Bayesian Lithology-Fluid Prediction and Simulation based on a Markov Chain Prior Model

Anne Louise Larsen

Formerly Norwegian University of Science and Technology, N-7491 Trondheim, Norway;

presently Schlumberger Stavanger Research, Risabergveien 3, N-4068 Stavanger, Norway.

E-mail: ALarsen2@stavanger.oilfield.slb.com

Marit Ulvmoen

Norwegian University of Science and Technology, N-7491 Trondheim, Norway.

E-mail: ulvmoen@stud.math.ntnu.no

Henning Omre

Norwegian University of Science and Technology, N-7491 Trondheim, Norway.

E-mail: omre@math.ntnu.no

Arild Buland

Statoil ASA, Forusbeen 50, N-4035 Stavanger, Norway.

E-mail: abu@statoil.com

(May 19, 2005)

ABSTRACT

A technique for Lithology-Fluid prediction and simulation from prestack seismic data is developed in a Bayesian framework. The objective is to determine the Lithology-Fluid classes along 1D profiles through a reservoir target zone. A stationary Markov chain prior model is used to model vertical continuity of Lithology-Fluid classes along the profile. The likelihood model entails strong spatial coupling related to the convolutional model. An approximation of the likelihood model provides an approximate posterior model which is a Markov chain. The posterior can be assessed by an exact and efficient recursive sampling algorithm. The Lithology-Fluid inversion approach is evaluated on a synthetic 1D profile inspired by a North Sea sand stone reservoir. With a realistic signal-to-noise ratio in the seismic data, the results are reliable. The proportion of Lithology-Fluid classes and the identification of the interfaces between zones are largely correct. The prediction uncertainty increases if the number of zones increases and zone thicknesses decreases. The study clearly demonstrates the impact of a vertically coupled Markov model for the Lithology-Fluid classes.

INTRODUCTION

Lithology and fluid prediction (LFP) is important both in exploration and development of petroleum reservoirs. Inference about the Lithology-Fluid (LF) characteristics is usually based on general geological experience and seismic data. The fact that a variety of LF characteristics may result in identical seismic data make this inference a challenging task, actually an ill-posed inverse problem. Bayesian approaches are frequently used to solve geophysical inverse problems, and the Bayesian framework is also applied in the current study.

In Mukerji et al. (2001) and Houck (2002) Bayesian AVO approaches are used to identify different combinations of LF classes on either side of an interpreted seismic reflector. Interpreted horizons are usually defined from peak or trough amplitudes

and do typically represent the top of a potential reservoir zone. LFP along a horizon is usually done location wise and do not include spatial aspects. In Eidsvik et al (2004) the case in Mukerji et al. (2001) was extended to account for lateral spatial coupling of the LF classes along the horizon. The lateral coupling was imposed using a Markov random field prior model for the LF classes. A serious problem with standard AVO methods is that they neglect multi layer interference and do not include deconvolution.

In the current study, inference of LF classes along a vertical earth profile is made. A time window of a true amplitude processed and offset to angle transformed seismic gather is used in the inversion. A Markov chain prior model is used to model vertical coupling of the LF classes along the profile. An approximate posterior model is defined, which can be assessed by an exact and very efficient recursive algorithm. Both LF profile predictions with associated uncertainties and simulation of possible LF profiles are provided. The development draws heavily on previous work on Bayesian AVO inversion, see Buland and Omre (2003) and Buland et al. (2003). In the current study the LF inference approach is evaluated on a synthetic case inspired by a North Sea sand stone reservoir, and highly reliable predictions of the LF profile are reported. Moreover, a set of LF profile realizations representing the prediction uncertainty is provided. These realizations can be considered as possible LF profiles. The major contributions of the study is a new Bayesian LF inversion approach of prestack seismic data, with introduction of an exact, efficient recursive algorithm for assessing the posterior model.

NOTATION

The focus of the study is on the Lithology-Fluid (LF) characteristics along a vertical profile through a reservoir or a general target zone. The LF classes can for example be gas-, oil-, and brine-saturated sand stone, and shale. The LF characteristics are discretized into a set of LF classes π_t at each reflection time t corresponding to the

seismic sampling. The spatial reference is ordered from top of the target zone and down, $t = 1, \ldots, T$. The complete LF profile is represented by a vector of discrete LF classes, termed $\boldsymbol{\pi}$.

The seismic data are assumed to be true amplitude processed such that the prestack amplitudes represent band-limited primary reflection strength. The data are also assumed to be prestack migrated and offset to angle transformed. The seismic prestack gather along the profile is represented in a vector termed $\mathbf{d}$ containing the seismic samples in a target time window for a set of reflection angles.

An isotropic, elastic medium is described by three elastic parameters. In this study, we represent the elastic properties by the logarithm of the P- and S-wave velocities and density. In order to link the LF classes of interest and the available seismic prestack data, the elastic properties are defined for each t along the profile. Let $\mathbf{m}_t$ represent the elastic properties at time t, and let the elastic properties along the complete profile be represented in a $3T$ dimensional vector termed $\mathbf{m}$.

The term $p(\cdot)$ is used as a generic term for probability. Hence $p(\pi_t)$ defines the probability for the various LF classes π_t, and correspondingly, $p(\boldsymbol{\pi})$ is a multivariate probability on the vector $\boldsymbol{\pi}$. For continuous variables $p(\cdot)$ will be the corresponding probability density function (pdf).

STOCHASTIC MODEL

The objective of the study is to predict the LF profile $\boldsymbol{\pi}$ with associated uncertainty and to generate realization of it based on the seismic prestack data $\mathbf{d}$. In a Bayesian inversion setting the complete solution is represented by the posterior pdf $p(\boldsymbol{\pi}|\mathbf{d})$. This posterior may be expressed as

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

where $p(\mathbf{d}|\boldsymbol{\pi})$ is the likelihood function, $p(\boldsymbol{\pi})$ is the prior pdf, and const is a normalizing constant which usually is very difficult to compute. Note that all pdf's are

multivariate since they are defined on vectors of random variables. Given the posterior pdf $p(\boldsymbol{\pi}|\mathbf{d})$, the probability of any LF combination along the profile given the seismic prestack data, can be determined. In particular, the most probable LF profile can be determined. Moreover, a set of realizations of the LF profiles representing the prediction uncertainty, can be generated. These realizations can be considered as possible LF profiles. The problem is, however, that the posterior pdf can only in very few cases be computed exactly.

Likelihood Model

The likelihood function, $p(\mathbf{d}|\boldsymbol{\pi})$, specifies how likeliy each LF profile $\boldsymbol{\pi}$ is given the prestack data $\mathbf{d}$. It contains both forward models representing the physical relation between $\boldsymbol{\pi}$ and $\mathbf{d}$, and associated uncertainties. The likelihood function above may naturally be decomposed into a rock physics likelihood function $p(\mathbf{m}|\boldsymbol{\pi})$ representing the link from LF classes $\boldsymbol{\pi}$ to elastic properties $\mathbf{m}$ (Mavko et al., 1998), and a seismic likelihood function $p(\mathbf{d}|\mathbf{m})$ representing the forward model link from $\mathbf{m}$ to seismic prestack data $\mathbf{d}$. The direct link from $\boldsymbol{\pi}$ to $\mathbf{d}$ is obtained by the integral over the $3T$ dimensional vector $\mathbf{m}$

$$p(\mathbf{d}|\boldsymbol{\pi}) = \int \cdots \int p(\mathbf{d}|\mathbf{m})\, p(\mathbf{m}|\boldsymbol{\pi})\, d\mathbf{m}. \tag{2}$$

The rock physics likelihood function can be factorized as

$$p(\mathbf{m}|\boldsymbol{\pi}) = \prod_t p(\mathbf{m}_t|\pi_t), \tag{3}$$

since the rock physics models only include variables in the same spatial location. For a given LF class π_t, the elastic properties $\mathbf{m}_t$ will normally vary around a central representative value making $p(\mathbf{m}_t|\pi_t)$ a uni-modal pdf. Randomizing over π_t will provide the marginal pdf $p(\mathbf{m}_t)$ which usually is multi-modal with one mode corresponding to each LF class.

