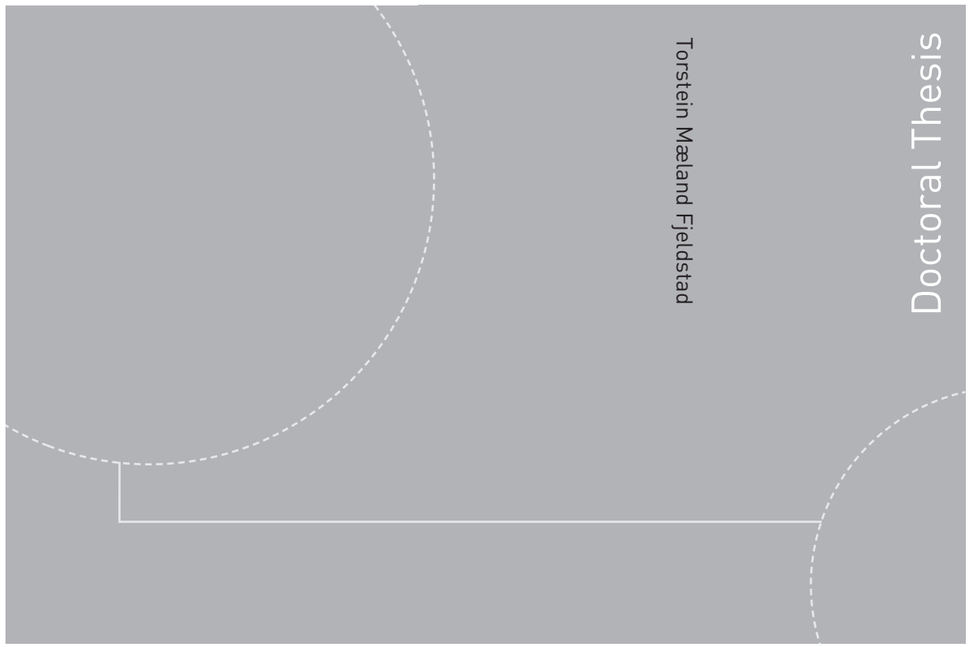
ISBN 978-82-326-4710-1 (printed version)
ISBN 978-82-326-4711-8 (electronic version)
ISSN 1503-8181



Doctoral theses at NTNU, 2020:181

Torstein Mæland Fjeldstad

Spatial Gaussian Mixture Models Applied to Bayesian Seismic Inversion

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of
Science and Technology
Faculty of Information Technology
and Electrical Engineering
Department of Mathematical Sciences

Doctoral theses at NTNU, 2020:181

NTNU

Torstein Mæland Fjeldstad

Spatial Gaussian Mixture Models Applied to Bayesian Seismic Inversion

Thesis for the degree of Philosophiae Doctor

Trondheim, May 2020

Norwegian University of Science and Technology
Faculty of Information Technology
and Electrical Engineering
Department of Mathematical Sciences

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the degree of Philosophiae Doctor

Faculty of Information Technology
and Electrical Engineering
Department of Mathematical Sciences

ISBN 978-82-326-4710-1 (printed version)
ISBN 978-82-326-4711-8 (electronic version)
ISSN 1503-8181

Doctoral theses at NTNU, 2020:181

Printed by Skipnes Kommunikasjon as

Preface

This thesis is submitted in partial fulfillment of the requirements for the degree of Philiosophiae Doctor (PhD) at the Norwegian University of Science and Technology (NTNU). The research is funded by the Uncertainty of Reservoir Evaluation (URE) initiative at the Department of Mathematical Sciences (IMF), NTNU.

First and foremost, I would like to thank my supervisor Professor Henning Omre for his excellent supervision and guidance during my time as a PhD student at IMF. Secondly, I would like to thank my co-supervisor Adjunct Professor Per Åge Avseth for valuable input during our collaboration and for providing the data. In addition, I am grateful for the collaboration with my co-authors Associate Professor Dario Grana, Professor Håkon Tjelmeland and Xin Luo, PhD. I am also grateful to Professor Jo Eidsvik for valuable discussions. I would also like to thank the statistics group at IMF for giving me the opportunity of lecturing the introductory course in statistics at NTNU.

In addition I would like to thank my friends and colleagues at IMF for creating a socially stimulating work environment. Finally, I am grateful to my family for their continuous support during my period as a PhD student.

Torstein Mæland Fjeldstad

Trondheim, February 2020

Thesis outline

Background

Paper I **Bayesian Gaussian Mixture Linear Inversion for Geophysical Inverse Problems**

Dario Grana, Torstein Fjeldstad and Henning Omre

Mathematical Geosciences vol. 49 (4). (2017)

Paper II **Bayesian Inversion of Convolved Hidden Markov Models With Applications in Reservoir Prediction**

Torstein Fjeldstad and Henning Omre

IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 3, pp. 1957-1968, March 2020

Paper III **Joint probabilistic petrophysics-seismic inversion based on Gaussian mixture and Markov chain prior models**

Torstein Fjeldstad and Dario Grana

Geophysics vol. 83 (1) (2018)

Paper IV **Bayesian model for lithology/fluid class prediction using a Markov mesh prior fitted from a training image**

Håkon Tjelmeland, Xin Luo and Torstein Fjeldstad

Geophysical Prospecting vol. 67 (3). (2019)

Paper V **A one-step Bayesian inversion framework for three-dimensional reservoir characterization based on a Gaussian mixture model – A Norwegian Sea demonstration**

Torstein Fjeldstad, Per Åge Avseth and Henning Omre

Submitted

Background

Introduction

Spatial models are becoming increasingly important in a wide variety of applications, such as epidemiology (Lawson, 2018), image processing (Hurn et al., 2003) and geoscience (Tarantola, 2005; Hansen et al., 2006). The objective is to predict a variable of interest in a spatial domain where exact observations of the variable of interest are not available. The observations are not necessarily made as a set of univariate independent realizations of the underlying variable but as weighted aggregates across space in addition to a measurement error. The class of problems to be studied has various names depending on the field of application, including inverse models (Tarantola, 2005) and switching state-space models (Frühwirth-Schnatter, 2006). A probabilistic approach appears to be well suited to describe the variable of interest, as it provides not only a point prediction but also the full probabilistic specification for the variable of interest.

Formally, we operate in a Bayesian inversion framework, where the ultimate objective is to assess the posterior probability density of the spatial variable given a set of measurements. The variable of interest is assigned a prior probability density, and a likelihood function relates the variable to the observations. According to Bayes' theorem, the posterior density is proportional to the product of the likelihood function and the prior density. The prior and likelihood models are defined from a set of model parameters based on parametric classes of models, which we, in this thesis, consider to

be fixed and known.

We focus on the assessment of spatial variables that arise from the characterization of a three-dimensional oil and gas reservoir. During the exploration and development phase of a reservoir, measurements of the subsurface are combined with geophysical and geological knowledge about the area of interest to predict the occurrence of hydrocarbons, such as oil and gas, in the subsurface. In the exploration phase, well observations are typically sparsely distributed in space, if they exist, and seismic measurements are heavily blurred images of the subsurface in the vertical direction. The latter is a result of the convolution of the reflection coefficients that occurs when the seismic waves propagate through the subsurface. The likelihood function is defined from a set of geophysical relations describing wave propagation in the subsurface, relating the model variables to the observations. We consider a so-called Gauss-linear likelihood model, where the measurements are assumed to be a linear translation of the underlying variable with a Gaussian error term.

Ideally, in a Bayesian framework, we seek an analytic expression for the posterior density of interest. However, such expressions are generally not available in closed form in reservoir characterization. If the likelihood model is Gauss-linear *and* the prior model is a Gaussian density, the posterior is Gaussian with analytic expressions for the mean vector and covariance matrix. This is because the Gaussian density is fully specified from the first two moments and is closed under linear combinations, marginalization and

conditioning. We refer to, for example, Cressie (1993) for an introduction to Gaussian random fields. A popular choice for interpolation of the spatial variable is kriging (Matheron, 1963). However, in high-dimensional problems, kriging is not necessarily feasible due to storage and memory limitations. Alternative approaches include techniques based on the Fourier transformation (Buland et al., 2003) or Gaussian Markov random field approximations (Rue and Held, 2005).

In geostatistical applications, the assumption of a Gauss-linear likelihood model discussed above is often adequate. The Gaussian prior model assumption, however, does not need to be satisfactory to represent prior knowledge of the variable of interest. In this thesis, we discuss a class of prior models that have marginal densities that are multimodal and skewed, namely, Gaussian mixture densities. Alternative models not discussed here include copula models (Bárdossy and Li, 2008), generalized linear modeling approaches (Diggle et al., 1998), uniform transformation models (De Oliveira et al., 1997) and selection Gaussian random fields (Omre and Rimstad, 2018)

The Gaussian mixture density is defined by a discrete latent variable and a set of conditional Gaussian densities and is a specific case of a finite mixture model (Titterington et al., 1985). In principle, any density can be approximated with a desired accuracy using a finite mixture model with the appropriate number of mixing components (Ferguson, 1973). Depending on the application, the latent variable may have an intuitive physical interpretation or may be a pure nuisance variable. In both cases, a mixture model

is well suited for classification of the underlying variable. One alternative is to model the latent discrete variable in the Gaussian mixture model by using a Markov random field model (Besag, 1974; Kindermann et al., 1980; Tjelmeland and Besag, 1998; Hurn et al., 2003). Other alternatives, such as pluri-Gaussian random fields (Matheron et al., 1987), the Markov mesh model (Abend et al., 1965) and multiple-point statistics (Caers, 2001; Strebelle, 2002; Ortiz and Deutsch, 2004), exist.

The continuous-valued variables of interest in reservoir characterization, i.e., elastic attributes such as pressure wave velocity and petrophysical properties such as water saturation, often appear as multimodal or skewed due to the presence of various lithology or fluid classes of the subsurface (Grana and Della Rossa, 2010; Rimstad et al., 2012). A Gaussian mixture prior density appears to be well suited to model this behavior.

