively using the extremely efficient forward-backward algorithm. The impact of the approximation in the likelihood model is evaluated empirically by comparing results from the approximate approach with results generated from the exact posterior model. The exact posterior is assessed by sampling using a sophisticated Markov chain Monte Carlo simulation algorithm. The simulation algorithm is iterative, and it requires considerable computer resources. Seven realistic evaluation models are defined, from which synthetic seismic data are generated. Using identical seismic data, the approximate marginal posterior is calculated and the exact marginal posterior is assessed. It is concluded that the approximate likelihood model preserves 50% to 90% of the information content in the exact likelihood model.

Key words Seismic inversion – Lithology-fluid prediction – Empirical evaluation – Bayesian model – Forward-backward algorithm

1 Introduction

Inversion of seismic AVO data into lithology-fluid (LF) characteristics in a petroleum reservoir is important both for exploration and production. In a Bayesian setting, prior information about the LF characteristics can be combined with rock physics and seismic likelihood models linking the seismic data to these characteristics. The posterior model contains the complete solution in the Bayesian setting. For Bayesian LF inversion approaches, see Eidsvik et al. [5], Avseth et al. [2], Larsen et al. [8], Hammer and Tjelmeland [7], González et al. [6], Buland et al. [3], Ulvmoen and Omre [10] and Ulvmoen et al. [11]. In the current study, we focus on the approaches in Larsen et al. [8] and Hammer and Tjelmeland [7].

In Larsen et al. [8], the prior for the LF classes is defined by a Markov chain model upward through a vertical profile. The likelihood model is defined by a rock physics term relating the LF classes to elastic material properties, and a seismic forward term relating these properties to seismic AVO data. The likelihood model is approximated such that the posterior follows a Markov chain model. With the approximate posterior on this Markov chain form, it can be calculated exactly using the recursive forward-backward algorithm which is extremely computer efficient. The speed of the algorithm is important especially if the methodology is extended into 3D, see Ulvmoen and Omre [10] and Ulvmoen et al. [11]. The impact of the approximation in the likelihood model on the results is unknown.

In Hammer and Tjelmeland [7], the seismic inversion is defined using the same prior and likelihood models as in Larsen et al. [8]. The inversion is, however, solved using the exact posterior model without any approximation. Direct sampling from the posterior model is infeasible; hence a sophisticated Markov chain Monte Carlo (McMC) algorithm is defined where changes for all the model variables are proposed in each location. The posterior is assessed by sampling which requires considerable computer resources. The McMC algorithm is iterative and it converges in the limit. The convergence rate is such that the simulation algorithm is feasible but still computer demanding.

The objective of the study is to evaluate the approximation within the likelihood model in Larsen et al. [8]. This is done empirically by comparing inversion results from the methodology in Larsen et al. [8] with results from the methodology in Hammer and Tjelmeland [7]

* *Present address:* Høgskolen i Oslo, N-0130 Oslo, Norway

using identical seismic data. By comparing results from the approximate posterior with results from the exact posterior, the impact of the approximation can be evaluated.

2 Stochastic model

The LF characteristics along a vertical profile through a reservoir target zone are of primary interest in the study. The LF characteristic in location t is denoted by π_t, and it can take one of the L classes $\pi_t \in \{\pi^1, \ldots, \pi^L\}$. The complete set of LF characteristics in the profile is denoted by $\boldsymbol{\pi} = \{\pi_1, \ldots, \pi_T\}$ with T defining the profile length. The inversion is of seismic prestack AVO data into LF characteristics. We denote these seismic data along the vertical profile by $\mathbf{d}$, and they contain seismic samples for a set of n reflection angles $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$.

The term $p(\cdot)$ is used as a generic term for probability. In particular, $p(\pi_t)$ denotes the probability of the various LF classes $\pi_t \in \{\pi^1, \ldots, \pi^L\}$, and $p(\boldsymbol{\pi})$ is the multivariate probability of the complete set of LF classes. Moreover, $p(\boldsymbol{\pi}|\mathbf{d})$ denotes the conditional probability of $\boldsymbol{\pi}$ given $\mathbf{d}$.

The inversion is defined in a Bayesian setting where the complete solution is the posterior probabilistic model defined by

$$p(\boldsymbol{\pi}|\mathbf{d}) = \text{const} \times p(\boldsymbol{\pi})\, p(\mathbf{d}|\boldsymbol{\pi}) \tag{1}$$

with $p(\boldsymbol{\pi})$ being the prior model for the LF classes and $p(\mathbf{d}|\boldsymbol{\pi})$ the likelihood model relating the seismic data to these classes. The normalizing constant is usually difficult to calculate directly.

2.1 Prior model

The prior model for the LF characteristics is defined as a stationary Markov chain model upward through the vertical profile. The Markov chain model is defined by an upward transition matrix $\mathbf{P}$ and the marginal probabilities $p(\pi_1)$, with the elements in $\mathbf{P}$ being the transition probabilities $p(\pi_t|\pi_{t-1})$ for all combinations of LF classes. As the Markov chain model is stationary, the conditional elements $p(\pi_t|\pi_{t-1})$ are independent of the location t; hence the marginal probabilities $p(\pi_t)$ are identical in each location. These marginals can then be calculated from $\mathbf{P}$, which fully specifies the stationary Markov chain model. The Markov chain model is written

$$p(\boldsymbol{\pi}) = \prod_t p(\pi_t|\pi_{t-1}) \tag{2}$$

with $p(\pi_1) = p(\pi_1|\pi_0)$ for notational convenience in the rest of the paper.