The seismic likelihood function $p(\mathbf{d}|\mathbf{m})$ is defined by the convolutional model. For a seismic time-angle gather, the convolutional model can be written in matrix-vector form (Buland and Omre, 2003)

$$\mathbf{d} = \mathbf{s} + \mathbf{e} = \mathbf{G}\mathbf{m} + \mathbf{e}, \tag{4}$$

where $\mathbf{s}$ represents the seismic prestack signal and $\mathbf{e}$ is an additive error term. The modeling matrix $\mathbf{G}$ is defined by

$$\mathbf{G} = \mathbf{WAD}, \tag{5}$$

where $\mathbf{W}$ is a matrix containing one wavelet for each angle trace, $\mathbf{A}$ is the matrix of angle dependent weak contrast Aki-Richards coefficients (Aki and Richards, 1980), and $\mathbf{D}$ is a differential matrix giving the contrasts of the elastic properties. The elastic properties $\mathbf{m}$ are represented by the logarithm of the P-wave velocity, S-wave velocity and the density along the profile. If both the error term $\mathbf{e}$ and the prior model for the elastic properties $\mathbf{m}$ are Gaussian, then the distribution for $\mathbf{m}$ given $\mathbf{d}$ is Gaussian with analytical expressions for the expectation vector and covariance matrix, see Buland and Omre (2003). The pdf $p(\mathbf{m}|\mathbf{d})$ represents the complete solution of an efficient Bayesian AVO inversion method.

Recall that the pdf $p(\mathbf{m})$ usually has multi-modal marginal pdf's $p(\mathbf{m}_t)$ and this is mathematically inconvenient when the posterior model shall be assessed. Therefore, let the prior distribution for $\mathbf{m}$ be approximated by a Gaussian pdf $p_*(\mathbf{m})$ with uni-modal marginal pdfs. This Gaussian pdf is parametrized by $(\boldsymbol{\mu}_*, \boldsymbol{\Sigma}_*, c_*(\tau))$ being expectation vector, inter-variable covariance matrix and spatial correlation function, respectively. Under the Gaussian prior model for $\mathbf{m}$ and the seismic likelihood model defined above, the corresponding posterior pdf for AVO inversion,

$$p_*(\mathbf{m}|\mathbf{d}) = \text{const} \times p(\mathbf{d}|\mathbf{m})\, p_*(\mathbf{m}), \tag{6}$$

is Gaussian with conditional expectation vector and covariance matrix analytically obtainable, see Buland and Omre (2003). Later, we will use that the seismic likelihood

function can be expressed by

$$p(\mathbf{d}|\mathbf{m}) = \text{const} \times \frac{p_*(\mathbf{m}|\mathbf{d})}{p_*(\mathbf{m})}. \tag{7}$$

Prior Model

The prior model for the LF vector $p(\boldsymbol{\pi})$ can in principle be any valid pdf describing the a priori knowledge of the LF profile before the seismic data is analyzed. The prior model may be subjectively defined based on general knowledge, or it may be estimated directly from a representative well through the target zone. Vertical continuity in the LF profile may be imposed by specifying spatial coupling in the prior model. Since the current model is defined for the 1D case, a Markov chain prior model for the LF vector may be used. This choice makes it possible to define an efficient algorithm to assess an approximation to the posterior pdf as demonstrated in the next section. In reservoir geology it may be natural to define this Markov chain upwards through the geological sequences, see Krumbein and Dacey (1969), hence opposite the ordering of the spatial reference defined in this study. The basic assumption for Markov chains is that the probability of an LF class to occur at level t, given the complete LF sequence below it, is only dependent on the LF class present at the level immediately below, i.e. at $t+1$. The Markov prior model is defined by a stationary transition matrix $\mathbf{P}$ and the marginal pdf $p(\pi_T)$. The elements of $\mathbf{P}$ are the conditional pdfs $p(\pi_t|\pi_{t+1})$ for all LF class combinations, but the stationarity assumption entails that the conditional pdfs are independent of t. The stationary $\mathbf{P}$ defines a stationary marginal pdf for the Markov chain, denoted by $p_s(\pi)$, which also is independent of t. In the current study, the marginal pdf $p(\pi_T)$ is set identical to $p_s(\pi)$. Note that under these assumptions, the LF sequence downwards along the profile, i.e. in the direction of the spatial reference defined here, will also be a stationary Markov chain. The associated transition matrix is defined by the conditional pdfs $p(\pi_t|\pi_{t-1}) = p(\pi_{t-1}|\pi_t)\, p_s(\pi_t)/p_s(\pi_{t-1})$. The parameters to be specified for the prior

model are the elements of the stationary transition matrix $\mathbf{P}$ for the stationary Markov chain upwards through the geological sequences. Under these assumptions one may write

$$p(\boldsymbol{\pi}) = \prod_t p(\pi_t|\pi_{t-1}), \tag{8}$$

with $p(\pi_1) = p(\pi_1|\pi_0)$ for notational convenience.

Posterior Model

The likelihood and prior models define the posterior pdf for the LF profile $\boldsymbol{\pi}$ given the seismic prestack data $\mathbf{d}$ by

$$p(\boldsymbol{\pi}|\mathbf{d}) = \text{const} \times \left[\int \cdots \int \frac{p_*(\mathbf{m}|\mathbf{d})}{p_*(\mathbf{m})}\, p(\mathbf{m}|\boldsymbol{\pi})\, d\mathbf{m} \right]\, p(\boldsymbol{\pi}), \tag{9}$$

where const is a normalizing constant which is very difficult to determine. Note further that the term within the parenthesis is an integral of dimension $3T$, which normally is large. A possible approximation is to retain only the diagonal elements in the covariance matrices in the Gaussian pdf's $p_*(\mathbf{m}|\mathbf{d})$ and $p_*(\mathbf{m})$, and use the factorizations in expressions (3) and (8). This approximation provides an approximate posterior pdf

$$\begin{aligned} \tilde{p}(\boldsymbol{\pi}|\mathbf{d}) &= \text{const} \times \prod_t \left[\iiint \frac{p_*(\mathbf{m}_t|\mathbf{d})}{p_*(\mathbf{m}_t)}\, p(\mathbf{m}_t|\pi_t)\, d\mathbf{m}_t \right]\, p(\boldsymbol{\pi}), \\ &= \text{const} \times \prod_t \tilde{p}(\mathbf{d}|\pi_t)\, p(\pi_t|\pi_{t-1}), \end{aligned} \tag{10}$$

with $p(\pi_1) = p(\pi_1|\pi_0)$ for notational convenience. The integral is now of dimension three and numerically tractable. The assumptions above entails that spatial dependence in $\mathbf{m}$ is ignored. Note, however, that the term within the parenthesis still is conditioned on the complete data vector $\mathbf{d}$, and the prior model for $\boldsymbol{\pi}$ remains spatially coupled. Note also that the normalizing constant is different from the one in expression (9) and is defined by

$$\sum_{\boldsymbol{\pi}} \tilde{p}(\boldsymbol{\pi}|\mathbf{d}) = 1, \tag{11}$$

where the sum is over all possible combinations of LF classes along the profile. This will in the general case be an enormous number of combinations, and the constant can not be determined in practical settings.