The prior and likelihood models are defined from a set of model parameters that have to be chosen. If the latent discrete random field is conditionally independent, parameter estimates of the mixing proportions are typically obtained by the expectation-maximization algorithm (Dempster et al., 1977) or by clustering methods (Hastie et al., 2009). For a one-dimensional convolved hidden Markov model, Lindberg and Omre (2015) assessed the vertical transition matrix. Two- and three-dimensional parameter inference for the discrete Markov random field model with an arbitrary neighborhood system is in practice restricted by the computationally intractable normalizing constant. In contrast, the normalizing constant for the Markov mesh

model is computationally tractable, which makes a Bayesian approach feasible (Luo and Tjelmeland, 2017). Bolin et al. (2019) considered parameter estimation for a latent Gaussian mixture density using a likelihood-based method. Rimstad and Omre (2010), Lindberg and Omre (2014) and Skauvold et al. (2016) considered estimation of the likelihood model parameters with applications in reservoir characterization.

The ultimate objective is to assess the posterior model for the variables of interest given the observations. Numerical methods based on sampling or optimization of the posterior are often applied to assess the posterior density of interest. We focus on Markov chain Monte Carlo methods to assess the posterior density of interest. See, e.g., Gilks et al. (1995) for an introduction to Markov chain Monte Carlo methods. These methods are often tailored for a specific problem, and they often require manual tuning to obtain a satisfactory convergence rate.

We have outlined some of the challenges of reservoir characterization, such as the choice of the likelihood and prior model. The objective of this thesis is to construct a class of computationally feasible models in a Bayesian spatial inversion framework for the joint characterization of lithology and fluid classes, petrophysical properties and elastic attributes in three-dimensional reservoir characterization. This thesis mainly discusses the choice of the prior model and efficient sampling methods to assess the posterior density for the spatial variables of interest. In the following, we describe the Bayesian inversion framework and introduce the necessary model

variables and parameters. The introduction serves as a motivation and overview of the thesis and outlines the class of models discussed in this thesis.

Notation

Unless otherwise specified, the variables to be discussed are defined on a discretized volume $\mathcal{L} = \{xyt : x = 1, \ldots, n_x, y = 1, \ldots, n_y, t = 1, \ldots, n_t\}$. Let $v \in \mathcal{L}$ denote a grid cell and let $n = n_x n_y n_t$. We consider a discretized spatial variable $\{z_v : v \in \mathcal{L}\}$ represented by the vector $\mathbf{z}$, where each $z_v \in \Omega_z$, and denote $\mathbf{z} \setminus z_v$ for all $v \in \mathcal{L}$ by $\mathbf{z}_{-v}$. We denote the collection of grid cells along a vertical profile at horizontal position xy by $\mathcal{L}_{xy\cdot}$. Furthermore, let $p(\cdot)$ denote an arbitrary probability density/mass function.

Inverse problems

In an inverse framework, the objective is to characterize an unobservable variable $\mathbf{z}$ given a set of measurements $\mathbf{y}$ that are related to $\mathbf{z}$ (Tarantola, 2005; Hansen et al., 2006). The measurements $\mathbf{y}$ in general represent a nonlinear transformation of the underlying variable $\mathbf{z}$, which we cannot observe directly, with an associated measurement error. In compact vector notation, the forward model is often expressed as

$$\big[\mathbf{y} \mid \mathbf{z}\big] = \mathcal{F}(\mathbf{z}) + \boldsymbol{\varepsilon}, \tag{1}$$

where $\mathcal{F}(\cdot)$ is the forward operator and $\boldsymbol{\varepsilon}$ is assumed to be a Gaussian error term. If $\mathcal{F}(\cdot)$ is a linear operator, we refer to the model as a Gauss-linear model, since Equation (1) is linear in the conditioning variable and includes a Gaussian error term. The forward operator may represent a physical model, for example, the wave equation. The unobserved variable $\mathbf{z}$ is often assessed based on the observations $\mathbf{y}$ by optimizing a misfit criterion including a regularizing factor or by using a Bayesian inversion framework where $\mathbf{z}$ is assigned a prior density.

Inverse problems may be formulated as optimization problems, for example, minimization of the squared error. These inverse problems are often ill posed, and the objective function needs to be regularized by adding a penalty term

$$\tilde{\mathbf{z}} = \underset{\mathbf{z}}{\operatorname{argmin}} \|\mathbf{y} - \mathcal{F}(\mathbf{z})\| + \lambda \|\mathbf{z} - \mathbf{z}_0\|. \tag{2}$$

Here, $\mathbf{z}_0$ is a prior guess on a reasonable solution, and the scalar λ is a regularization factor. The most common norm in Equation (2) is the L^2-norm; however, other norms may also be considered. The objective function in Equation (2) is often assessed by iterative methods such as steepest descent or conjugate gradient methods (Nocedal and Wright, 2006), genetic algorithms (Goldberg, 1989; Sen and Stoffa, 1996) or simulated annealing (Kirkpatrick et al., 1983). Note that only a point prediction $\tilde{\mathbf{z}}$ is obtained.

We focus on the Bayesian framework, since it provides the full posterior probability density function, not only a point prediction $\tilde{\mathbf{z}}$. In a Bayesian

framework, the objective is to assess the posterior density

$$p\big(\mathbf{z} \mid \mathbf{y}\big) \propto p\big(\mathbf{y} \mid \mathbf{z}\big)p(\mathbf{z}), \tag{3}$$

where $p\big(\mathbf{y} \mid \mathbf{z}\big)$ is the likelihood function and $p(\mathbf{z})$ is the prior density. Note that the prior $p(\mathbf{z})$ regularizes the solution. Assessment of the posterior density is computationally challenging and often requires approximations and/or simulation-based techniques. As our problem is formulated in a Bayesian inversion setting where the main focus is on high-dimensional variable prediction and not low-dimensional model parameter inference, assigning an informative prior may be a necessity to obtain the solution of the inverse problem. Such prior knowledge may be available from geological experience based on comparable areas and/or physical models.

For a Gauss-linear likelihood function and a Gaussian prior, an analytic expression for the posterior density is available. The choice of a Gauss-linear likelihood function and a Gaussian prior model is motivated by the fact that the Gaussian density is closed under linear transformations, marginalization and conditioning. In geostatistical applications, the Gauss-linear likelihood assumption is often adequate; however, prior knowledge of the underlying random field $\mathbf{z}$ need not be adequately described by a Gaussian prior density due to multimodality and skewness. Predictions based on the Gaussian prior tend to be too smooth and cause a regression towards the global mean.

We focus on a specific class of prior models that satisfies the multimodality and skewness properties, namely, the Gaussian mixture density.

We introduce a discrete-valued latent variable $\mathbf{s}$ associated with $\mathbf{z}$. The latent variable $\mathbf{s}$, where each $s_v \in \Omega_s = \{1, \ldots, L\}$ for $v \in \mathcal{L}$, acts as a mode indicator or discrete switching state space variable (Frühwirth-Schnatter, 2006) for $\big[\mathbf{z} \mid \mathbf{s}\big]$. Note that the model parameters describing the conditional density $p\big(\mathbf{z} \mid \mathbf{s}\big)$ depend on the corresponding value of $\mathbf{s}$. One of the most common models is the Gaussian mixture density

$$p(\mathbf{z}) = \sum_{\mathbf{s} \in \Omega_s^n} p\big(\mathbf{z} \mid \mathbf{s}\big) p(\mathbf{s}), \tag{4}$$

where $p\big(\mathbf{z} \mid \mathbf{s}\big)$ is a Gaussian density with mean and covariance given by the corresponding value of $\mathbf{s}$. Traditionally, the Gaussian mixture prior density is assumed not to have any spatial structure of $\mathbf{s}$, e.g., conditional independence $p(\mathbf{s}) = \prod_{i=1}^{n} p(s_i)$. We relax this assumption and consider a spatial model $p(\mathbf{s})$ to impose spatial structure.

Markov chain Monte Carlo simulation

The posterior density of interest defined in Equation (3) is often not analytically tractable, and stochastic simulation methods have to be applied. Simulation-based assessment is often performed by constructing a Markov chain with a stationary or equilibrium distribution equal to the target density $p\big(\mathbf{z} \mid \mathbf{y}\big)$. A popular way to choose such a Markov chain is by the use of the Markov chain Monte Carlo Metropolis-Hastings algorithm. We refer to Gilks et al. (1995) for an introduction to simulation techniques. The

following description defines the algorithm.

First, initialize $\mathbf{z}^1$ with $p\big(\mathbf{z}^1 \mid \mathbf{y}\big) > 0$. Then, for $i = 2$ and until convergence, the following two steps are carried out at each iteration i:

Proposal step : Propose $\mathbf{z}$ according to the proposal density $q(\cdot)$

$$\mathbf{z} \sim q(\mathbf{z} \mid \mathbf{z}^{i-1}). \tag{5}$$

Accept/reject step : Set $\mathbf{z}^i = \mathbf{z}$ with probability

$$\alpha = \min\left\{1, \frac{p\big(\mathbf{y} \mid \mathbf{z}\big)p(\mathbf{z})}{p\big(\mathbf{y} \mid \mathbf{z}^{i-1}\big)p\big(\mathbf{z}^{i-1}\big)} \times \frac{q(\mathbf{z}^{i-1} \mid \mathbf{z})}{q(\mathbf{z} \mid \mathbf{z}^{i-1})}\right\} \tag{6}$$

and $\mathbf{z}^i = \mathbf{z}^{i-1}$ else.