2.2 Likelihood model

In order to link the LF classes to the seismic data, a set of three elastic material properties is introduced. These properties are P-wave velocity (v_p), S-wave velocity (v_s) and density (ρ). Let $\mathbf{m}_t$ represent the log-transform of the three elastic material properties in location t, and let the complete set in the vertical profile be denoted by $\mathbf{m}$. The likelihood model is defined as the integral over $\mathbf{m}$ like in Larsen et al. [8]

$$p(\mathbf{d}|\boldsymbol{\pi}) = \int p(\mathbf{d}|\mathbf{m})\, p(\mathbf{m}|\boldsymbol{\pi})\, d\mathbf{m} \tag{3}$$

where $p(\mathbf{d}|\mathbf{m})$ is a seismic likelihood model and $p(\mathbf{m}|\boldsymbol{\pi})$ is a rock physics likelihood model. The integral is over all configurations of the three elastic variables, which may be computer demanding to calculate.

The seismic data are defined by the convolution model like in Buland and Omre [4] given by

$$\mathbf{d} = \mathbf{WADm} + \mathbf{e} \tag{4}$$

where $\mathbf{W}$ is a block diagonal convolution matrix containing one wavelet for each reflection angle in $\boldsymbol{\theta}$, $\mathbf{A}$ is a matrix of angle-dependent weak contrast Aki-Richards coefficients, see Aki and Richards [1], $\mathbf{D}$ is a differential matrix giving the contrasts in $\mathbf{m}$, and $\mathbf{e}$ is an error term. We define $\mathbf{e}$ as a mixture of wavelet colored and white noise defined by the relation

$$\mathbf{e} = \mathbf{We}_1 + \mathbf{e}_2 \tag{5}$$

with $\mathbf{e}_1$ and $\mathbf{e}_2$ being Gaussian white noise given by

$$\mathbf{e}_i \sim \mathcal{N}(\mathbf{0}, \sigma_i^2\mathbf{I});\ i = 1, 2 \tag{6}$$

where $\mathbf{I}$ is the identity matrix. It follows from the relations above that the seismic likelihood model is Gauss-linear given by

$$[\mathbf{d}|\mathbf{m}] \sim p(\mathbf{d}|\mathbf{m}) = \mathcal{N}(\mathbf{WADm}, \sigma_1^2\mathbf{WW}' + \sigma_2^2\mathbf{I}) \tag{7}$$

because the noise in $\mathbf{e}$ is Gaussian.

The rock physics likelihood model is factorized as

$$p(\mathbf{m}|\boldsymbol{\pi}) = \prod_t p(\mathbf{m}_t|\pi_t), \tag{8}$$

and the marginals $p(\mathbf{m}_t|\pi_t)$ are assigned Gaussian distributions

$$[\mathbf{m}_t|\pi_t] \sim \mathcal{N}(\boldsymbol{\mu}_{\mathbf{m}_t|\pi_t}, \boldsymbol{\Sigma}_{\mathbf{m}_t|\pi_t}) \tag{9}$$

with the expectation vector $\boldsymbol{\mu}_{\mathbf{m}_t|\pi_t}$ and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{m}_t|\pi_t}$ assumed to be known. In Larsen et al. [8] the rock physics likelihood model is defined locationwise using an empirical relation between $\mathbf{m}_t$ and π_t. This empirical relation leads to more general rock physics likelihood models, such that well observations can be used directly without imposing a Gaussian distribution. The inversion approach in Hammer and Tjelmeland [7] does, however, demand the rock physics likelihood model to be Gaussian. In order to compare the two approaches, we consider a Gaussian distribution in this study.

2.3 Posterior model

The posterior model is fully defined by the prior and likelihood models defined above, and given by

$$p(\boldsymbol{\pi}|\mathbf{d}) = \text{const}\times \prod_t p(\pi_t|\pi_{t-1}) \times \int p(\mathbf{d}|\mathbf{m}) \prod_t p(\mathbf{m}_t|\pi_t)\, d\mathbf{m}. \tag{10}$$

As both the seismic and rock physics likelihood models are assigned Gaussian distributions, the integral over $\mathbf{m}$ is analytically obtainable. The normalizing constant can not, however, be calculated directly as it is defined as the sum over all configurations of the LF classes.

3 Assessment of posterior model

The posterior model is fully defined in Expression (10). It contains a high-dimensional integral and a normalizing constant which may be difficult to calculate. In Larsen et al. [8], an approximation of the seismic likelihood model is defined. Using this approximation, the dimension of the integral is reduced, and the resulting posterior is a non-stationary Markov chain model. With the posterior being a Markov chain model, it can be calculated extremely fast using the forward-backward algorithm where the posterior model including the normalizing constant is calculated recursively.

In Hammer and Tjelmeland [7] no approximation is made, and samples from the posterior model are generated using a computer demanding McMC algorithm. In the simulation algorithm, direct calculation of the normalizing constant is avoided. The algorithm is iterative, and it converges in the limit. The two approaches are discussed below.

3.1 Assessment of approximate posterior model

The seismic convolution model is given as $\mathbf{d} = \mathbf{WADm}+\mathbf{e}$ in Expression (4). In Larsen et al. [8], the elastic material properties in $\mathbf{m}$ are approximated by a Gaussian distribution $p_*(\mathbf{m})$ given by

$$\mathbf{m} \sim p_*(\mathbf{m}) = \mathcal{N}(\boldsymbol{\mu}_\mathbf{m}, \boldsymbol{\Sigma}_\mathbf{m}) \tag{11}$$

where $\boldsymbol{\mu}_\mathbf{m}$ and $\boldsymbol{\Sigma}_\mathbf{m}$ can be calculated as the first two moments of

$$p(\mathbf{m}) = \sum_{\boldsymbol{\pi}} p(\boldsymbol{\pi})\, p(\mathbf{m}|\boldsymbol{\pi}). \tag{12}$$