Special Cases

In order to evaluate the characteristics of the approximate pdf in expression (10), consider two extreme cases of information content in the seismic prestack data $\mathbf{d}$. Assume firstly that $\mathbf{d}$ is completely unrelated to $\mathbf{m}$ and $\boldsymbol{\pi}$, i.e. it is independent of $\mathbf{m}$ and $\boldsymbol{\pi}$. Then $p_*(\mathbf{m}_t|\mathbf{d}) = p_*(\mathbf{m}_t)$ and the value of the parenthesis is one, hence the approximate posterior $\tilde{p}(\boldsymbol{\pi}|\mathbf{d})$ equals the prior $p(\boldsymbol{\pi})$ which is the correct solution. Assume secondly that $\mathbf{d}$ is very conclusive about $\boldsymbol{\pi}$, which entails that $p_*(\mathbf{m}_t|\mathbf{d})$ will have a lot of mass around $\mathbf{m}_t$ representative for the true π_t, and little mass elsewhere. Hence for the true π_t both $p_*(\mathbf{m}_t|\mathbf{d})$ and $p(\mathbf{m}_t|\pi_t)$ will be large for similar $\mathbf{m}_t$ and hence the product be large. For other π_t, $p_*(\mathbf{m}_t|\mathbf{d})$ will be almost zero where $p(\mathbf{m}_t|\pi_t)$ is large, and hence the product be small. Consequently $\tilde{p}(\boldsymbol{\pi}|\mathbf{d})$ will be large for the true $\boldsymbol{\pi}$ and small elsewhere, which is correct. This demonstrates that the approximate posterior pdf in expression (10) provides reliable solutions for extreme cases of information content in $\mathbf{d}$.

If the prior model is chosen to be without spatial coupling, then $p(\boldsymbol{\pi})$ factorizes location wise in expression (10). The factorization entails that the approximate posterior can be computed independently, location wise for each t under this uncoupled prior model as

$$\tilde{p}(\pi_t|\mathbf{d}) = \text{const}_t \times \tilde{p}(\mathbf{d}|\pi_t)\, p(\pi_t), \tag{12}$$

where const_t is a normalizing constant which can be computed from the sum over the LF classes for each t, such that

$$\sum_{\pi_t} \tilde{p}(\pi_t|\mathbf{d}) = 1. \tag{13}$$

In the current study there are four LF classes; hence const_t is easily available.

ASSESSMENT OF POSTERIOR MODEL

The approximate posterior pdf $\tilde{p}(\boldsymbol{\pi}|\mathbf{d})$ with a Markov chain prior model in expression (10) contains an unknown normalizing constant which appears to be very difficult to determine since it requires summation over all possible LF combinations in $\boldsymbol{\pi}$. The posterior pdf $\tilde{p}(\boldsymbol{\pi}|\mathbf{d})$, and its marginal posterior pdfs $\tilde{p}(\pi_t|\mathbf{d})$ for all t in particular, can however be computed by a recursive algorithm, see Scott (2002). Moreover, simulations of $\boldsymbol{\pi}$ from the approximate posterior pdf can be generated. This algorithm can be used because the approximate likelihood model factorizes and the prior model is a Markov chain which entails that the posterior model is a Markov chain. The algorithm is exact and the computing time increases linearly in the length of the profile T and quadratically in the number of LF classes.

The algorithm is termed the forward-backward recursive algorithm in Scott (2002) which translates into upward-downward recursion when applied to a seismic profile through a target zone. The algorithm has the following steps:

1. Initiate $\tilde{p}(\pi_T|\mathbf{d})$ by updating the prior $p(\pi_T)$ by likelihood $\tilde{p}(\mathbf{d}|\pi_T)$.

2. Iterate upward by updating the prior pdfs by all likelihoods below them. Then, at level 1, $\tilde{p}(\pi_1|\mathbf{d})$ is the correct marginal pdf.

3. Iterate downward by updating the corresponding upward pdfs by all the likelihoods above them. Then, at all levels, $\tilde{p}(\pi_t|\pi_{t-1}, \mathbf{d})$ are the correct transition probabilities in the non-stationary Markov chain including the likelihoods at all levels. The marginal pdf $\tilde{p}(\pi_t|\mathbf{d})$ can be determined and realizations can be generated.

The algorithm is fully defined in Appendix A.

The reliability of the approximate posterior pdf $\tilde{p}(\boldsymbol{\pi}|\mathbf{d})$ can only be evaluated empirically. A test study inspired by a North Sea sand stone petroleum reservoir is defined in order to do so.

The reference reservoir contains four LF classes, gas-, oil-, and brine-saturated sand stone, and shale. The target zone is defined from 2000 ms to 2879 ms, discretized into units of one ms, which entails that $T = 880$. The profile is considerably longer than the thickness of a typical North Sea reservoir, but the length is extended in order to obtain repeatability of many LF sequences in the test. In the implementation of the algorithm a small border zone is included to avoid border effects in the results.

The prior model for the LF classes is defined as a stationary Markov chain upwards through the reference reservoir, with transition matrix

$$\mathbf{P} = \begin{pmatrix} 0.980 & 0 & 0 & 0.020 \\ 0.015 & 0.970 & 0 & 0.015 \\ 0.002 & 0.008 & 0.980 & 0.010 \\ 0.007 & 0.007 & 0.036 & 0.950 \end{pmatrix},$$

with rows and colomns corresponding to gas-, oil-, and brine-saturated sand stone, and shale respectively. From $\mathbf{P}$ the stationary pdf $p_s(\pi)$, which represents the proportions of LF classes, can be computed and it turns out to be (0.2326,0.1558,0.3932,0.2184). Note that the transition matrix contains several zero elements representing impossible upwards transitions. In particular, brine can never be directly above neither gas nor oil, and only shale can be directly above gas. The current transition matrix is defined based on general reservoir experience. In studies of a specific reservoir $\mathbf{P}$ may be estimated from data collected in wells as demonstrated in Eidsvik et al. (2004). In Figure 1 one realization of an LF sequence generated from $\mathbf{P}$ with $p(\pi_1)$ set identical to $p_s(\pi)$ is presented in the left-most display. Note how LF sequences with sand filled

with brine-oil-gas overlaid by shale tends to appear. The sequence in Figure 1 is used as the reference LF profile and denoted by $\boldsymbol{\pi}^R$.

The rock physics likelihood function $p(\mathbf{m}_t|\pi_t)$ is a function in the three elastic properties for each LF class. The likelihood may be defined by rock physics relations like the Gassmann relation, see Mavko et al. (1998), or as a set of representative samples. In the current study 1500 samples of properties for each LF class are used, see Figure 2. A stochastic rock physics model is applied to make samples typical for North Sea reservoirs, but real measurements could just as well have been used. The averages of P- and S-wave velocity and density for each LF class are $(3156, 1804, 2184)$, $(3211, 1767, 2275)$, $(3373, 1758, 2313)$ and $(3527, 1903, 2555)$ for gas-, oil-, and brine-saturated sand stone, and shale respectively. The reference elastic property profile $\mathbf{m}^R$ is constructed from the reference LF profile by using the average values of $\mathbf{m}_t$ for each LF class and add appropriate independent heterogenity to make $\mathbf{m}^R$ appear typical, see Figure 1.