After convergence and thinning, an ensemble of approximately independent realizations $\mathbf{z}^1, \ldots, \mathbf{z}^B$ is generated from the posterior density $p\big(\mathbf{z} \mid \mathbf{y}\big)$. Based on these realizations, a set of summary statistics is computed. For a continuous-valued $\mathbf{z}$, the most common choices of summary statistics include the expected value and the standard deviation. However, since we are interested in multimodal and skewed densities, we also consider the marginal maximum posterior (MMAP) predictor,

$$\widehat{\mathbf{z}} = \left\{\widehat{z}_v = \underset{z_v \in \Omega_z}{\operatorname{argmax}}\, p\big(z_v \mid \mathbf{y}\big); v \in \mathcal{L}\right\}, \tag{7}$$

which is often referred to as the (marginal) mode. For a discrete-valued $\mathbf{z}$, we consider the MMAP predictor

$$\widehat{\mathbf{p}} = \left\{\widehat{p}_v = \underset{l \in \Omega_z}{\operatorname{argmax}}\, \frac{1}{B}\sum_{i=1}^{B} 1\left(z_v^i = l\right); v \in \mathcal{L}\right\}. \tag{8}$$

The main challenge in a Markov chain Monte Carlo algorithm is to specify a proposal density $q(\cdot)$ in Equation (5) such that the Markov chain converges within a given computational budget. The naïve approach is to propose based on the prior model $p(\mathbf{z})$; however, an unsatisfactory acceptance rate and poor mixing can be expected in high-dimensional problems with spatial coupling in the prior and likelihood models. A similar conclusion is also valid for a Gibbs update (Gilks et al., 1995), where the elements of $\mathbf{z}$ are updated one at a time conditional on every other element, e.g., $z_j \sim q(z_j \mid \mathbf{z}_{-j}^{i-1})$ and $\mathbf{z} = \left(z_1^{i-1}, \ldots, z_{j-1}^{i-1}, z_j, z_{j+1}^{i-1}, \ldots, z_n^{i-1}\right)$. There exists a vast literature on various algorithms to overcome these convergence problems in various applications, such as delayed rejection (Trias et al., 2009) or Hamiltonian/hybrid methods (Neal, 2012; Betancourt, 2017).

Reservoir characterization

We now discuss inverse problems that arise from problems in reservoir characterization. In an exploration phase, it is important to characterize reservoir variables of the subsurface to construct initial models for the proportion of hydrocarbons and fluid flow in the subsurface. Quantitative interpretation (Avseth et al., 2005) has become the industry standard for quantifying reservoir variables over the last decades. These models are used in a decision theory framework to decide whether a reservoir is to be developed for commercial use.

We focus on characterizing the subsurface discretized on the three-dimensional cube $\mathcal{L}$ and represent the subsurface by three variables. Let $\boldsymbol{\kappa}$ denote the lithology/fluid classes of the subsurface. The lithology/fluid classes may also be referred to as facies or 'rock types'. For each grid cell $v \in \mathcal{L}$, we assume $\kappa_v \in \Omega_\kappa = \{1, \ldots, L\}$; that is, κ_v takes one out of L categorical values, for example shale, gas sandstone or brine sandstone. Next, let $\mathbf{r}$ denote the petrophysical properties, such as the porosity ϕ, the clay volume/proportion c or the water saturation s_w of the subsurface. The petrophysical properties are generally constrained to be on the unit interval; that is, $r_v \in \Omega_r = [0, 1]$ for each $v \in \mathcal{L}$. We consider a one-to-one transformation of $\mathbf{r}$ with support on $\mathbb{R}$ to avoid restricted support; however, we refer to petrophysical properties as $\mathbf{r}$ for ease of discussion. Finally, let $\mathbf{m}$ denote the elastic attributes of the subsurface, where $m_v \in \Omega_m = [0, \infty)$ for each $m_v \in \mathcal{L}$. The elastic material properties are generally described by the pressure wave velocity v_p, the shear wave velocity v_s and the density ρ of the subsurface. Alternative parametrizations exist, for example, by a combination of the above-mentioned parameters such as the P-impedance ρv_p, the S-impedance ρv_s or the Poisson ratio v_p/v_s. In practice, we consider the logarithm of the elastic attributes, but we refer to it as $\mathbf{m}$. In the geoscience literature, these variables are often referred to as (model) parameters; however, we refer to them as (model) variables, as they are in principle observable by well measurements. In summary, we focus on the assessment of $\boldsymbol{\kappa}, \mathbf{r}$ and $\mathbf{m}$ defined on $\mathcal{L}$.

The motion of the seismic waves propagating through the layers of the subsurface is described by the wave equation. The resulting seismic data consist only of information on contrasts in the subsurface and relate the two-way traveltime to the seismic amplitude of the elastic attributes of interest. In marine seismic acquisition, reflections of seismic waves in the subsurface are recorded at a set of receivers for various incident angles. Each incident angle represents the angle at which a ray-path impinges upon a line normal to an interface. Note that the dimension of the variables of interest is often larger than the dimension of the acquired observations, since we have a limited number of incident angles. We consider seismic amplitude versus offset (AVO) observations, denoted by $\mathbf{d}$, discretized on $\mathcal{L}$. The Gauss-linear likelihood model is defined as in Buland and Omre (2003),

$$\big[\mathbf{d} \mid \boldsymbol{\kappa}, \mathbf{r}, \mathbf{m}\big] = \mathbf{WADm} + \boldsymbol{\varepsilon}_{\mathbf{d}}, \tag{9}$$

where $\mathbf{W}$ is a dense matrix representing the convolution kernel, $\mathbf{A}$ is a matrix approximation of the Aki-Richards coefficients valid for weak contrasts (Aki and Richards, 1980), $\mathbf{D}$ is a matrix approximation of the derivative, and $\boldsymbol{\varepsilon}_{\mathbf{d}}$ is a zero-mean Gaussian error vector. Note that Equation 9 implies that $p\big(\mathbf{d} \mid \boldsymbol{\kappa}, \mathbf{r}, \mathbf{m}\big) = p\big(\mathbf{d} \mid \mathbf{m}\big)$, since $\mathbf{m}$ is a canonical variable for $\mathbf{d}$.

In Figure 1, we present an example of a wavelet in the vertical direction used in reservoir characterization. Considering a specific depth v, the observation d_v depends on past ($< v$) and future ($> v$) values of the elastic attributes along the vertical profile. For the presented wavelet, each obser-

vation d_v depends on the elastic attributes in a time window in a vertical direction of approximately 200 ms. The recorded seismic data are heavily processed before they are used for quantitative interpretation by accounting for various effects from the acquisition procedure, such as the removal of multiples and attenuation. We do not consider processing problems but account for them by including the modeling error $\boldsymbol{\varepsilon}_{\mathbf{d}}$.

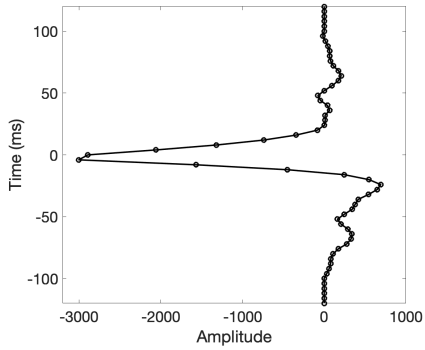


Figure 1: Wavelet used in reservoir characterization.

The resolution of seismic images is also often less than the desired resolution. The reason why the resolution of the seismic signal is less than the desired resolution in the vertical direction is the convolution. In Figure 2, we display near- and far-angle AVO observations of a seismic volume. The spatial continuity of the reservoir variable has a far greater lateral extent than vertical extent. Since seismic observations are band-limited, thin horizontal layers need not be easily observable in the seismic observations for

a given vertical trace but may be observable across a cross-section.

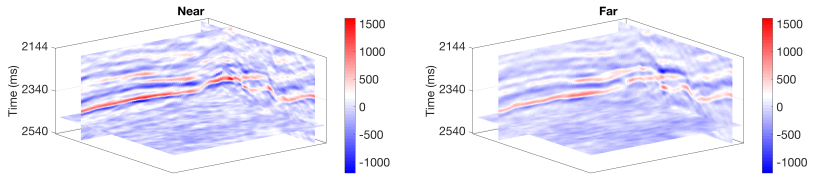


Figure 2: Near- and far-angle 3D volumes of seismic AVO measurements.

The main objective of reservoir characterization is to describe the reservoir properties $\boldsymbol{\kappa}, \mathbf{r}$ and $\mathbf{m}$ from geophysical measurements $\mathbf{d}$. Assessment of these variables is phrased in three distinct inverse problems: assessment of elastic attributes given seismic data (seismic or elastic inversion), assessment of petrophysical properties (petrophysical or rock physics inversion), and assessment of lithology/fluid classes (lithology/fluid class or facies inversion). Traditionally, these inverse problems are often solved in a stepwise procedure. First, the seismic data are used to obtain a prediction of the elastic attributes. The predictions are next used as an input in a petrophysical inversion framework (Doyen, 2007). Finally, the predicted petrophysical properties are used to classify the lithology/fluid classes. Such a stepwise procedure is known to often underpredict the uncertainty. Recently, joint approaches, where the variables are assessed simultaneously, have been introduced (Rimstad and Omre, 2010; Grana and Della Rossa, 2010). In this thesis, we consider a joint approach.

More specifically, we consider a sequential model (Figure 3) relating

seismic data $\mathbf{d}$ to the variables of interest $\boldsymbol{\kappa}, \mathbf{r}$ and $\mathbf{m}$. This is specified as the following sequential decomposition of the prior density:

$$p(\boldsymbol{\kappa}, \mathbf{r}, \mathbf{m}) = p\big(\mathbf{m} \mid \boldsymbol{\kappa}, \mathbf{r}\big) p\big(\mathbf{r} \mid \boldsymbol{\kappa}\big) p(\boldsymbol{\kappa}), \tag{10}$$

where $p\big(\mathbf{m} \mid \boldsymbol{\kappa}, \mathbf{r}\big)$ is the rock physics model, $p\big(\mathbf{r} \mid \boldsymbol{\kappa}\big)$ is the petrophysics model and $p(\boldsymbol{\kappa})$ is the lithology/fluid class model. We assume the rock physics model and the petrophysics model to be Gaussian random fields. For $p(\boldsymbol{\kappa})$, we consider various models that will be discussed later.

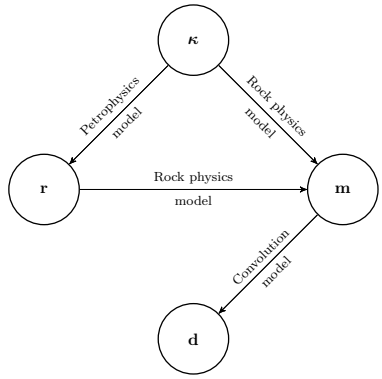


Figure 3: Directed acyclic graph relating the model variables and observations.