The resulting posterior

$$p_*(\mathbf{m}|\mathbf{d}) = \text{const} \times p_*(\mathbf{m})\, p(\mathbf{d}|\mathbf{m}) \tag{13}$$

is then Gaussian with expectation vector and covariance matrix analytically obtainable as the seismic likelihood $p(\mathbf{d}|\mathbf{m})$ is Gauss-linear, see Expression (7). The seismic likelihood model is rewritten as the ratio of the Gaussian posterior and prior models, then approximated by the product

$$\tilde{p}(\mathbf{d}|\mathbf{m}) = \text{const} \times \prod_t \frac{p_*(\mathbf{m}_t|\mathbf{d})}{p_*(\mathbf{m}_t)}, \tag{14}$$

see Larsen et al. [8]. In the approximation, the spatial dependencies within $p_*(\mathbf{m}|\mathbf{d})$ and $p_*(\mathbf{m})$ are removed such that the corresponding covariance matrices in the marginals are 3×3 block diagonal containing only the intervariable dependencies. Note, however, that all the spatial dependencies are included in the calculation of $p_*(\mathbf{m}|\mathbf{d})$ prior to the approximation such that this posterior is calculated given the full profile of seismic data. The exact posterior model is given in Expression (10), and the corresponding approximate posterior model is now on product form

$$\tilde{p}(\boldsymbol{\pi}|\mathbf{d}) = \text{const} \times \prod_t p(\pi_t|\pi_{t-1}) \times \int \frac{p_*(\mathbf{m}_t|\mathbf{d})}{p_*(\mathbf{m}_t)}\, p(\mathbf{m}_t|\pi_t)\, d\mathbf{m}_t \tag{15}$$

with the integral being of dimension three which is numerically tractable. Further, as the rock physics likelihood $p(\mathbf{m}_t|\pi_t)$ is assigned a Gaussian distribution, these integrals can be calculated analytically. The approximate posterior now follows a Markov chain model; hence the marginal approximate posterior probabilities $\tilde{p}(\pi_t|\mathbf{d})$ can be calculated recursively using the forward-backward algorithm, see e.g. Scott [9]. Further, the locationwise most probable LF profile is given as the LF class in each location with highest marginal probability. A thorough description of the recursive forward-backward algorithm is given in Larsen et al. [8].

3.2 Assessment of exact posterior model

The exact solution to the inversion problem is the posterior model in Expression (10). In Hammer and Tjelmeland [7], McMC simulation is used to generate samples from the posterior model; hence direct calculation of the normalizing constant is omitted. Brute force McMC assessment is not feasible due to the complexity of the posterior model. The McMC simulation algorithm is implemented by introducing an auxiliary variable $\mathbf{z}$ defined by

$$\mathbf{z} = \mathbf{ADm} + \mathbf{e}_1 \tag{16}$$

such that the convolution model for the seismic data in Expression (4) is rewritten

$$\mathbf{d} = \mathbf{Wz} + \mathbf{e}_2. \tag{17}$$

By introducing $\mathbf{z}$, the relations above obtain Gauss-linear distributions given by

$$[\mathbf{z}|\mathbf{m}] \sim p(\mathbf{z}|\mathbf{m}) = \mathcal{N}(\mathbf{ADm}, \sigma_1^2\mathbf{I}) \tag{18}$$

and

$$[\mathbf{d}|\mathbf{z}] \sim p(\mathbf{d}|\mathbf{z}) = \mathcal{N}(\mathbf{W}\mathbf{z}, \sigma_2^2\mathbf{I}) \tag{19}$$

where $\mathbf{I}$ is the identity matrix.

The posterior is redefined by introducing the integral over $\mathbf{z}$

$$p(\boldsymbol{\pi}|\mathbf{d}) = \text{const} \times \prod_t p(\pi_t|\pi_{t-1}) \times \int \left(\int p(\mathbf{d}|\mathbf{z})\, p(\mathbf{z}|\mathbf{m})\, d\mathbf{z} \right) \prod_t p(\mathbf{m}_t|\pi_t)\, d\mathbf{m} \tag{20}$$

with all prior and likelihood models defined above. To simulate from the posterior, an McMC algorithm consisting of two steps in each iteration is defined. Firstly, a Gibbs step is used to update $\mathbf{m}$ and $\mathbf{z}$ jointly conditioned on $\boldsymbol{\pi}$ and $\mathbf{d}$. This first step is done efficiently because the resulting distribution is Gaussian. Secondly, a Metropolis-Hastings step is used. In this step, new values for $\boldsymbol{\pi}$ and $\mathbf{m}$ are proposed from a proposal distribution $q(\boldsymbol{\pi}, \mathbf{m}|\mathbf{z})$ which is an approximation of the corresponding conditional distribution $p(\boldsymbol{\pi}, \mathbf{m}|\mathbf{z})$. This conditional distribution can not be calculated directly. The proposed values are accepted with a Metropolis-Hastings acceptance probability such that the simulated classes $\boldsymbol{\pi}$ are realizations from the exact posterior $p(\boldsymbol{\pi}|\mathbf{d})$ after convergence. The McMC algorithm converges relatively fast and mixes well; hence the results are reliable. The algorithm is, however, still computer demanding compared to the approximate approach. The marginal posterior probabilities $p(\pi_t|\mathbf{d})$ are estimated by sampling based inference by counting the number of each of the LF classes in each location after convergence of the McMC simulation algorithm. We term this the exact posterior $p(\boldsymbol{\pi}|\mathbf{d})$ is spite of the McMC error. The locationwise most probable LF profile is further obtained by choosing the LF class in each location with most frequent occurrences. A thorough description of the simulation algorithm is given in Hammer and Tjelmeland [7].

4 Empirical evaluation

We evaluate the approximation in the likelihood model empirically by defining seven realistic evaluation models from which we generate synthetic seismic data. Using the synthetic data, we calculate the approximate posterior $\tilde{p}(\boldsymbol{\pi}|\mathbf{d})$ and assess the exact posterior $p(\boldsymbol{\pi}|\mathbf{d})$. The approximation is then evaluated by a set of evaluation criteria.

4.1 Evaluation models

The evaluation models are selected such that they span a set of realistic earth models. The transition matrix defining the Markov chain prior model is defined on the basis of general reservoir experience, see Larsen et al. [8], the rock physics models are chosen such that the variability corresponds with theory, see Avseth et al. [2], and the observation errors are chosen such that realistic signal-to-noise ratios are obtained.