The seismic likelihood function $p(\mathbf{d}|\mathbf{m})$ is defined with a forward model as in Buland and Omre (2003). The seismic signal is modeled for angles $\theta = (0, 10, 20, 30, 40)$ degrees using Ricker wavelets with 30 Hz center frequency for all angles and white Gaussian errors are assumed. The reference seismic signal $\mathbf{s}^R$ is constructed from the reference elastic property profile and the forward model, see Figure 1. The associated reference prestack seismic data $\mathbf{d}^R$ is obtained by adding a white Gaussian error component with variance 0.0009, see Figure 1. The signal-to-noise ratio in the data turns out to be 2.3.

The objective of the test is to assess the approximate posterior pdf $\tilde{p}(\boldsymbol{\pi}|\mathbf{d}^R)$ and compare it to the reference LF profile $\boldsymbol{\pi}^R$. In order to perform the LF inversion the approximate Gaussian prior model $p_*(\mathbf{m})$, defined by the parameters $(\mu_*, \boldsymbol{\Sigma}_*, c_*(\tau))$, must be defined. The former two parameters are estimated by standard estimators from simulations of $\mathbf{m}$ based on the prior Markov chain model for $\boldsymbol{\pi}$ and the rock physics likelihood model $p(\mathbf{m}_t|\pi_t)$ as described above. The parameter values obtained

are

$$\boldsymbol{\mu}_* = (8.114, 7.493, 7.764)^T, \tag{14}$$

and

$$\boldsymbol{\Sigma}_* = \begin{pmatrix} 0.0022 & 0.0017 & 0.0023 \\ 0.0017 & 0.0022 & 0.0019 \\ 0.0023 & 0.0019 & 0.0034 \end{pmatrix}, \tag{15}$$

corresponding to the logarithm of the P-wave velocity, S-wave velocity, and density. Further a second-order exponential spatial correlation function with approximate range of 6 ms is used, see Buland and Omre (2003). The approximate likelihood $\tilde{p}(\mathbf{d}|\pi_t)$ in expression (10) includes a three-dimensional integral which is numerically determined by the average of the functional value in 1500 samples of $\mathbf{m}_t$ specified for the given π_t.

In addition to the test described above, results from noise free inversion based on the approximate posterior pdf $\tilde{p}(\boldsymbol{\pi}|\mathbf{s}^R)$ are presented. These results expose the reliability of the approximation in the likelihood model, see expression (10), since the error term is excluded. Lastly, the inversion based on a prior model without spatial coupling is performed. The results expose the importance of including spatial coupling in the prior model.

RESULTS WITH DISCUSSION

The target of the study is to assess the approximate posterior pdf $\tilde{p}(\boldsymbol{\pi}|\mathbf{d}^R)$, see expression (10), and to evaluate how well it reproduces the reference LF profile $\boldsymbol{\pi}^R$. By applying the upward-downward recursive algorithm the associated marginal pdfs $\tilde{p}(\pi_t|\mathbf{d}^R)$ for all t can be determined. The predicted LF profile, $\tilde{\boldsymbol{\pi}}$ is defined by the

location wise maximum posterior criterion

$$\tilde{\pi}_t = \text{argmax}_{\pi_t}\{\tilde{p}(\pi_t|\mathbf{d}^R)\} \text{ for all } t.$$

Hence the LF class predicted at each t is the LF class with largest marginal approximate posterior pdf. Moreover, realizations $\boldsymbol{\pi}^S$ from the approximate posterior pdf $\tilde{p}(\boldsymbol{\pi}|\mathbf{d}^R)$ can be generated by the algorithm. A set of realizations represents the uncertainty in the prediction and can be interpreted as possible outcomes of the LF profile given the prestack seismic data.

Figure 3 contains a comparison between the predicted LF profile $\tilde{\boldsymbol{\pi}}$ and the reference LF profile $\boldsymbol{\pi}^R$. A display of the reference seismic data $\mathbf{d}^R$ is also included in the figure. Moreover, a classification matrix for $\boldsymbol{\pi}^R$ versus $\tilde{\boldsymbol{\pi}}$ is displayed. Perfect prediction would make $\tilde{\boldsymbol{\pi}}$ and $\boldsymbol{\pi}^R$ identical, and all non-diagonal terms in the classification matrix would be zero. The prediction is not perfect, but it must be characterized as very reliable given a signal-to-noise ratio of 2.3.

All shale zones are correctly predicted, although one additional shale zone is introduced in a thick brine-saturated sand zone at 2630 ms. All hydrocarbon zones, i.e. sand zones with gas or oil, are identified although once with fairly wrong thickness at 2250 ms, and once with wrong hydrocarbon type at 2510 ms. On level-by-level basis, 86 % of the hydrocarbon is correctly identified. From the classification matrix it can be seen that the proportion of LF classes are largely correct, except for an increase in brine-saturated sand on the expense of oil-saturated sand. This deviation is caused by the misclassification at 2260 ms.

Figure 4 contains a display of the approximate marginal posterior pdf $\tilde{p}(\pi_t|\mathbf{d}^R)$ for all t with the reference LF profile $\boldsymbol{\pi}^R$ marked. These marginal pdfs provide the basis for the prediction and is associated with the prediction uncertainty. Perfect prediction would make the marginal pdfs binary zero-one functions, with value one in the actual reference LF class. The marginal pdfs appear as somewhat smoothed across LF class borders and this is caused by non-perfect deconvolution. The consequence is

uncertainty in the detection of interfaces between zones. In general, the LF predictions appears as very conclusive. It is so even for the few zones that are misclassified.

Figure 5 contains a comparison between the synthetic seismic signal corresponding to the predicted LF profile $\tilde{\boldsymbol{\pi}}$, and the reference seismic signal $\mathbf{s}^R$. Note that the latter includes the seismic response from the heterogeneity within each LF class. Perfect prediction would make the predicted seismic signal equal to $\mathbf{s}^R$ except for the effect related to the heterogeneity within each LF class. The two seismic signals are very similar except for at levels with transitions into and out of the few zones that are wrongly classified.

Figure 6 contains eight simulated LF profiles $\boldsymbol{\pi}^S$ from the approximate posterior pdf $\tilde{p}(\boldsymbol{\pi}|\mathbf{d}^R)$. The reference LF profile $\boldsymbol{\pi}^R$ is also presented in the right-most display. The simulated LF profiles represent the variability in the approximate posterior pdf. Ideally this variability should be identical to the variability in the correct posterior pdf in expression (9), but this identity is difficult to demonstrate. The approximation in expression (10) is ignoring some spatial correlation in the seismic properties; hence it is expected to have somewhat higher variability than the correct one. Each realization appear as more heterogeneous than the predicted LF profile as it should do, but it also look more heterogeneous than the reference LF profile. This is as expected from the discussion above.

Figure 7 contains results from the LF prediction based on the reference seismic signal $\mathbf{s}^R$; hence based on the approximate posterior pdf $\tilde{p}(\boldsymbol{\pi}|\mathbf{s}^R)$. This figure has layout identical to Figure 3. The results expose the reliability of the approximate inversion approach since the effect of the seismic noise is removed. Note, however, that the colored variability caused by heterogeneity within the LF classes is still included in the test. The prediction is almost perfect, and deviates from the reference profile $\boldsymbol{\pi}^R$ only in the top zone and a thin zone around 2360 ms. By simulating LF profiles based on this model one can demonstrate that the approximation of the posterior pdf represents almost no uncertainty.