Rock physics models (Mavko et al., 2009) describe the relationships between reservoir properties, such as lithology/fluid classes $\boldsymbol{\kappa}$, petrophysical properties $\mathbf{r}$, and elastic attributes $\mathbf{m}$. Houck (2002) discuss the importance of considering rock physics model uncertainties in a lithology/fluid prediction setting. The uncertainty arises from the potential mis-specification of the physical relationships (Mukerji et al., 2001; Avseth et al., 2005; Bachrach, 2006). Landrø (2001) considered a linear approximation of the rock physics model, and various linearized rock physics models are discussed in Grana (2016). We restrict ourselves to linear rock physics models. In Figure 4, we display a reference classification of the lithology/fluid classes at a well location together with the water saturation and log P-impedance measurements. In this example, there are two different types of sandstone, namely, brine-filled (water-filled) and gas-filled. The value of the observed water saturation log is strongly dependent on the corresponding reference lithology/fluid class log. For example, for a gas sandstone layer, the observed values of water saturation are close to zero, while they are close to unity for a brine sandstone layer. Note that the histograms are observed to be bimodal and skewed. In general, both the petrophysical properties and elastic attributes are often observed to be multimodal and skewed due to the presence of various fluid and saturation effects of the subsurface. Finally, it is well known that some lithology/fluid transitions in $\boldsymbol{\kappa}$ are geophysically non-valid due to gravitational sorting. For example, brine sandstone can never occur above gas sandstone.

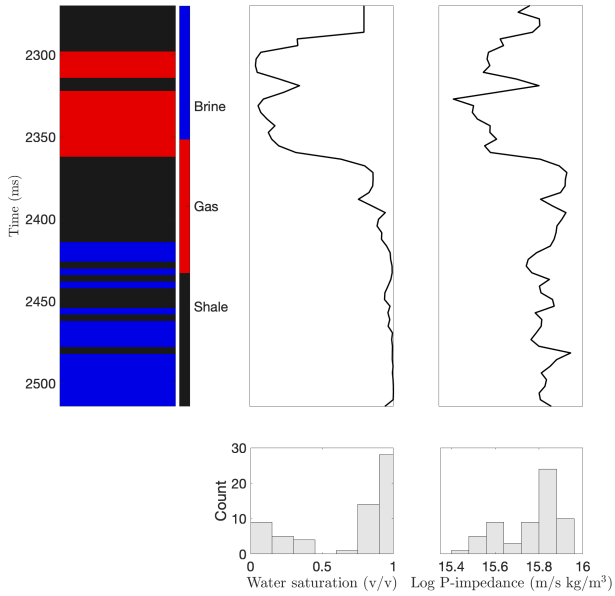


Figure 4: Well measurements: Reference classification of the lithology/fluid classes (left), water saturation (middle) and log P-impedance (right). Histograms of the continuous-valued properties are displayed.

Bayesian spatial inversion

Recall the definition of the likelihood model in Equation (9) and the prior model in Equation (10). We consider the following probabilistic model:

$$\begin{aligned} [\boldsymbol{\kappa}, \mathbf{r}, \mathbf{m}] &\sim p(\boldsymbol{\kappa}, \mathbf{r}, \mathbf{m}) \qquad \text{prior model} \\ \big[\mathbf{d} \mid \boldsymbol{\kappa}, \mathbf{r}, \mathbf{m}\big] &\sim p\big(\mathbf{d} \mid \mathbf{m}\big) \quad \text{likelihood model} \end{aligned} . \tag{11}$$

The ultimate objective is to assess the variables $\boldsymbol{\kappa}, \mathbf{r}$ and $\mathbf{m}$ in an inverse setting. We refer to the characterization of $\boldsymbol{\kappa}, \mathbf{r}$ and $\mathbf{m}$ as variable prediction since these quantities are in principle observable. Note that the prior density $p(\boldsymbol{\kappa}, \mathbf{r}, \mathbf{m})$ includes both discrete- and continuous-valued variables.

Recall the definitions of the likelihood model in Equation (9) and the prior model in Equation (10). In a Bayesian framework, the objective is to assess the posterior of the variables given the observations,

$$p\big(\boldsymbol{\kappa}, \mathbf{r}, \mathbf{m} \mid \mathbf{d}\big) = \text{const} \times p\big(\mathbf{d} \mid \mathbf{m}\big) p\big(\mathbf{m} \mid \boldsymbol{\kappa}, \mathbf{r}\big) p\big(\mathbf{r} \mid \boldsymbol{\kappa}\big) p(\boldsymbol{\kappa}), \tag{12}$$

where const is a normalizing constant that ensures the left-hand side of Equation (12) is a valid density. Finding the normalizing constant is in practice not computationally feasible since it requires evaluating the normalizing constant, that is, the marginal likelihood $p(\mathbf{d})$.

We operate mainly in an empirical Bayes setting, where the model parameters are fixed and known based on geophysical knowledge from comparable geological areas. The posterior densities of the reservoir variables are given as follows:

$$\begin{array}{rl} p\big(\boldsymbol{\kappa} \mid \mathbf{d}\big) \propto p\big(\mathbf{d} \mid \boldsymbol{\kappa}\big) p(\boldsymbol{\kappa}) & \text{lithology/fluid class inversion} \\ p\big(\mathbf{r}, \mathbf{m} \mid \mathbf{d}\big) \propto p\big(\mathbf{d} \mid \mathbf{m}\big) p(\mathbf{r}, \mathbf{m}) & \text{petrophysical and elastic inversion} \end{array}. \tag{13}$$

Next, we discuss the assessment of these inverse problems separately.

Lithology/fluid class inversion

First, we consider the problem of lithology/fluid class inversion. Recall that the posterior of interest is given as

$$p\big(\boldsymbol{\kappa} \mid \mathbf{d}\big) \propto p\big(\mathbf{d} \mid \boldsymbol{\kappa}\big) p(\boldsymbol{\kappa}). \tag{14}$$

In the following, we discuss the likelihood model $p\big(\mathbf{d} \mid \boldsymbol{\kappa}\big)$ and various choices for the prior density $p(\boldsymbol{\kappa})$.

Likelihood model

The likelihood defined in Equation (14) is given as

$$p\big(\mathbf{d} \mid \boldsymbol{\kappa}\big) = \int p\big(\mathbf{d} \mid \mathbf{m}\big) \int p\big(\mathbf{m} \mid \boldsymbol{\kappa}, \mathbf{r}\big) p\big(\mathbf{r} \mid \boldsymbol{\kappa}\big) \mathrm{d}\mathbf{r}\, \mathrm{d}\mathbf{m}, \tag{15}$$

which is analytically tractable, since the densities are all Gaussian densities. Note that the likelihood $p\big(\mathbf{d} \mid \boldsymbol{\kappa}\big)$ cannot be written as a product of the marginal conditional densities $p\big(d_v \mid \kappa_v\big)$. The reason is that each datum d_v is dependent on a large subset of $\boldsymbol{\kappa}$ because of the convolution in $p\big(\mathbf{d} \mid \mathbf{m}\big)$.

Prior models

To build initial models for the reservoir, it is important to model the proportion of hydrocarbons and the lateral connectivity and channels where the fluids are allowed to flow. We refer to Rimstad et al. (2012) for a discussion of its importance. We focus on constructing a prior model $p(\boldsymbol{\kappa})$ that follows

geophysical properties such as the ordering of the lithology/fluid classes, non-valid transitions and spatial connectivity. Recall that the lateral extent of the lithology/fluid classes is far greater than the vertical extent. Next, we discuss the one-dimensional vertical model based on a Markov chain model and the more complex spatially coupled models in two and three dimensions based on the Markov random field model and the Markov mesh model.

Markov chain model

In the following, we consider only a one-dimensional vertical profile discretized on $\mathcal{L}_{xy\cdot} = \{xyt : t = 1, \ldots, n_t\}$ at a given horizontal position xy and assume that $p(\boldsymbol{\kappa})$ follows a first-order Markov chain. The use of Markov chain models in geophysical applications dates back to Krumbein and Dacey (1969). Formally, the first-order Markov chain prior model on $\mathcal{L}_{xy\cdot}$ is defined as

$$p(\boldsymbol{\kappa}) = p(\kappa_1) \times \prod_{t=2}^{n_t} p\big(\kappa_t \mid \kappa_{t-1}\big), \tag{16}$$

where $p(\kappa_1)$ is the stationary probability and $p\big(\kappa_t \mid \kappa_{t-1}\big)$ is the transition probability of going from a specific class at layer $t-1$ to a specific class at layer t. Markov chain prior models have become increasingly popular in geoscience applications during the last two decades due to their interpretability and intuitive understanding of the ordering and sorting of lithology/fluid classes in the prior model (Elfeki and Dekking, 2001; Eidsvik et al., 2004;

Larsen et al., 2006; Ulvmoen and Omre, 2010; Grana and Della Rossa, 2010; Rimstad and Omre, 2013; Connolly and Hughes, 2016; de Figueiredo et al., 2019). This implies that it is possible to restrict non-valid geophysical transitions, such as brine sandstone above gas sandstone. We refer to Moja et al. (2018) for a discussion of non-stationary Markov chain models, where various marginal probability and conditional independence properties are discussed. Equation (16) can be extended to a higher-order Markov chain to incorporate, for example, knowledge of the minimum thickness of each layer. We refer to Talarico et al. (2019) for a discussion of higher-order Markov chains with applications in reservoir characterization.

Markov random field model

Consider the undirected graph $\mathcal{G} = \{\mathcal{L}, \mathcal{E}\}$, where $\mathcal{E} \subseteq \{\{u, v\} : u, v \in \mathcal{L}, u \neq v\}$ defines the set of edges. A clique is defined to be a subset $c \in \mathcal{L}$ of $\mathcal{G}$ such that every pair of distinct vertices in the clique are adjacent. A maximal clique c is a clique that is not a subset of another clique, and we denote the set of maximal cliques of $\mathcal{G}$ by $\mathcal{C}$. Let $n(v) \in \mathcal{L}$ be the set of neighbors of each $v \in \mathcal{L}$.