We consider test cases of length $T = 100$. To compensate for this relatively short profile length, we generate ten independent data sets for each of the seven models. We consider the four LF classes gas-saturated sandstone, oil-saturated sandstone, brine-saturated sandstone and shale, such that $\pi_t \in \{\text{SG}, \text{SO}, \text{SB}, \text{SH}\}$. The upward transition probabilities between these classes in the prior Markov chain model is defined by the transition matrix

$$\mathbf{P} = \begin{pmatrix} 0.9441 & 0 & 0 & 0.0559 \\ 0.0431 & 0.9146 & 0 & 0.0424 \\ 0.0063 & 0.0230 & 0.9422 & 0.0284 \\ 0.0201 & 0.0202 & 0.1006 & 0.8591 \end{pmatrix}$$

with rows and columns corresponding to SG, SO, SB and SH, respectively. The corresponding marginal probabilities are $p(\pi_1) = (0.2419, 0.1552, 0.3830, 0.2199)$. The elements with zero probability in the transition matrix correspond to impossible upward transitions ensuring that SB not can be directly above SG nor SO, and that SO not can be directly above SG. We term this model a geological prior model. We also evaluate the influence of the geological prior on the results, and introduce a non-informative prior model with transition matrix

$$\mathbf{P} = \begin{pmatrix} 0.91 & 0.03 & 0.03 & 0.03 \\ 0.03 & 0.91 & 0.03 & 0.03 \\ 0.03 & 0.03 & 0.91 & 0.03 \\ 0.03 & 0.03 & 0.03 & 0.91 \end{pmatrix}$$

and marginal distribution $p(\pi_1) = (0.25, 0.25, 0.25, 0.25)$. This model is termed a uniform prior model.

The rock physics likelihood model $p(\mathbf{m}_t|\pi_t)$ is a function of $\log(v_p)$, $\log(v_s)$ and $\log(\rho)$. The expectation vectors and covariance matrices for these are

$$\boldsymbol{\mu}_{\mathbf{m}_t|\text{SG}} = (8.0522,\ 7.4922,\ 7.6880)$$
$$\boldsymbol{\mu}_{\mathbf{m}_t|\text{SO}} = (8.0707,\ 7.4716,\ 7.7295)$$
$$\boldsymbol{\mu}_{\mathbf{m}_t|\text{SB}} = (8.1211,\ 7.4668,\ 7.7460)$$
$$\boldsymbol{\mu}_{\mathbf{m}_t|\text{SH}} = (8.1664,\ 7.5464,\ 7.8456)$$

$$\boldsymbol{\Sigma}_{\mathbf{m}_t|\text{SG}} = 10^{-3} \times \begin{pmatrix} 0.9610 & 0.8879 & 0.1162 \\ 0.8879 & 1.0699 & 0.1032 \\ 0.1162 & 0.1032 & 0.1352 \end{pmatrix}$$

$$\boldsymbol{\Sigma}_{\mathbf{m}_t|\text{SO}} = 10^{-3} \times \begin{pmatrix} 0.7279 & 0.7796 & 0.0930 \\ 0.7796 & 1.0513 & 0.0858 \\ 0.0930 & 0.0858 & 0.0804 \end{pmatrix}$$

$$\boldsymbol{\Sigma}_{\mathbf{m}_t|\text{SB}} = 10^{-3} \times \begin{pmatrix} 0.4688 & 0.6440 & 0.0783 \\ 0.6440 & 1.0631 & 0.0819 \\ 0.0783 & 0.0819 & 0.0637 \end{pmatrix}$$

$$\boldsymbol{\Sigma}_{\mathbf{m}_t|\text{SH}} = 10^{-3} \times \begin{pmatrix} 1.8981 & 2.9115 & 0.6157 \\ 2.9115 & 4.6322 & 0.9438 \\ 0.6157 & 0.9438 & 0.2286 \end{pmatrix}$$

with rows and columns corresponding to $\log(v_p)$, $\log(v_s)$ and $\log(\rho)$, respectively. Figure 1 contains 500 independent samples of v_p, v_s and ρ generated from $p(\mathbf{m}_t|\pi_t)$. The pattern corresponds with rock physics theory, see e.g. Avseth et al. [2]. The rock physics model with the parametrization above is termed medium variability. We evaluate the impact of the rock physics variability on the inversion results by using one rock physics model with small variability and one with large variability. For the first case the covariance matrices above are divided by two, and for the latter the covariance matrices are multiplied by two.

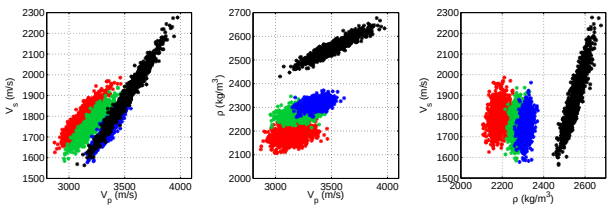

Fig. 1 Independent realizations of v_p, v_s and ρ simulated from the rock physics likelihood model $p(\mathbf{m}_t|\pi_t)$ with SG, SO, SB and SH colored red, green, blue and black, respectively

The seismic data are defined by the convolution model $\mathbf{d} = \mathbf{WADm} + \mathbf{e}$, see Expression (4). The data contain seismic samples for a set of reflection angles, and we consider the five angles $\boldsymbol{\theta} = (0, 10, 20, 30, 40)$. To construct $\mathbf{W}$, we use a Ricker wavelet with frequency $\phi = 0.11$ Hz and length 21, see Figure 2, for all the reflection angles.

The observation error is given by the relation $\mathbf{e} = \mathbf{We}_1 + \mathbf{e}_2$, see Expression (5). We let $\sigma_2 = 0.01\sigma_1$ for all noise levels, such that the colored part is most influential.