Figure 8 contains results from the LF prediction based on the seismic data $\mathbf{d}^R$ but with a prior model without spatial coupling, see expression (12). The results expose the influence of the spatially coupled Markov chain prior model. The predicted LF profile appears as highly variable and unreliable caused by the white noise in the seismic data. Note in particular that oil-saturated sand is predicted nowhere. The results in Figure 8 demonstrates the importance of using a spatially coupled prior model. The imposed continuity of the LF profile strongly increases the robustness of the inversion.

SUMMARY AND CONCLUSIONS

In the Lithology-Fluid (LF) prediction and simulation the objective is to determine the LF classes in a profile through a target zone based on prestack seismic data. A 1D Bayesian inversion model along the profile is introduced and a stationary Markov chain prior model is used to model the vertical coupling along the profile. The likelihood model may be decomposed into a rock physics likelihood model, and a seismic likelihood model. The rock physics model is heterogenous and non-linear and is represented by a set of samples, and the seismic model is based on the convolution model with linearized reflection coeffisients, which entails strong spatial coupling in the likelihood model. An approximate likelihood model provides an approximate posterior model which is assessed by a fast recursive sampling algorithm.

The LF prediction and simulation approach is evaluated on a reference case inspired by a North Sea sand stone reservoir, which contains four LF classes, gas-, oil-, and brine-saturated sand stone, and shale. The results of the LF predictions are surprisingly reliable for a signal-to-noise ratio of 2.3 in the seismic data. All hydrocarbon bearing zones are identified, although with some misclassifications between gas and oil. The uncertainty in the LF prediction increases if the number of zones increases and the zone thicknesses decrease. By representing the LF predictions by a set of LF

profile simulations, the approximate posterior model appears to provide realizations that are somewhat too heterogenuous.

The approximate likelihood model appears as remarkable reliable. The results also show that the spatial coupling in the prior model is very important. Without coupling the predictions are highly influenced by seismic noise.

The upward-downward recursive algorithm is exact and very efficient. The time consuming part is the calculation of the three dimensional integral in the approximate likelihood expression, since that is done numerically based on a relatively large set of sample values. Given the approximate likelihood values, the inversion can be done in less than a second on an average sized computer for the case used in this study.

ACKNOWLEDGEMENTS

The research is funded by the URE activity at NTNU and a Statoil grant.

REFERENCES

Aki, K., and Richards, P. G., 1980, Quantitative seismology: W.H.Freemann & co.

Buland, A., Kolbjornsen, O., and Omre, H., 2003, Rapid spatially coupled AVO inversion in the Fourier domain: Geophysics, **68**, 824-836.

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

Eidsvik, J., Avseth, P., Omre, H., Mukerji, T., and Mavko, G., 2004, Stochastic reservoir characterization using prestack seismic data: Geophysics, **69**, 978-993.

Eidsvik, J., Mukerji, T., and Switzer, P., 2004, Estimation of geological attributes from well log: An application of hidden Markov chains: Mathematical Geology, **36**, 379-397.

Houck, R. T., 2002, Quantifying the uncertainty in an AVO interpretation: Geophysics, **67**, 117-125.

Krumbein, Y. C., and Dacey, M. F., 1969, Markov chains and embedded Markov chains in geology: Mathematical Geology, **1**, 79-96.

Mukerji, T., Jorstad, A., Avseth, P., Mavko, G., and Granli, J. R., 2001, Mapping lithofacies and pore-fluid probabilities in a North Sea reservoir: Seismic inversions and statistical rock physics: Geophysics, **66**, 988-1001.

Mavko, G., Mukerji, T., and Dvorkin, J., 1988, The rock physics handbook: Cambridge University Press.

Scott, S. L., 2002, Bayesian methods for hidden Markov models: Recursive computing for the 21th century: Journal of American Statistical Association, **97**, 337-351.

APPENDIX: UPWARD-DOWNWARD RECURSIVE ALGORITHM

The pdf of interest is

$$\begin{aligned}\tilde{p}(\boldsymbol{\pi}|\mathbf{d}) &= \text{const} \times \prod_t \tilde{p}(\mathbf{d}|\pi_t)\, p(\pi_t|\pi_{t-1}), \\ &= \prod_t \tilde{p}(\pi_t|\pi_{t-1}, \mathbf{d}),\end{aligned} \tag{A-1}$$

with $p(\pi_1) = p(\pi_1|\pi_0)$ and $\tilde{p}(\pi_1|\pi_0, \mathbf{d}) = \tilde{p}(\pi_1|\mathbf{d})$ for notational convenience. This pdf can be shown to be a non-stationary Markov chain.

Algorithm

Upward recursion

Initiate

$$p^u(\pi_T) = \text{const} \times \tilde{p}(\mathbf{d}|\pi_T)\, p(\pi_T)$$

$$\text{const from} \quad \textstyle\sum_{\pi_T} p^u(\pi_T) = 1$$

Iterate $t = T-1, \ldots, 1$

$$p^u(\pi_{t+1}, \pi_t) = \text{const} \times \tilde{p}(\mathbf{d}|\pi_t)\, p(\pi_t|\pi_{t+1})\, p^u(\pi_{t+1})$$

$$\text{const from} \quad \textstyle\sum_{\pi_{t+1}} \sum_{\pi_t} p^u(\pi_{t+1}, \pi_t) = 1$$

$$p^u(\pi_t) = \textstyle\sum_{\pi_{t+1}} p^u(\pi_{t+1}, \pi_t)$$

end iterate

Downward recursion

Initiate

$$p^d(\pi_1) = p^u(\pi_1)$$

generate π_1^s from $p^d(\pi_1)$

Iterate $t = 2, \ldots, T$

$$p^d(\pi_t|\pi_{t-1}) = \frac{p^u(\pi_{t-1},\pi_t)}{p^u(\pi_{t-1})}$$

$$p^d(\pi_t) = \sum_{\pi_{t-1}} p^d(\pi_t|\pi_{t-1})\, p^d(\pi_{t-1})$$

generate π_t^s from $p^d(\pi_t|\pi_{t-1}^s)$

end iterate

End Algorithm

It can be shown, see Scott (2002), that

$$\tilde{p}(\pi_1|\mathbf{d}) = p^d(\pi_1)$$
$$\tilde{p}(\pi_t|\pi_{t-1}, \mathbf{d}) = p^d(\pi_t|\pi_{t-1}) \text{ for all t}$$

hence $\tilde{p}(\pi_t|\mathbf{d}) = p^d(\pi_t)$ for all t and $\boldsymbol{\pi}^s = (\pi_1^s, \ldots, \pi_T^s)$ is a realization from $\tilde{p}(\boldsymbol{\pi}|\mathbf{d})$.

FIGURES

FIG. 1. Reference LF profile $\boldsymbol{\pi}^R$ with gas-saturated sand stone (white), oil-saturated sand stone (light gray), brine-saturated sand stone (dark gray), and shale (black); elastic properties profile $\boldsymbol{m}^R$ represented by P-wave velocity, S-wave velocity, and density; seismic signal profile $\boldsymbol{s}^R$; and seismic data profile $\boldsymbol{d}^R$.

FIG. 2. Elastic properties $\boldsymbol{m}$ represented by P-wave velocity, S-wave velocity, and density given gas-saturated sand stone ($\cdot$), oil-saturated sand stone ($\circ$), brine-saturated sand stone ($+$), and shale ($\times$) simulated from a rock physics model.