We specify the Markov random field prior model on $\mathcal{L}$ based on the Gibbs formulation (Besag, 1974; Kindermann et al., 1980)

$$p(\boldsymbol{\kappa}) = \text{const}^{-1} \times \exp\left\{ -\sum_{c \in \mathcal{C}} \Phi_c(\boldsymbol{\kappa}_c) \right\}, \tag{17}$$

where $\Phi_c(\boldsymbol{\kappa}_c)$ is the clique potential function for the maximal clique

c and $\boldsymbol{\kappa}_c = \{\kappa_v; v \in c\}$. The normalizing constant const $= \sum_{\boldsymbol{\kappa}' \in \Omega_\kappa^n} \exp\left\{-\sum_{c \in \mathcal{C}} \Phi_c(\boldsymbol{\kappa}'_c)\right\}$ is not computationally tractable, since the sum is exponential in the number of grid cells. For sufficiently small grids, computation of the normalizing constant may be computationally tractable using recursions relying on factorisable models (Reeves and Pettitt, 2004; Friel and Rue, 2007). The Markov chain model defined previously is an example of a one-dimensional vertical Markov random field with a neighborhood system consisting of the two closest neighbors. We refer to Stoehr (2017) and references therein for an overview of Markov random fields.

For a fixed set of model parameters, simulation from a Markov random field is often performed using the Metropolis-Hastings algorithm since the troublesome normalizing constant cancels in the acceptance ratio in Equation (6). Simulation from a Markov random field is often done by simulation from the (pointwise) Markov formulation

$$p\big(\kappa_v \mid \boldsymbol{\kappa}_{-v}\big) = p\Big(\kappa_v \mid \boldsymbol{\kappa}_{n(v)}\Big) \tag{18}$$

until convergence. This is a Gibbs step, since we sample from the set of full-conditional densities. The correspondence between the Gibbs formulation in Equation (17) defined by the clique system and the set of the full-conditional densities $p\Big(\kappa_v \mid \boldsymbol{\kappa}_{n(v)}\Big)$ defined by the neighborhood system is given by the Hammersley-Clifford theorem (Hammersley and Clifford, 1971). An alternative to simulation based on Equation (18) is to simulate each vertical profile jointly conditioned on every other vertical profile, e.g.,

perform simulation based on joint vertical block update. That is, simulate from $p\big(\boldsymbol{\kappa}_{xy\cdot} \mid \boldsymbol{\kappa}_{-xy\cdot}\big)$, where $\boldsymbol{\kappa}_{xy\cdot} = (\kappa_{xy1}, \ldots, \kappa_{xyn_t})$ and $\boldsymbol{\kappa}_{-xy\cdot} = \boldsymbol{\kappa} \setminus \boldsymbol{\kappa}_{xy\cdot}$. For a first-order Markov random field, the conditional density for each vertical profile can be expressed in factorial form as

$$p\big(\boldsymbol{\kappa}_{xy\cdot} \mid \boldsymbol{\kappa}_{-xy\cdot}\big) = p\big(\kappa_{xy1} \mid \boldsymbol{\kappa}_{-xy\cdot}\big) \times \prod_{t=2}^{n_t} p\big(\kappa_{xyt} \mid \kappa_{xy,t-1}, \boldsymbol{\kappa}_{-xy\cdot}\big), \tag{19}$$

which makes sequential simulation feasible. Extensions of Equation (19) for Markov random fields based on a larger neighborhood system are valid.

Markov mesh model

An alternative to the Markov random field parametrization in two dimensions is the Markov mesh model (Abend et al., 1965; Cressie and Davidson, 1998; Stien and Kolbjørnsen, 2011; Luo and Tjelmeland, 2017). We consider a lexiographic ordering of the nodes in $\mathcal{L}_{x\cdot\cdot} = \{xyt : y = 1, \ldots, n_y, t = 1, \ldots, n_t\}$ for a fixed x and assume κ_v to be a binary random variable for each $v \in \mathcal{L}_{x\cdot\cdot}$. The dimension of the lattice is $n_y \times n_t$. Let

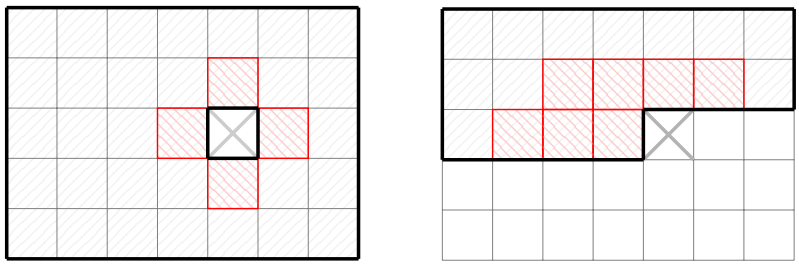
$$\rho(v) = \{(k, l) \in \mathcal{L}_{x\cdot\cdot} : n_t k + l < n_t y + t\} \tag{20}$$

be the set of predecessors for each $v = (y, t) \in \mathcal{L}_{x\cdot\cdot}$. The Markov mesh model is then defined as

$$p(\boldsymbol{\kappa}) = \prod_{v \in \mathcal{L}_{x\cdot\cdot}} p\Big(\kappa_v \mid \boldsymbol{\kappa}_{\rho(v)}\Big) = \prod_{v \in \mathcal{L}_{x\cdot\cdot}} p\Big(\kappa_v \mid \boldsymbol{\kappa}_{\nu(v)}\Big), \tag{21}$$

where $\nu(v) \subseteq \rho(v)$ is called the sequential neighborhood of $v \in \mathcal{L}_{x^{\cdot\cdot}}$ and is defined as a translation of a template sequential neighborhood (Luo and Tjelmeland, 2017).

In regard to simulation, there are several approaches for simulating from the Markov mesh model. Simulation based on the Markov mesh formulation is traditionally done by sequentially simulating from $p\Big(\kappa_v \mid \boldsymbol{\kappa}_{\nu(v)}\Big)$, while simulation based on the Markov random field formulation is performed by the full conditionals $p\big(\kappa_v \mid \boldsymbol{\kappa}_{-v}\big)$; see Figure 5. Note that any random field specified from the Markov mesh model is a Markov random field; however, the contrary is not true. An arbitrary Markov random field defined from the Gibbs formulation (Equation (17)) may not be able to be phrased as a Markov mesh model. In summary, simulation from a Markov mesh model is feasible by simulation from an extended version of the conditional Markov chains defined in Equation (19).



(a) Markov random field model **(b)** Markov mesh model

Figure 5: Conditioning set based on the Markov random field and the Markov mesh model. The grid cell $v \in \mathcal{L}_{x\cdot\cdot}$ to be simulated is marked by a gray cross, the full conditioning set ($\boldsymbol{\kappa}_{-v}$ and $\boldsymbol{\kappa}_{\rho(v)}$) is in light gray and the corresponding reduced conditioning set ($\boldsymbol{\kappa}_{n(v)}$ and $\boldsymbol{\kappa}_{\nu(v)}$) is in red.

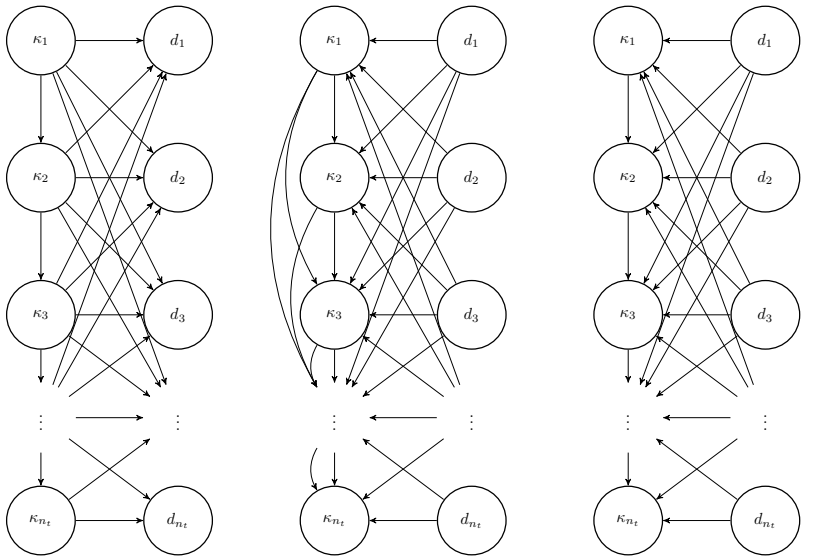
Posterior model

The vertical spatial blurring and averaging caused by the convolution in the likelihood model defined in Equation (9) restricts the analytical and fast simulation-based assessment of $p\big(\boldsymbol{\kappa} \mid \mathbf{d}\big)$ defined in Equation (14). Consider the one-dimensional vertical case based on a first-order Markov chain prior defined on $\mathcal{L}_{xy\cdot}$: $p\big(\boldsymbol{\kappa} \mid \mathbf{d}\big) \propto p\big(\mathbf{d} \mid \boldsymbol{\kappa}\big) \times p(\kappa_1) \times \prod_{t=2}^{n_t} p\big(\kappa_t \mid \kappa_{t-1}\big)$, where we have omitted xy for ease of notation. Because of the ordering and possible non-valid transitions in the prior model, numerical optimization of the posterior model is challenging. It can be demonstrated that the convolution in the likelihood model defines a convolved hidden Markov model (Lindberg

and Omre, 2014), where each datum d_v is dependent on a large subset of $\boldsymbol{\kappa}$. In Figure 6a, the model is visualized as a first-order Markov chain for $\boldsymbol{\kappa}$ and a convolutional model $p\big(\mathbf{d} \mid \boldsymbol{\kappa}\big)$. The exact inverse model $p\big(\boldsymbol{\kappa} \mid \mathbf{d}\big)$ is then a higher-order Markov chain (Figure 6b), since the convolutional model extends the simpler first-order neighborhood structure that occurs in the well-known hidden Markov model (Frühwirth-Schnatter, 2006). Assessment of a kth-order Markov chain generally requires $\mathcal{O}(L^{k+1})$ operations, which becomes computationally unfeasible even for modest values of k. The exact posterior model is approximated by constructing a kth-order Markov chain approximation (Figure 6c) of the form

$$\tilde{p}\left(\boldsymbol{\kappa} \mid \mathbf{d}\right) = \tilde{p}\left(\boldsymbol{\kappa}_{1:k} \mid \mathbf{d}\right) \times \prod_{t=k+1}^{n_t} \tilde{p}\left(\kappa_t \mid \boldsymbol{\kappa}_{(t-k):(t-1)}, \mathbf{d}\right), \tag{22}$$

which is assessed analytically by a recursive algorithm (Reeves and Pettitt, 2004; Friel and Rue, 2007). Various approximations exist here, for example, based on truncation of the likelihood or more complex approximations (Rimstad and Omre, 2013). Finally, the approximate density $\tilde{p}\left(\boldsymbol{\kappa} \mid \mathbf{d}\right)$ is used as an independent proposal density in a Markov chain Monte Carlo Metropolis-Hastings algorithm (Equation (5)) to simulate based on $p\big(\boldsymbol{\kappa} \mid \mathbf{d}\big)$. This is a so-called independent proposal density, since each proposal density is independent of the current value of $\boldsymbol{\kappa}$. In two and three dimensions, the approximation defined in Equation (22) is constructed by also conditioning on the neighboring traces based on Equation (19).