		Noise level			
		No: NN SN 4.2	Small: NS SN 2.2	Medium: NM SN 1.3	Large: NL SN 0.53
Rock physics variability	Small: VS	-	-	VS	-
	Medium: VM	NN	NS	BC	NL
	Large: VL	-	-	VL	-

Table 1 Evaluation models used in empirical study

We evaluate the impact of observation error on the model, and consider noise levels chosen such that a wide range of signal-to-noise (SN) ratios are obtained. The SN ratios are calculated as the ratio of the variance in the signal divided by the variance in the noise; hence low values indicate a large noise component while high values indicate a small noise component. We let the noise term contain the variance in $\mathbf{m}$ in addition to the variance in $\mathbf{e}$, and the signal contain the seismic convolution without variance in $\mathbf{m}$ nor $\mathbf{e}$.

By varying the SN ratios and the variability in the rock physics model, we obtain the six models described in Table 1. For all these models, the geological prior model is used. For the uniform prior we only evaluate the base case model, BC, which is the model with medium variability in the rock physics model and SN 1.3.

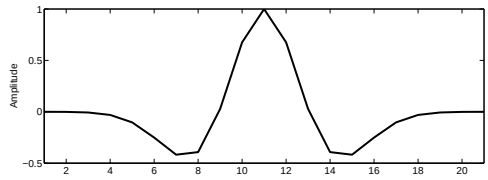

Fig. 2 Ricker wavelet with frequency $\phi = 0.11$ Hz and length 21

4.2 Evaluation criteria

Focus of the study is on the approximation in the likelihood model. We have defined seven realistic evaluation models, from which we generate synthetic seismic data. Using these data, we calculate the approximate marginal posterior $\tilde{p}(\pi_t|\mathbf{d})$ and assess the exact marginal posterior $p(\pi_t|\mathbf{d})$, then evaluate the approximation using criteria defined below.

We generate ten independent reference profiles $\boldsymbol{\pi}^R$ from the prior model for each of the seven evaluation models. Synthetic seismic data are then generated for each reference profile.

It should be kept in mind that focus is on the quality of the approximation of the exact likelihood model. The information content about the reference LF profile in the exact posterior will vary in the different evaluation models. The ability of the approximate model to reflect information content in the exact posterior is much more important than the ability to reproduce the reference profile.

For one of the reference profiles in each evaluation model, the marginal posterior, a set of realizations and the locationwise most probable LF prediction are displayed using both inversion approaches. The marginal posteriors provide the basis for the prediction and are associated with the prediction uncertainty. If the approximation is perfect the approximate and exact marginal posteriors have identical shape, and in perfect inversion the marginal probabilities have value one for the reference class and zero for the other classes in each location. The realizations represent the variability in the posteriors. Ideally, the variability in the realizations from the

approximate and exact approaches is identical. The locationwise most probable LF prediction is the LF profile with highest marginal posterior probability in each location. If the approximation is perfect the locationwise most probable LF predictions are identical for the two approaches, and in perfect inversion the most probable LF predictions are identical to the reference profile.

We define a confusion rate matrix $C = [c_{i,j}]$ with elements

$$c_{i,j} = \frac{\sum_{t=1}^{T} \tilde{p}(\pi_t = j|\mathbf{d})\, \mathrm{I}(\pi_t^R = i)}{\sum_{t=1}^{T} p(\pi_t = j|\mathbf{d})\, \mathrm{I}(\pi_t^R = i)} \tag{21}$$

where $I(A)$ is an indication function taking value 1 if A is true and 0 otherwise, and where i and j indicate the different LF classes in $\pi_t \in \{\mathrm{SG}, \mathrm{SO}, \mathrm{SB}, \mathrm{SH}\}$. Row i in the matrix contains the ratio of the approximate and the exact posterior probabilities summed over the locations where i is the true class in the reference profile $\boldsymbol{\pi}^R$. Note that $p(\boldsymbol{\pi}|\mathbf{d})$ is the exact solution; hence the elements give the relative deviation of the approximate model from this exact solution. If the approximation is perfect, the numerator and denominator are equal in all elements of the confusion rate matrix such that the approximate posterior captures all the information in the exact posterior. In perfect inversion, the numerator and denominator are both equal to one on the diagonal and zero in the rest of the matrix.

The probability of correct classification is denoted δ and defined by

$$\delta_p = \frac{1}{T} \sum_{t=1}^{T} p(\pi_t = \pi_t^R|\mathbf{d}). \tag{22}$$

This probability calculated from the the exact posterior $p(\boldsymbol{\pi}|\mathbf{d})$, the approximate posterior $\tilde{p}(\boldsymbol{\pi}|\mathbf{d})$ and the prior $p(\boldsymbol{\pi})$ is denoted by δ_p, $\delta_{\tilde{p}}$ and δ_π, respectively. If the approximation is perfect, δ calculated from the two posterior approaches are equal, and in perfect inversion they are equal to one. The probability of correct classification from the prior model, δ_π, contains the marginal prior probabilities weighted by the number of each of the LF classes in the reference $\boldsymbol{\pi}^R$. Ideally, δ_p and $\delta_{\tilde{p}}$ should be much larger than δ_π.

The approximation is within the likelihood model; hence the amount of information in the exact likelihood captured by the approximation is of interest. The probability of correct classification coming from the likelihood model only is calculated by the difference $\delta_p - \delta_\pi$. Then, by the ratio

$$\Delta = \frac{\delta_{\tilde{p}} - \delta_\pi}{\delta_p - \delta_\pi} \tag{23}$$

the amount of this probability in the exact likelihood captured by the approximate likelihood is given. If the approximation is good, the numerator is equal to the denominator; hence the ratio equals one.