FIG. 3. Location wise Map LF profile prediction $\tilde{\boldsymbol{\pi}}$, with gas-saturated sand stone (white), oil-saturated sand stone (light gray), brine-saturated sand stone (dark gray), and shale (black). Reference LF profile $\boldsymbol{\pi}^R$; reference seismic data $\boldsymbol{d}^R$; and classification matrix of $\boldsymbol{\pi}^R$ versus $\tilde{\boldsymbol{\pi}}$ where gas-, oil-, and brine-saturated sand stone, and shale are denoted SG, SO, SB, and SH respectively.

FIG. 4. Profile of marginal approximate posterior pdf $\tilde{p}(\pi_t|\boldsymbol{d}^R)$ for gas-, oil-, and brine-saturated sand stone, and shale with reference LF profile $\boldsymbol{\pi}^R$ marked on respective axis.

FIG. 5. Seismic signal computed from predicted LF profile without inter-heterogenity in seismic properties; reference seismic signal $\boldsymbol{s}^R$; and reference seismic data $\boldsymbol{d}^R$.

FIG. 6. Independent realizations of LF profiles $\boldsymbol{\pi}^S$ from approximate posterior pdf $\tilde{p}(\boldsymbol{\pi}|\boldsymbol{d}^R)$; and reference LF profile $\boldsymbol{\pi}^R$.

FIG. 7. Results from case without observation error, i.e. inversion based on reference seismic signal $\boldsymbol{s}^R$. Location wise Map LF profile prediction $\tilde{\boldsymbol{\pi}}$, with gas-saturated

sand stone (white), oil-saturated sand stone (light gray), brine-saturated sand stone (dark gray), and shale (black). Reference LF profile $\boldsymbol{\pi}^R$; reference seismic signal $\boldsymbol{s}^R$; and classification matrix of $\boldsymbol{\pi}^R$ versus $\tilde{\boldsymbol{\pi}}$ where gas-, oil-, and brine-saturated sand stone, and shale are denoted SG, SO, SB, and SH respectively.

FIG. 8. Results from case with prior model without spatial coupling. Location wise Map LF profile prediction $\tilde{\boldsymbol{\pi}}$, with gas-saturated sand stone (white), oil-saturated sand stone (light gray), brine-saturated sand stone (dark gray), and shale (black). Reference LF profile $\boldsymbol{\pi}^R$; reference seismic data $\boldsymbol{d}^R$; and classification matrix of $\boldsymbol{\pi}^R$ versus $\tilde{\boldsymbol{\pi}}$ where gas-, oil-, and brine-saturated sand stone, and shale are denoted SG, SO, SB, and SH respectively.

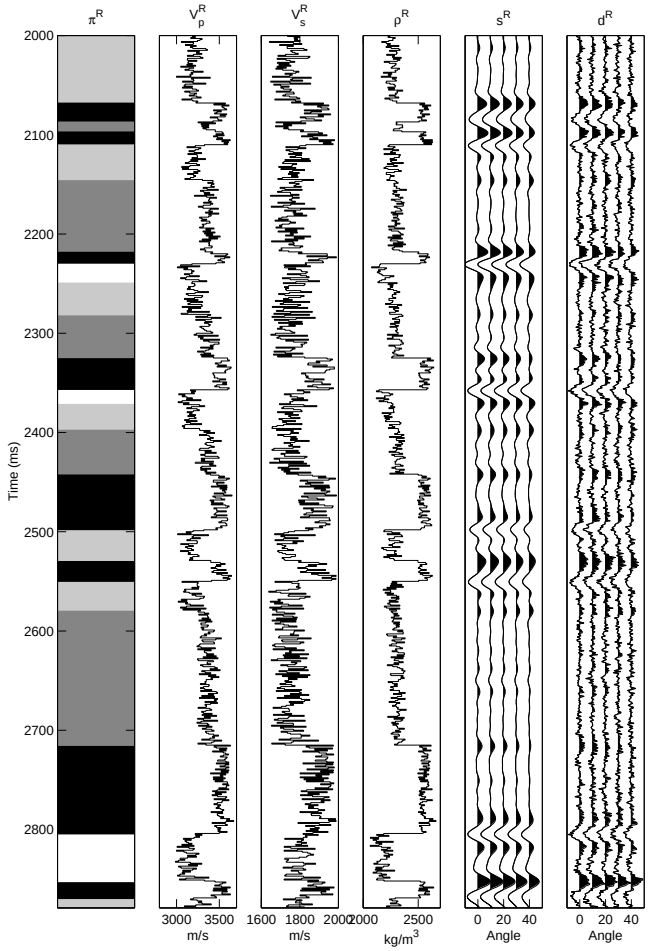


FIG. 1. Reference LF profile $\boldsymbol{\pi}^R$ with gas-saturated sand stone (white), oil-saturated sand stone (light gray), brine-saturated sand stone (dark gray), and shale (black); elastic properties profile $\boldsymbol{m}^R$ represented by P-wave velocity, S-wave velocity, and density; seismic signal profile $\boldsymbol{s}^R$; and seismic data profile $\boldsymbol{d}^R$.

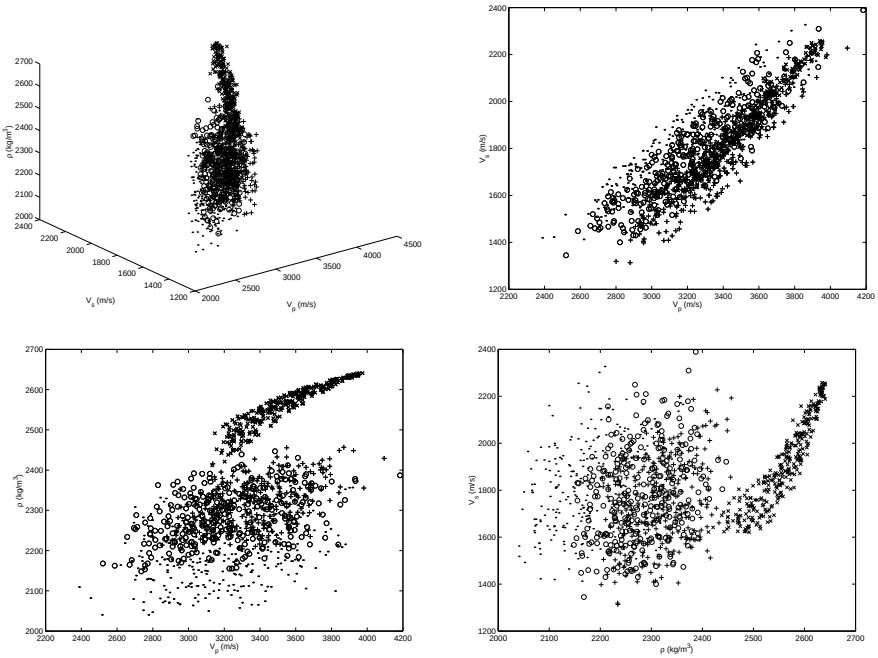
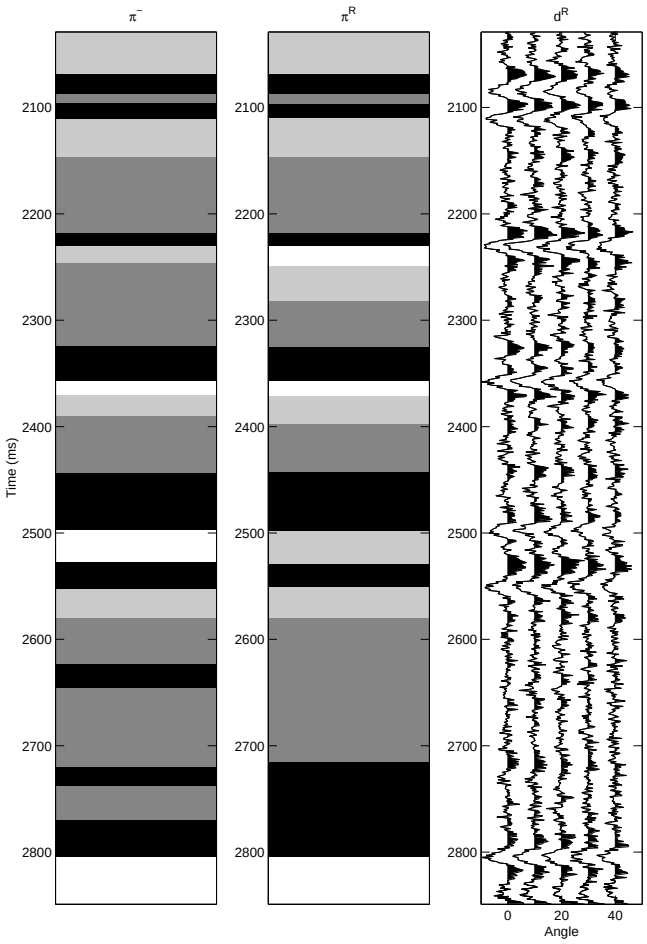