(a) Forward model **(b)** Inverse model **(c)** Approximate model ($k = 1$)

Figure 6: Dependence structure of the forward, inverse and first-order approximate inverse model for the convolved hidden Markov model.

Petrophysical and elastic inversion

Thus far, we have only discussed the evaluation of $p(\boldsymbol{\kappa} \mid \mathbf{d})$. Recall that one of the objectives is to construct a class of continuous-valued prior models that includes multimodal and skewed marginal densities. We consider the Gaussian mixture prior density, which has become increasingly popular in the last decade in reservoir characterization (Grana and Della Rossa, 2010).

Likelihood model

Recall that the likelihood model is defined as

$$p\big(\mathbf{d} \mid \mathbf{r}, \mathbf{m}\big) = p\big(\mathbf{d} \mid \mathbf{m}\big). \tag{23}$$

Note that the likelihood cannot be factorized by including only the set of univariate densities $p\big(d_v \mid m_v\big)$ because of the convolution, as was the case for Equation (15).

Prior model

We consider the following spatially coupled Gaussian mixture prior density:

$$p(\mathbf{r}, \mathbf{m}) = \sum_{\boldsymbol{\kappa} \in \Omega_\kappa^n} p\big(\mathbf{r}, \mathbf{m} \mid \boldsymbol{\kappa}\big) p(\boldsymbol{\kappa}), \tag{24}$$

where $p\big(\mathbf{r}, \mathbf{m} \mid \boldsymbol{\kappa}\big)$ is a Gaussian density for which the mean and covariance are defined from the corresponding value of $\boldsymbol{\kappa}$. The high-dimensional Gaussian mixture model includes spatial coupling because of the spatial coupling in $p(\boldsymbol{\kappa})$. One interpretation of $\boldsymbol{\kappa}$ is that it acts as a spatially connected mode identifier that assigns a corresponding conditional Gaussian density $p\big(\mathbf{r}, \mathbf{m} \mid \boldsymbol{\kappa}\big)$ for each value of $\boldsymbol{\kappa}$. Furthermore, $p\big(\mathbf{r}, \mathbf{m} \mid \boldsymbol{\kappa}\big)$ may be dependent on a (parametric) spatial correlation function.

Posterior model

For the Gauss-linear likelihood model defined in Equation (9), the Gaussian mixture density is a conjugate prior model. Hence, the posterior density is

a Gaussian mixture density, where the model parameters have analytic expressions given by the likelihood and prior model parameters. The posterior density is given by

$$p\big(\mathbf{r}, \mathbf{m} \mid \mathbf{d}\big) = \sum_{\boldsymbol{\kappa} \in \Omega_{\kappa}^{n}} p\big(\mathbf{r}, \mathbf{m} \mid \boldsymbol{\kappa}, \mathbf{d}\big) p\big(\boldsymbol{\kappa} \mid \mathbf{d}\big). \tag{25}$$

Note that the mixing weights $p\big(\boldsymbol{\kappa} \mid \mathbf{d}\big)$ are identical to the posterior density obtained by the lithology/fluid class inversion in Equation (14). One consequence of Equation (25) is that the posterior model is analytically tractable, since the set of densities $p\big(\mathbf{r}, \mathbf{m} \mid \boldsymbol{\kappa}, \mathbf{d}\big)$ are Gaussian densities having analytical expressions. However, the dimensions of $\boldsymbol{\kappa}, \mathbf{r}$ and $\mathbf{m}$ make it unfeasible to compute all possible configurations of $\boldsymbol{\kappa}$. Thus, we are only able to generate an ensemble of realizations from these posterior densities by simulation. It can be verified that the Gaussian mixture density is closed under marginalization. Thus, the marginal densities $p\big(r_v \mid \mathbf{d}\big)$ and $p\big(m_v \mid \mathbf{d}\big)$ are univariate Gaussian mixture densities.

Summary of papers

The objective of this PhD thesis is to construct a class of computationally feasible models in a Bayesian spatial inverse framework for the joint characterization of lithology/fluid classes, petrophysical properties and elastic attributes in three-dimensional reservoir characterization.

In the following, we summarize the papers included in this thesis. They

constitute the scientific contribution of this PhD thesis. Each paper is considered to be self-contained; hence, the papers can be read in an arbitrary order.

PAPER I

Bayesian Gaussian Mixture Linear Inversion for Geophysical Inverse Problems

Dario Grana, Torstein Fjeldstad and Henning Omre

Published in Mathematical Geosciences vol. 49 (4) (2017)

The problems of both seismic inversion and petrophysical inversion are cast in Bayesian frameworks based on spatially coupled Gaussian mixture prior models. The two inverse problems are both phrased based on a Gauss-linear likelihood model and Gaussian mixture prior density. That is, the proposed workflow is valid for both seismic and petrophysical inversion. The spatially coupled Gaussian mixture prior is discussed and demonstrated on two real case studies jointly with lithology/fluid class prediction. The results based on the Gaussian mixture prior density are compared to results based on a simpler Gaussian prior model for two real case studies, and an improvement for the root-mean-squared error, typically in the range $10-30$ %, is obtained.

Main contribution: rigorous formulation of petrophysical and seismic inversions based on spatially coupled Gaussian mixture prior densities.

PAPER II

Bayesian Inversion of Convolved Hidden Markov Models With Applications in Reservoir Prediction

Torstein Fjeldstad and Henning Omre

To appear in IEEE Transactions on Geoscience and Remote Sensing

Two classes of approximations for a long-memory 1D Markov chain for the convolved hidden Markov model based on a Gauss-linear likelihood model are proposed. These approximate low-order Markov chains are assessed analytically by the forward-backward algorithm. Assessment of the correct posterior model is performed by an independent proposal Markov chain Monte Carlo algorithm together with a delayed rejection step. The two proposed approximations are validated on a set of synthetic cases, and their performance is compared. The favorable approximation is demonstrated in a real 1D case study from the Norwegian Sea to predict lithology/fluid classes and elastic attributes.

Main contribution: two classes of kth-order Markov chain approximations are introduced, and their performance is discussed.

PAPER III

Joint probabilistic petrophysics-seismic inversion based on Gaussian mixture and Markov chain prior models

Torstein Fjeldstad and Dario Grana

Published in Geophysics vol. 83 (1) (2018)

The two distinct inverse problems defined in Paper I are combined into a joint probabilistic model that allows for joint Bayesian characterization of lithology/fluid classes, petrophysical properties and elastic attributes based on the Gauss-linear likelihood model. The posterior model in 2D is assessed using a block-Gibbs simulation algorithm, where each vertical profile is updated conditionally on the remaining vertical profiles. The proposed methodology is demonstrated on seismic data from a 2D cross-section from the Norwegian Sea reservoir to predict the occurrence of gas sandstone.

Main contribution: sequential specification of a spatially coupled Bayesian model for joint prediction of lithology/fluid class, petrophysical properties and elastic attributes.

PAPER IV

Bayesian model for lithology/fluid class prediction using a Markov mesh prior fitted from a training image

Håkon Tjelmeland, Xin Luo and Torstein Fjeldstad

Published in Geophysical Prospecting vol. 67 (3) (2019)

Two classes of prior models for inversion of binary lithology/fluid classes based on seismic data in a real 2D case study are discussed. The first prior model is based on a manually specified Markov random field with a first-order neighborhood system for the lithology/fluid classes. The second prior model is based on a Markov mesh prior model, where the neighborhood and the associated model parameters are estimated from a training image in an empirical Bayesian framework. These prior models are combined with a Gauss-linear likelihood model to assess the posterior density of the lithology/fluid classes. The difference between the posteriors for the marginal probabilities is small but observable. However, the connectivity in the posterior realizations is vastly different. The larger neighborhood of the Markov mesh model enables the identification of long-range connectivity, which is important for modeling fluid flow.

Main contribution: impact of a neighborhood system with a complex structure in reservoir characterization.

PAPER V

A one-step Bayesian inversion framework for three-dimensional reservoir characterization based on a Gaussian mixture model – A Norwegian Sea demonstration

Torstein Fjeldstad, Per Åge Avseth and Henning Omre

Submitted

Joint spatial Bayesian inversion of lithology/fluid classes, petrophysical properties and elastic attributes is considered in three dimensions based on a Gauss-linear likelihood model. A recursive algorithm that translates the Gibbs formulation for a Markov random field into a set of vertical Markov chain transition probabilities is proposed. The proposed algorithm is demonstrated on a 3D gas discovery from the Norwegian Sea. We compare the proposed model to a simpler model where the vertical profiles are assumed to be conditionally independent of every other vertical profile. Both methods are validated at a blind well location, and the proposed model based on a Markov random field prior obtains a reduction of the root-mean-squared error of up to 60 %. Posterior realizations from the 3D model are observed with a stronger lateral connectivity.