4.3 Results with discussion

Figures 3 to 9 contain marginal posteriors, 200 independent realizations and locationwise most probable LF predictions from the exact and approximate approach for each of the seven evaluation models. For each model, one of the ten reference realizations is shown. The marginal posteriors for the approximate and exact approaches tend to have similar shapes, often with probabilities close to one for the reference class, and probabilities close to zero for the rest of the classes. The probabilities in the approximate posterior $\tilde{p}(\boldsymbol{\pi}|\mathbf{d})$ are often smoother between the layers than the exact posterior $p(\boldsymbol{\pi}|\mathbf{d})$, making the transitions more non-distinct. Most of the uncertainty is between the layers in both approaches. In the approximation in the likelihood model, the spatial correlations between the elastic material properties are ignored; hence the realizations from the approximate approach are expected to have more variability than the ones from the exact approach. This is verified in all the realizations. For both posterior approaches, the amount of variability increases with increasing noise levels. The most probable LF predictions from the two approaches often look similar, but the ones from the approximate model are more heterogeneous than the ones from the exact model. The predicted profiles generated from the exact posterior tend to look more like the true LF profile than the ones from the approximate posterior.

Table 2 contains the confusion rate matrix for each of the evaluation models, calculated based on the ten independent reference realizations for each model. The diagonal elements contain the probability of correct prediction for each of the LF classes; hence these are of particular interest. The numerator is mostly smaller than the denominator in the diagonal elements such that the probability of correct classification is higher in the exact than the approximate posterior. Note, however, that this often shifts in the diagonal elements for SG. The exact posterior is the optimal solution in the Bayesian approach; hence ratios larger than one does not mean that the approximate approach is better than the exact one. What is important is how much the numerator and denominator deviate from each other. For all models with geological prior, SO is the diagonal class with largest deviation between numerator and denominator. Note, however, that SO has smallest marginal prior probability in the geological prior. In the approximate approach, regression towards the dominant classes is therefore more common than in the exact approach. Note also that this is not so in the model with uniform prior, where each class has equal marginal probability and the numerator and denominator are almost equal for SO.

Table 3 contains the probabilities of correct classification, δ_p, $\delta_{\tilde{p}}$ and δ_π, for the exact posterior, the approximate posterior and the prior, respectively. The probabilities are calculated for each of the seven models, with

		SG	SO	SB	SH
BC	SG	0.8178/0.6696	0.0843/0.3078	0.0819/0.0191	0.0160/0.0034
	SO	0.4841/0.3525	0.3632/0.5585	0.1334/0.0867	0.0193/0.0023
	SB	0.0962/0.0067	0.2295/0.0733	0.5749/0.8830	0.0994/0.0371
	SH	0.0153/0.0006	0.0153/0.0055	0.0537/0.0409	0.9157/0.9530

		SG	SO	SB	SH
NN	SB	0.9978/0.9939	0.0021/0.0061	0.0001/0.0000	0.0000/0.0000
	SO	0.5673/0.0046	0.3802/0.9898	0.0498/0.0055	0.0026/0.0000
	SB	0.0009/0.0000	0.0282/0.0001	0.9671/0.9996	0.0038/0.0003
	SH	0.0006/0.0000	0.0060/0.0002	0.0392/0.0031	0.9541/0.9966

		SG	SO	SB	SH
NS	SG	0.5592/0.8655	0.4087/0.1342	0.0213/0.0001	0.0109/0.0002
	SO	0.2861/0.1939	0.3824/0.8020	0.3213/0.0032	0.0103/0.0009
	SB	0.0055/0.0151	0.1171/0.1216	0.7687/0.8172	0.1087/0.0461
	SH	0.0024/0.0063	0.0111/0.0085	0.0702/0.0229	0.9162/0.9622

		SG	SO	SB	SH
NL	SG	0.6961/0.6545	0.1306/0.2355	0.1471/0.1073	0.0262/0.0027
	SO	0.4594/0.2646	0.2811/0.3525	0.2094/0.3668	0.0501/0.0161
	SB	0.0703/0.0489	0.0686/0.1245	0.6977/0.7472	0.1634/0.0793
	SH	0.0656/0.0147	0.0232/0.0171	0.1871/0.1400	0.7241/0.8282

		SG	SO	SB	SH
VS	SG	0.7137/0.7722	0.0812/0.1987	0.1839/0.0283	0.0212/0.0008
	SO	0.3689/0.1910	0.1506/0.4881	0.3865/0.3031	0.0939/0.0178
	SB	0.0715/0.0059	0.3228/0.0702	0.5046/0.8960	0.1011/0.0279
	SH	0.0137/0.0012	0.0331/0.0072	0.1974/0.0654	0.7558/0.9261

		SG	SO	SB	SH
VL	SG	0.8730/0.8641	0.0673/0.1143	0.0290/0.0147	0.0306/0.0069
	SO	0.3780/0.1365	0.5027/0.7423	0.0981/0.1098	0.0212/0.0113
	SB	0.0249/0.0007	0.1976/0.0350	0.6543/0.9221	0.1232/0.0421
	SH	0.0209/0.0011	0.0509/0.0061	0.1774/0.0691	0.7508/0.9237

		SG	SO	SB	SH
BC, uniform prior	SG	0.5719/0.7204	0.2990/0.2569	0.1060/0.0223	0.0231/0.0005
	SO	0.1268/0.3404	0.5347/0.5730	0.2803/0.0858	0.0582/0.0008
	SB	0.0849/0.0349	0.2386/0.1383	0.5421/0.7976	0.1344/0.0292
	SH	0.0258/0.0009	0.0788/0.0129	0.1228/0.1818	0.7726/0.8044

Table 2 Confusion rate matrix with approximate (numerator) and exact (denominator) posterior probabilities for each of the seven evaluation models

each δ calculated based on the ten reference realizations. We see that the probability of correct classification always is larger in the exact than in the approximate posterior. The final column in Table 3 contains Δ, being the relative loss in classification probability by using the approximate likelihood model. The ratio has lowest value for the model with small variability in the rock physics model; hence the amount of information in the true likelihood captured by the approximation is lowest in this model. This is also the model where $\delta_{\bar{p}}$ has lowest value indicating that this is the most difficult model for the approximate approach. Note that the noise level in $\mathbf{e}$ has to be large in this model in order to obtain SN1.3. The consequences of the approximation are also, however, large for the model with large variability in the rock physics model. We see from this that the consequences of the approximation are large in both models where the variability in the rock physics model has been altered. In the model with uniform prior both approaches are relatively poor, and the approximate model captures most of the information in the exact approach. The zero elements in the geological prior puts constraints on the model, which generally makes the classification an easier problem than with a uniform prior. The consequences of the approximation are smallest in the model with lowest SN ratio. This is also the model where the exact approach has lowest probability of correct classification. The approximate approach does not, however, have especially low probability of correct classification in this model; hence the ratio Δ has highest value here. Overall, the approximate

likelihood model preserves between 50% and 90% of the information content in the exact likelihood model.