FIG. 2. Elastic properties $\boldsymbol{m}$ represented by P-wave velocity, S-wave velocity, and density given gas-saturated sand stone ($\cdot$), oil-saturated sand stone ($\circ$), brine-saturated sand stone ($+$), and shale ($\times$) simulated from a rock physics model.



$\boldsymbol{\pi}^R \backslash \tilde{\boldsymbol{\pi}}$	SG	SO	SB	SH	$\sum$
SG	58	17	3	0	78
SO	29	121	40	5	195
SB	0	0	282	24	306
SH	1	0	37	204	242
$\sum$	88	138	362	233	

FIG. 3. Location wise Map LF profile prediction $\tilde{\boldsymbol{\pi}}$, with gas-saturated sand stone (white), oil-saturated sand stone (light gray), brine-saturated sand stone (dark gray), and shale (black). Reference LF profile $\boldsymbol{\pi}^R$; reference seismic data $\boldsymbol{d}^R$; and classification matrix of $\boldsymbol{\pi}^R$ versus $\tilde{\boldsymbol{\pi}}$ where gas-, oil-, and brine-saturated sand stone, and shale are denoted SG, SO, SB, and SH respectively.

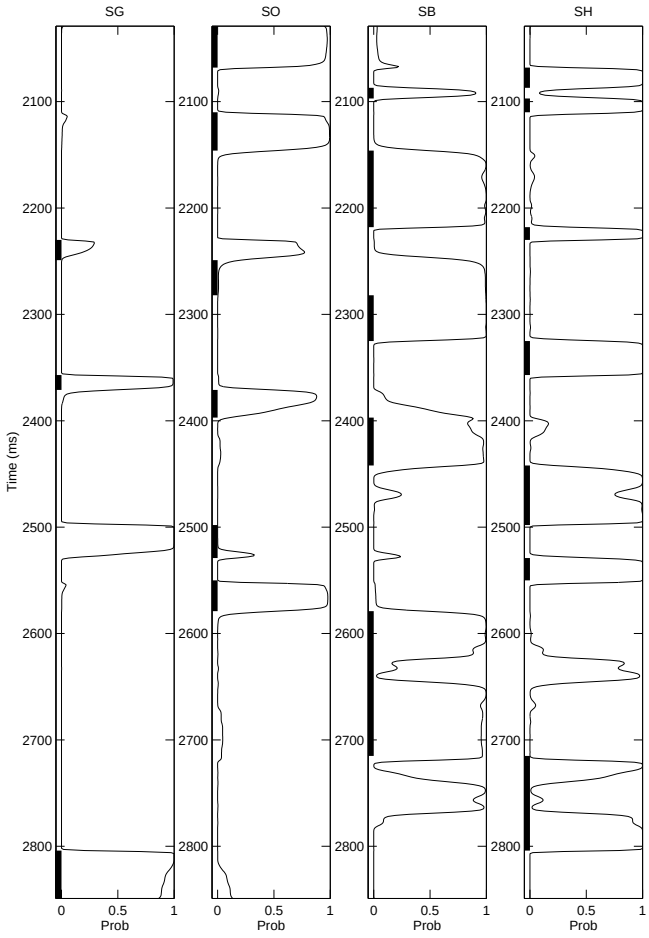


FIG. 4. Profile of marginal approximate posterior pdf $\tilde{p}(\pi_t|\boldsymbol{d}^R)$ for gas-, oil-, and brine-saturated sand stone, and shale with reference LF profile $\boldsymbol{\pi}^R$ marked on respective axis.

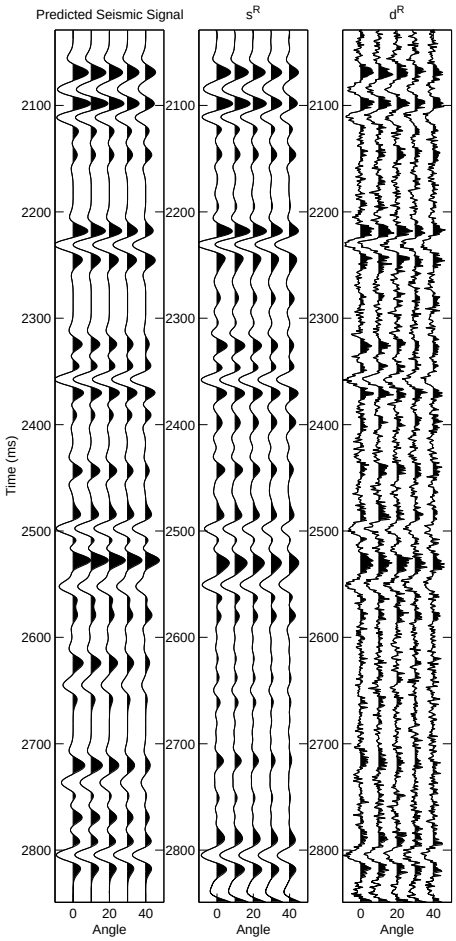


FIG. 5. Seismic signal computed from predicted LF profile without inter-heterogenity in seismic properties; reference seismic signal $\boldsymbol{s}^R$; and reference seismic data $\boldsymbol{d}^R$.