Main contribution: a recursive algorithm that translates a 3D Markov random field into a set of conditional vertical 1D Markov chains.

In Table 1, we present a short summary of the included variables and posterior densities of interest. In summary, the main scientific contributions of this thesis are the following: **Paper I** formalizes the Gaussian mixture prior density in cases of petrophysical and seismic inversion (Equation (24)) and is written for a geoscience audience. **Paper II** presents two classes of statistical approximations of Equation (22) and discusses their performance as proposal densities. Specifically, their impact on the acceptance rate (Equation (6)) is discussed. The focus is on the statistical properties of the proposed approximations, and the paper is therefore aiming at a statistical audience. **Paper III** presents a sequential model for joint lithology/fluid class and petrophysical and elastic inversion of Equation (13) and is written for a geophysical audience. A linearized statistical rock physics model and block-Gibbs simulation algorithm are discussed. **Paper IV** introduces the use of the Markov mesh prior model (Equation (21)) in seismic inversion. Its impact on connectivity in a real case study is studied and compared with a simpler manually specified Markov random field. The paper targets a geoscience audience. **Paper V** can be viewed as the natural extension of previous papers and discusses the assessment of Equation (12). An algorithm to simulate a three-dimensional Markov random field (Equation (17)) based on recursive simulations of vertical Markov chains is presented. The 3D Markov random field posterior model (Equation (17)) is compared to a set of 1D vertical Markov chain posterior models (Equation (16)) at a blind well location. The 3D Markov random field prior model obtained a

reduction of up to 60 % in root-mean-squared error for the petrophysical properties compared to the alternative trace-independent model. The paper is written for a geophysical audience with a strong background in statistical modeling.

Table 1: Summary of the models discussed in Papers I-V.

Paper	**Included variables**	**Number of classes**	**Dimension of reservoir model**	**Posterior densities of interest**
I	$\boldsymbol{\kappa}$ and $\mathbf{r}$, and κ and $\mathbf{m}$	$L \geq 2$	1D	$p(\boldsymbol{\kappa}, \mathbf{r} \mid \mathbf{m})$ and $p(\boldsymbol{\kappa}, \mathbf{m} \mid \mathbf{d})$
II	$\boldsymbol{\kappa}$ and $\mathbf{m}$	$L \geq 2$	1D	$p(\boldsymbol{\kappa}, \mathbf{m} \mid \mathbf{d})$
III	$\boldsymbol{\kappa}, \mathbf{r}$ and $\mathbf{m}$	$L \geq 2$	2D	$p(\boldsymbol{\kappa}, \mathbf{r}, \mathbf{m} \mid \mathbf{d})$
IV	$\boldsymbol{\kappa}$	$L = 2$	2D	$p(\boldsymbol{\kappa} \mid \mathbf{d})$
V	$\boldsymbol{\kappa}, \mathbf{r}$ and $\mathbf{m}$	$L \geq 2$	3D	$p(\boldsymbol{\kappa}, \mathbf{r}, \mathbf{m} \mid \mathbf{d})$

References

Abend, K., Harley, T. J., and Kanal, L. N. (1965). Classification of binary random patterns. *IEEE Trans. Information Theory*, 11:538–544.

Aki, K. and Richards, P. (1980). *Quantitative Seismology: Theory and Methods*. W. H. Freeman and Co., New York.

Avseth, P., Mukerji, T., and Mavko, G. (2005). *Quantitative Seismic Interpretation: Applying Rock Physics Tools to Reduce Interpretation Risk*. Cambridge University Press.

Bachrach, R. (2006). Joint estimation of porosity and saturation using stochastic rock-physics modeling. *Geophysics*, 71(5):O53–O63.

Besag, J. (1974). Spatial Interaction and the Statistical Analysis of Lattice Systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, 36(2):pp. 192–236.

Betancourt, M. (2017). A conceptual introduction to hamiltonian monte carlo.

Bolin, D., Wallin, J., and Lindgren, F. (2019). Latent Gaussian random field mixture models. *Computational Statistics & Data Analysis*, 130(C):80–93.

Buland, A., Kolbjørnsen, O., and Omre, H. (2003). Rapid spatially coupled avo inversion in the fourier domain. *Geophysics*, 68(3):824–836.

Buland, A. and Omre, H. (2003). Bayesian linearized AVO inversion. *Geophysics*, 68(1):185–198.

Bárdossy, A. and Li, J. (2008). Geostatistical interpolation using copulas. *Water Resources Research*, 44(7).

Caers, J. (2001). Geostatistical reservoir modelling using statistical pattern recognition. *Journal of Petroleum Science and Engineering*, 29:177–188.

Connolly, P. A. and Hughes, M. J. (2016). Stochastic inversion by matching to large numbers of pseudo-wells. *Geophysics*, 81:M7–M22.

Cressie, N. (1993). *Statistics for spatial data*. Wiley series in probability and mathematical statistics: Applied probability and statistics. J. Wiley.

Cressie, N. and Davidson, J. L. (1998). Image analysis with partially ordered markov models. *Computational Statistics & Data Analysis*, 29(1):1 – 26.

de Figueiredo, L. P., Grana, D., Roisenberg, M., and Rodrigues, B. B. (2019). Gaussian mixture markov chain monte carlo method for linear seismic inversion. *Geophysics*, 84(3):R463–R476.

De Oliveira, V., Kedem, B., and Short, D. A. (1997). Bayesian prediction of transformed gaussian random fields. *Journal of the American Statistical Association*, 92(440):1422–1433.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38.

Diggle, P. J., Tawn, J. A., and Moyeed, R. A. (1998). Model-based geostatistics. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 47(3):299–350.

Doyen, P. (2007). *Seismic reservoir characterization: an earth modelling perspective*. Education tour series. EAGE publications.

Eidsvik, J., Mukerji, T., and Switzer, P. (2004). Estimation of Geological Attributes from a Well Log: An Application of Hidden Markov Chains. *Mathematical Geology*, 36(3):379–397.

Elfeki, A. M. M. and Dekking, M. (2001). A markov chain model for subsurface characterization: Theory and applications. *Mathematical Geology*, 33:569–589.

Ferguson, T. S. (1973). A bayesian analysis of some nonparametric problems. *The Annals of Statistics*, 1(2):209–230.

Friel, N. and Rue, H. (2007). Recursive computing and simulation-free inference for general factorizable models. *Biometrika*, 94(3):661–672.

Frühwirth-Schnatter, S. (2006). *Finite Mixture and Markov Switching Models*. Springer Series in Statistics. Springer New York.

Gilks, W., Richardson, S., and Spiegelhalter, D. (1995). *Markov Chain Monte Carlo in Practice*. Chapman & Hall/CRC Interdisciplinary Statistics. Taylor & Francis.

Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st edition.

Grana, D. (2016). Bayesian linearized rock-physics inversion. *Geophysics*, 81:D625–D641.

Grana, D. and Della Rossa, E. (2010). Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion. *Geophysics*, 75:O21–O37.

Hammersley, J. M. and Clifford, P. E. (1971). Markov random fields on finite graphs and lattices. Unpublished manuscript.

Hansen, T. M., Journel, A. G., Tarantola, A., and Mosegaard, K. (2006). Linear inverse Gaussian theory and geostatistics. *Geophysics*, 71:R101–R111.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data mining, Inference, and Prediction*. Springer Series in Statistics. Springer, New York, 2nd edition.

Houck, R. T. (2002). Quantifying the uncertainty in an AVO interpretation. *Geophysics*, 67(1):117–125.

Hurn, M. A., Husby, O. K., and Rue, H. (2003). *A Tutorial on Image Analysis*, pages 87–141. Springer New York, New York, NY.

Kindermann, R., Snell, J., and Society, A. M. (1980). *Markov Random Fields and Their Applications*. AMS books online. American Mathematical Society.

Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220(4598):671–680.

Krumbein, W. C. and Dacey, M. F. (1969). Markov chains and embedded Markov chains in geology. *Mathematical Geology*, 1:79–96.

Landrø, M. (2001). Discrimination between pressure and fluid saturation changes from time-lapse seismic data. *Geophysics*, 66(3):836–844.

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:R69–R78.

Lawson, A. (2018). *Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology, Third Edition*. Chapman & Hall/CRC Interdisciplinary Statistics. CRC Press.

Lindberg, D. and Omre, H. (2014). Blind categorical deconvolution in two-level hidden markov models. *IEEE Transactions on Geoscience and Remote Sensing*, 52:7435–7447.

Lindberg, D. and Omre, H. (2015). Inference of the transition matrix in convolved hidden markov models and the generalized baum–welch algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, 53:6443–6456.

Luo, X. and Tjelmeland, H. (2017). Prior specification for binary markov mesh models. *Statistics and Computing*, 29:367–389.

Matheron, G. (1963). Principles of geostatistics. *Economic Geology*, 58(8):1246–1266.

Matheron, G., Beucher, H., de Fouquet, C., Galli, A., Guérillot, D., and Ravenne, C. (1987). Conditional simulation of the geometry of fluvio-deltaic reservoirs. *Soc. Petrol. Eng. (SPE)*, 16753,.

Mavko, G., Mukerji, T., and Dvorkin, J. (2009). *The Rock Physics Handbook*. Cambridge University Press, second edition.

Moja, S., Asfaw, Z., and Omre, H. (2018). Bayesian inversion in hidden markov models with varying marginal proportions. *Mathematical Geosciences*, 51.

Mukerji, T., Jørstad, 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.

Neal, R. (2012). Mcmc using hamiltonian dynamics. *Handbook of Markov Chain Monte Carlo*.

Nocedal, J. and Wright, S. (2006). *Numerical optimization*. Springer series in operations research and financial engineering. Springer, New York, NY, 2. ed. edition.

Omre, H. and Rimstad, K. (2018). Bayesian spatial inversion and conjugate selection gaussian prior models.

Ortiz, J. M. and Deutsch, C. V. (2004). Indicator Simulation Accounting for Multiple-Point Statistics. *Mathematical Geology*, 36(5):545–565.