	δ_p	$\delta_{\tilde{p}}$	δ_π	Δ
BC	0.7846	0.6647	0.2673	0.7681
NN	0.9963	0.8744	0.2765	0.8307
NS	0.8573	0.6842	0.2781	0.7010
NL	0.7029	0.6631	0.3012	0.9010
VS	0.7937	0.5343	0.2719	0.5029
VL	0.8771	0.6675	0.2829	0.6473
BC, uniform prior	0.7168	0.5975	0.2500	0.7444

Table 3 Probability of correct classification, δ_p, $\delta_{\tilde{p}}$ and δ_π, for exact posterior, approximate posterior and prior, respectively; and relative loss in classification probability by using the approximate likelihood model, Δ, for each of the seven evaluation models

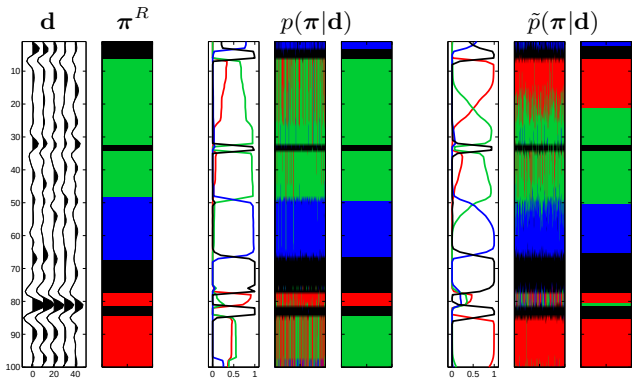

Fig. 3 Evaluation model BC: Seismic data $\mathbf{d}$ and reference LF profile $\boldsymbol{\pi}^R$ with SG, SO, SB and SH colored red, green, blue and black, respectively; marginal posterior, 200 independent realizations and locationwise most probable LF prediction from exact posterior model $p(\boldsymbol{\pi}|\mathbf{d})$; and marginal approximate posterior, 200 independent realizations and locationwise most probable LF prediction from approximate posterior model $\tilde{p}(\boldsymbol{\pi}|\mathbf{d})$

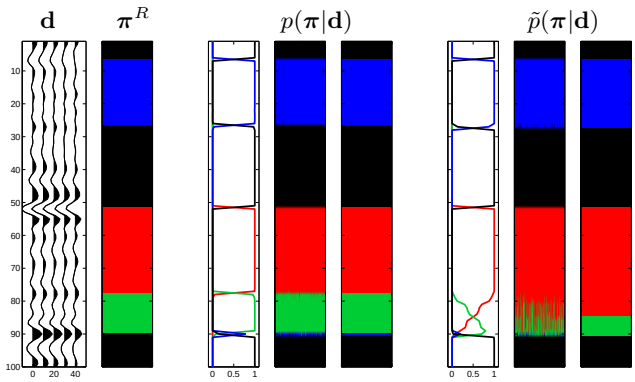

Fig. 4 Evaluation model NN: See caption in Figure 3

5 Closing remarks

The approximation within the likelihood model in Larsen et al. [8] is evaluated in an empirical study by comparing inversion results from the methodology in Larsen et al. [8] with results from the exact approach in Hammer and Tjelmeland [7]. Seven realistic evaluation models are defined, from which synthetic seismic data are generated. Using identical seismic data, the approximate marginal posterior is calculated and the exact marginal posterior is assessed.

The shapes of the marginal approximate and exact posteriors are similar for all evaluation models; hence the approximation appears reliable. The variability in the realizations from the approximate approach is larger than the variability in the corresponding realizations from the exact approach due to the approximation in the likelihood model, where spatial correlations between the elastic material properties are ignored. For both approximate and exact approaches, the amount of variability increases with increasing noise levels. Regression towards the dominant class is more common in the approximate than the exact approach. The probability of correct classification is larger for the exact than the approximate approach in all the evaluation models, and the consequences of the approximation is largest for the models where the variability in the rock physics model has been altered.

The main result of the study is that the approximate likelihood preserves between 50% and 90% of the information content in the exact likelihood model. The approximate approach therefore appears as reliable for realistic LF inversions, although the exact posterior provides somewhat better results.

It takes more than thousand times more computing time to generate results from the McMC simulation algorithm than from the recursive forward-backward algorithm. Extension of the McMC simulation algorithm into large 3D target zones will not be feasible due to the computer demanding algorithm. The approximate methodology in Larsen et al. [8] is possible to extend to 3D where an iterative simulation algorithm must be used to assess the posterior. The algorithm is, however, iterative only in 2D as the third dimension is calculated recursively by the forward-backward algorithm, see Ulvmoen and Omre [10].

Acknowledgements The work is funded by the URE-initiative at the Norwegian University of Science and Technology.

References

1. Aki, K. and Richards, P. G. [1980]. *Quantitative seismology: Theory and methods*, W. H. Freeman and Company.
2. Avseth, P., Mukerji, T. and Mavko, G. [2005]. *Quantitative Seismic interpretation : Applying rock physics tools to reduce interpretation risk*, Cambridge University Press.