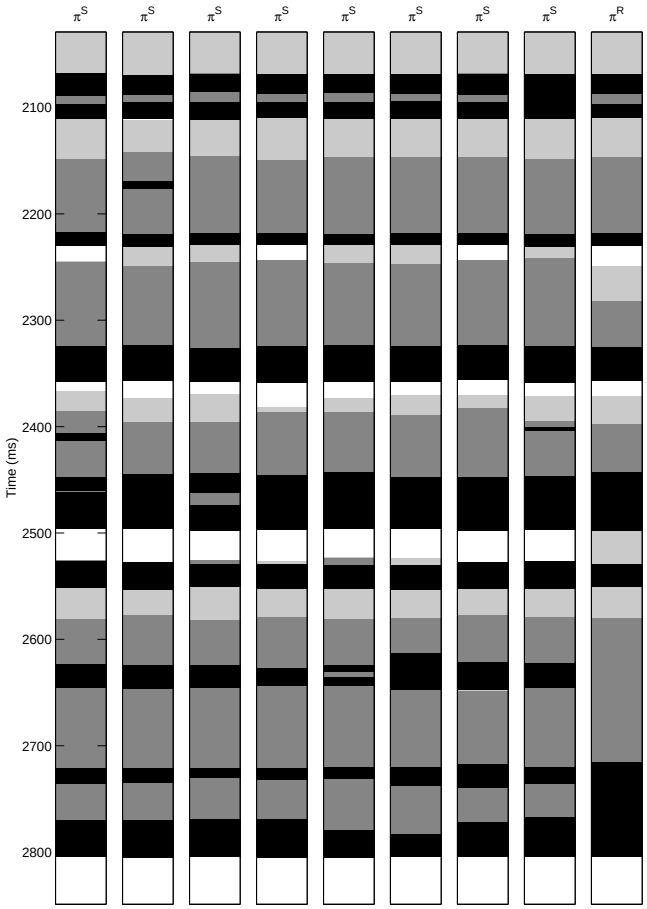
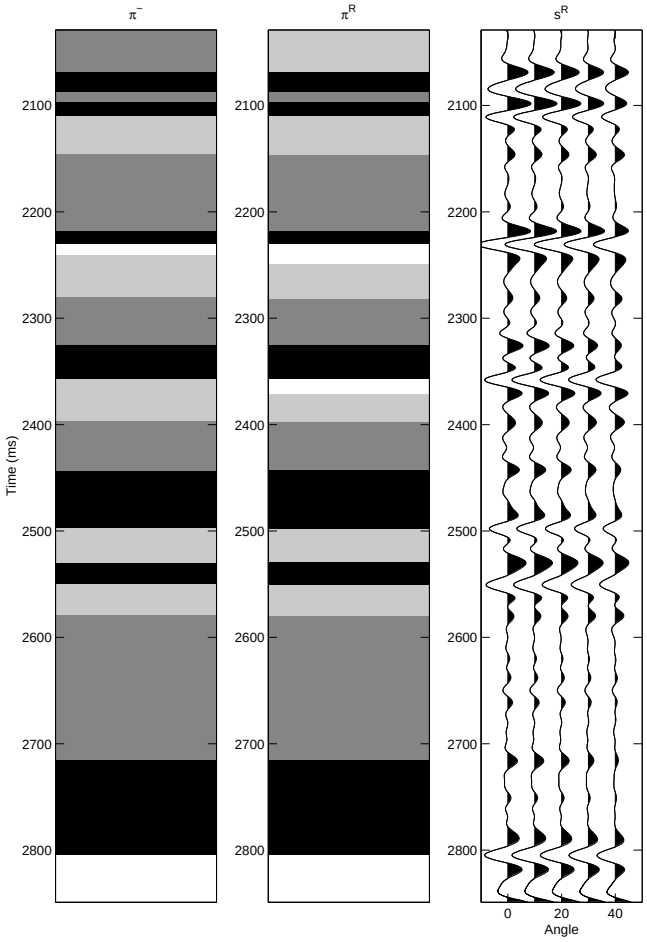
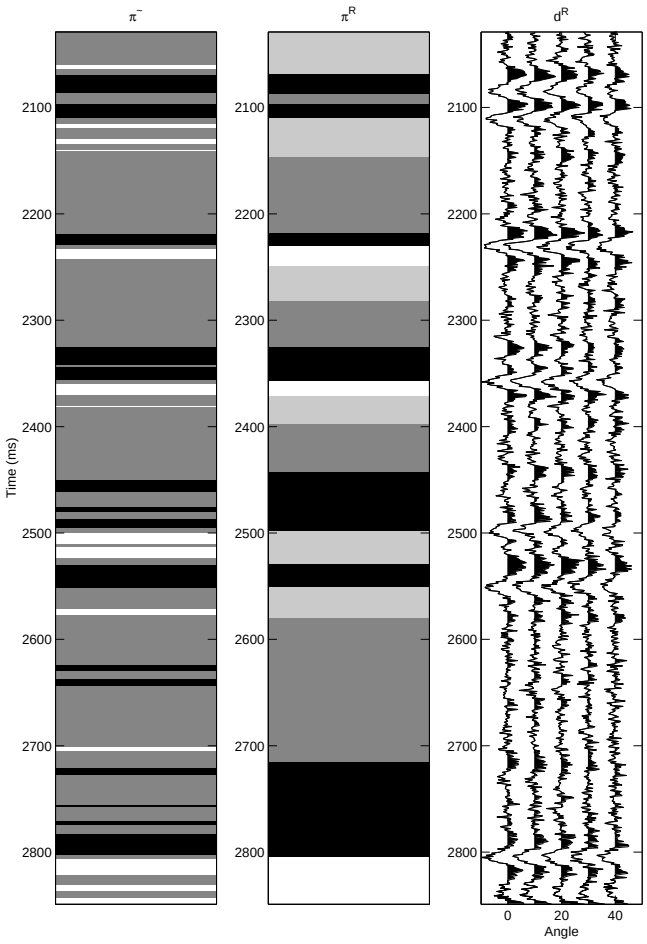


FIG. 6. Independent realizations of LF profiles $\boldsymbol{\pi}^S$ from approximate posterior pdf $\tilde{p}(\boldsymbol{\pi}|\boldsymbol{d}^R)$; and reference LF profile $\boldsymbol{\pi}^R$.



$\boldsymbol{\pi}^R \backslash \tilde{\boldsymbol{\pi}}$	SG	SO	SB	SH	$\sum$
SG	55	23	0	0	78
SO	0	150	45	0	195
SB	0	0	306	0	306
SH	0	3	1	238	242
$\sum$	55	176	352	238	

FIG. 7. Results from case without observation error, i.e. inversion based on reference seismic signal $\boldsymbol{s}^R$. Location wise Map LF profile prediction $\tilde{\boldsymbol{\pi}}$, with gas-saturated sand stone (white), oil-saturated sand stone (light gray), brine-saturated sand stone (dark gray), and shale (black). Reference LF profile $\boldsymbol{\pi}^R$; reference seismic signal $\boldsymbol{s}^R$; and classification matrix of $\boldsymbol{\pi}^R$ versus $\tilde{\boldsymbol{\pi}}$ where gas-, oil-, and brine-saturated sand stone, and shale are denoted SG, SO, SB, and SH respectively.



$\boldsymbol{\pi}^R \backslash \tilde{\boldsymbol{\pi}}$	SG	SO	SB	SH	$\sum$
SG	47	0	31	0	78
SO	41	0	153	1	195
SB	3	0	292	11	306
SH	0	0	95	147	242
$\sum$	91	0	571	159	

FIG. 8. Results from case with prior model without spatial coupling. Location wise Map LF profile prediction $\tilde{\boldsymbol{\pi}}$, with gas-saturated sand stone (white), oil-saturated sand stone (light gray), brine-saturated sand stone (dark gray), and shale (black). Reference LF profile $\boldsymbol{\pi}^R$; reference seismic data $\boldsymbol{d}^R$; and classification matrix of $\boldsymbol{\pi}^R$ versus $\tilde{\boldsymbol{\pi}}$ where gas-, oil-, and brine-saturated sand stone, and shale are denoted SG, SO, SB, and SH respectively.