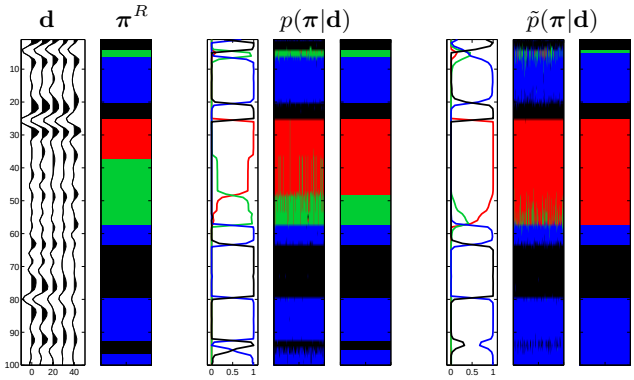

Fig. 5 Evaluation model NS: See caption in Figure 3

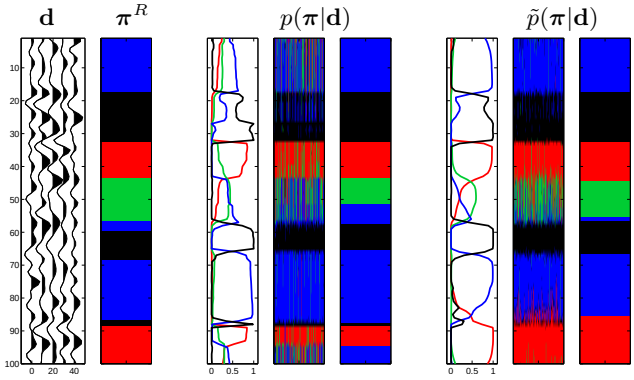

Fig. 6 Evaluation model NL: See caption in Figure 3

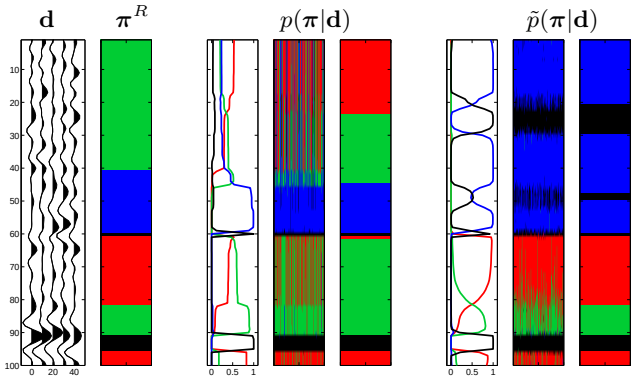

Fig. 7 Evaluation model VS: See caption in Figure 3

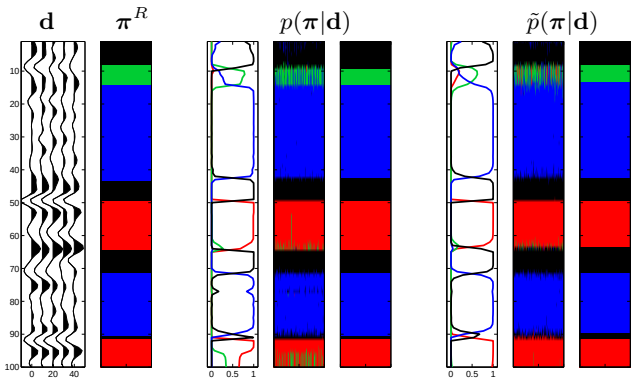

Fig. 8 Evaluation model VL: See caption in Figure 3

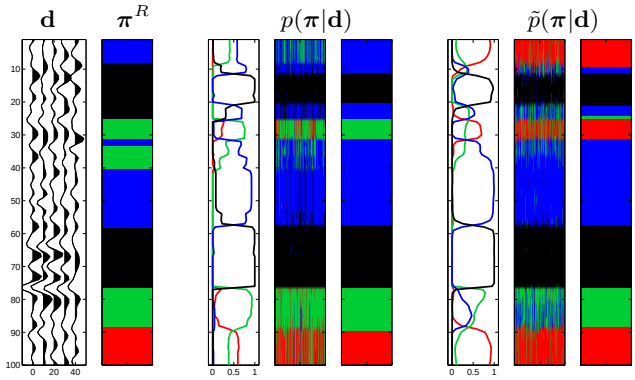

Fig. 9 Evaluation model BC with uniform prior: See caption in Figure 3

3. Buland, A., Kolbjørnsen, O., Hauge, R., Skjæveland, Ø. and Duffaut, K. [2008]. Bayesian lithology and fluid prediction from seismic prestack data, *Geophysics* **73**: C13–C21.
4. Buland, A. and Omre, H. [2003]. Bayesian linearized AVO inversion, *Geophysics* **68**: 185–198.
5. Eidsvik, J., Avseth, P., Omre, H., Mukerji, T. and Mavko, G. [2004]. Stochastic reservoir characterization using prestack seismic data, *Geophysics* **69**: 978–993.
6. González, E. F., Mukerji, T. and Mavko, G. [2008]. Seismic inversion combining rock physics and multiple-point geostatistics, *Geophysics* **73**: R11–R21.
7. Hammer, H. and Tjelmeland, H. [2008]. Approximative forward-backward algorithm for a three layer hidden Markov model - with applications to seismic inversion, *Technical Report S4-2008*, Department of Mathematical Sciences, Norwegian University of Science and Technology. Download from `http://www.math.ntnu.no/preprint/statistics/`.
8. 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 issue 5**: R69–R78.
9. Scott, A. L. [2002]. Bayesian Methods for Hidden Markov Models: Recursive Compution in the 21st Century, *Journal of the American Statistical Association* **97**: 337–351.
10. Ulvmoen, M. and Omre, H. [2009]. Improved resolution in Bayesian lithology/fluid inversion from prestack seismic data and well observations: Part I - Methodology, *Submitted for publication*. p. Preprint at `http://www.math.ntnu.no/ure/`.
11. Ulvmoen, M., Omre, H. and Buland, A. [2009]. Improved resolution in Bayesian lithology/fluid inversion from prestack seismic data and well observations: Part II - Real case study, *Submitted for publication*. p. Preprint at `http://www.math.ntnu.no/ure/`.