Reeves, R. and Pettitt, A. N. (2004). Efficient recursions for general factorisable models. *Biometrika*, 91:751–757.

Rimstad, K., Avseth, P., and Omre, H. (2012). Hierarchical bayesian lithology/fluid prediction: A north sea case study. *GEOPHYSICS*, 77(2):B69–B85.

Rimstad, K. and Omre, H. (2010). Impact of rock-physics depth trends and Markov random fields on hierarchical Bayesian lithology/fluid prediction. *Geophysics*, 75:R93–R108.

Rimstad, K. and Omre, H. (2013). Approximate posterior distributions for convolutional two-level hidden Markov models. *Computational Statistics & Data Analysis*, 58:187–200.

Rue, H. and Held, L. (2005). *Gaussian Markov Random Fields: Theory and Applications*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. CRC Press.

Sen, M. and Stoffa, P. (1996). Bayesian inference, gibbs' sampler and uncertainty estimation in geophysical inversion. *Geophysical Prospecting*, 44(2):313–350.

Skauvold, J., Eidsvik, J., and Theune, U. (2016). A parametric model for seismic wavelets—with estimation and uncertainty quantification. *Geophysical Journal International*, 205(2):796–809.

Stien, M. and Kolbjørnsen, O. (2011). Facies modeling using a markov mesh model specification. *Mathematical Geosciences*, 43(6):611.

Stoehr, J. (2017). A review on statistical inference methods for discrete markov random fields.

Strebelle, S. (2002). Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology*, 34(1):1–21.

Talarico, E., Leäo, W., and Grana, D. (2019). Comparison of recursive neural network and markov chain models in facies inversion. 2019(1):1–5.

Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics.

Titterington, D., Titterington, P., Smith, A., and Makov, U. (1985). *Statistical Analysis of Finite Mixture Distributions*. Applied section. Wiley.

Tjelmeland, H. and Besag, J. (1998). Markov random fields with higher-order interactions. *Scandinavian Journal of Statistics*, 25(3):415–433.

Trias, M., Vecchio, A., and Veitch, J. (2009). Delayed rejection schemes for efficient markov-chain monte-carlo sampling of multimodal distributions.

Ulvmoen, M. and Omre, H. (2010). Improved resolution in Baeysian lithology/fluid inversion from seismic prestack data and well observations: Part I - Methodology. *Geophysics*, 75:R21–R35.

Paper I

Bayesian Gaussian Mixture Linear Inversion for Geophysical Inverse Problems

Dario Grana, Torstein Fjeldstad and Henning Omre

Mathematical Geosciences vol. 49 (4) (2017)

Paper II

Bayesian Inversion of Convolved Hidden Markov Models With Applications in Reservoir Prediction

Torstein Fjeldstad and Henning Omre

IEEE Transactions on Geoscience and Remote Sensing,

vol. 58, no. 3, pp. 1957-1968, March 2020

Paper III

Joint probabilistic petrophysics-seismic inversion based on Gaussian mixture and Markov chain prior models

---

Torstein Fjeldstad and Dario Grana

Geophysics vol. 83 (1) (2018)

GEOPHYSICS, VOL. 83, NO. 1 (JANUARY-FEBRUARY 2018); P. R31–R42, 12 FIGS., 2 TABLES.
10.1190/GEO2017-0239.1

Joint probabilistic petrophysics-seismic inversion based on Gaussian mixture and Markov chain prior models

Torstein Fjeldstad[1] and Dario Grana[2]

ABSTRACT

Seismic reservoir characterization focuses on the prediction of reservoir properties based on the available geophysical and petrophysical data. The inverse problem generally includes continuous properties, such as petrophysical and elastic attributes, and discrete properties, such as lithology/fluid classes. We have developed a joint probabilistic inversion methodology for the prediction of petrophysical and elastic properties and lithology/fluid classes that combined statistical rock physics and Bayesian seismic inversion. The elastic attributes depend on continuous petrophysical variables, such as porosity and clay content, and discrete lithology/fluid classes, through a nonlinear rock-physics relationship together. The seismic model relates the elastic attributes, such as velocities and density, to their seismic response (reflectivity, traveltime, and amplitudes). The advantage of our integrated approach is that the inversion method accounts for the uncertainty associated to each step of the modeling workflow. The lithology/fluid classes are assigned by a Markov random field prior model to capture vertical continuity and vertical sorting of the lithology/fluid classes. Because rock and fluid properties are in general not Gaussian, a spatially coupled Gaussian mixture prior model based on the lithology/fluid classes is constructed. The forward geophysical operator includes a lithology-/fluid-dependent rock physics model and a linearized seismic model based on the convolution of the seismic wavelet with the reflectivity coefficient series. The solution of the inverse problem consists of the posterior distributions of petrophysical and elastic properties and lithology/fluid classes. We proposed an efficient Markov chain Monte Carlo algorithm to sample from the posterior models and assess the uncertainty. Our methodology is demonstrated on a seismic cross section from a survey in the Norwegian Sea, and it shows promising results consistent with well-log data measured at the well location as well as reliable prediction uncertainties.

INTRODUCTION

In reservoir characterization, the prediction of hydrocarbon presence in the subsurface given geophysical observations is a problem of utmost importance. Geophysical observations may include data from seismic surveys and exploration wells. Quantitative interpretation (Avseth et al., 2005) is an important tool in oil and gas recovery to predict reservoir properties, evaluate the associated uncertainty, and build models for fluid flow simulations. We focus on the lithology/fluids classification, and the predictions of rock properties and elastic attributes. Indeed, the assessment of the uncertainty is important to reduce risk and increase the expected revenue.

Petrophysics models describe the relations between rock and fluid properties and measured petrophysical well logs, whereas rock-physics models link the rock and fluid properties to the elastic attributes (Mavko et al., 2009). Petrophysics and rock-physics relations might be different in different lithology/fluid classes. Common petrophysics and rock-physics models can be either theoretical models (Mavko et al., 2009) or empirical models calibrated from well log measurements. These models are generally nonlinear; therefore, the classic linear inverse theory cannot be applied. However, many of these models can be linearized within a small range of the model properties, e.g., using Taylor's expansion (Grana, 2016).

A common geophysical data type in reservoir characterization is amplitude variation with offset (AVO) observations. The seismic response can be generally approximated using a convolutional model of the source wavelet with the vector of reflection coefficients at the interfaces. The reflection coefficients are a function of the elastic

Manuscript received by the Editor 21 April 2017; revised manuscript received 9 July 2017; published ahead of production 25 October 2017; published online 15 December 2017.
[1]NTNU – Norwegian University of Technology and Science, Department of Mathematical Sciences, Faculty of Information Technology and Electrical Engineering, Trondheim, Norway. E-mail: torstein.fjeldstad@ntnu.no.
[2]University of Wyoming, Department of Geology and Geophysics, School of Energy Resources, Laramie, Wyoming, USA. E-mail: dgrana@uwyo.edu.

properties of the rock, i.e., velocities and density, and for weak elastic contrasts, they can be computed using linear approximations (Aki and Richards, 1980).

In this paper, we focus on the assessment of the spatial distribution of lithology/fluids, rock and fluid properties (such as porosity), and elastic attributes (such as acoustic impedance), conditioned by the measured seismic data. Mathematically, this problem can be considered as three distinct inverse problems; however, they are strongly related. We will refer to the inverse problems as, respectively, lithology/fluid, petrophysical, and seismic inversion. In the first problem, the model variable to be predicted is a discrete property representing an indicator for each of the lithology/fluids classes. In the second problem, it is a set of continuous reservoir properties bounded between zero and one, and in the third problem, it is a set of continuous nonnegative elastic properties. Assessment of these inverse problems is a major challenge in reservoir characterization because they are often ill-posed (Tarantola, 2005). We operate in a Bayesian framework, which is a popular choice in reservoir characterization for property prediction and uncertainty quantification (Scales and Tenorio, 2000; Ulrych et al., 2001). The solution of the inverse problem is represented by a probability density function, not just a single predicted value. As discussed in Houck (2002), it is important to combine rock physics and seismic uncertainties in quantitative interpretation to obtain posterior densities with realistic uncertainty. For deterministic methods, we refer to Aster et al. (2005), and the references therein.

Buland and Omre (2003) propose a Bayesian framework for seismic inversion, in which they focused on the prediction of elastic attributes. They assumed a Gaussian prior model and linearized seismic forward model together with Gaussian error terms; thus, their posterior density can be assessed analytically. Hansen et al. (2006) propose a sequential simulation approach for the seismic inversion problem. Larsen et al. (2006) and Ulvmoen and Omre (2010) consider lithology/fluid predictions based on a Markov chain prior model for the lithology/fluids classes inspired by Eidsvik et al. (2004). Rimstad and Omre (2013) later formalize this to a convolved hidden Markov model. Gunning and Glinsky (2007), Grana and Della Rossa (2010), and Rimstad and Omre (2010) also assess the rock properties. Jullum and Kolbjørnsen (2016) propose a series of local Gaussian approximations of the likelihood function to assess the posterior model for rock properties and elastic attributes. Connolly and Hughes (2016) propose an acceptance/rejection sampling technique in which a set of lithology/fluids classes is generated from the prior model together with a set of synthetic observations and is updated until convergence. The lithology/fluids classes are accepted based on a misfit criterion of the corresponding synthetic observations and the observed seismic observations. We refer to González et al. (2008) and Bosch et al. (2010) for a discussion of geostatistical methods applied to seismic inversion.

For a non-Gaussian prior and likelihood model, Markov chain Monte Carlo (MCMC) methods (Robert and Casella, 2005) are often required because analytical expressions for the posterior model do not exist except for very particular cases. In MCMC methods, a sample of the variables of interest is generated from a proposal density and accepted or rejected based on the likelihood of observing the measured data given the sampled variable. Unfortunately, the assessment of geophysical inverse problems is in general not feasible by straightforward single-site MCMC methods because of the spatial coupling in the likelihood function; i.e., designing a suitable proposal density in the MCMC algorithm is a challenging problem.

The inverse problem is also challenging for the presence of the discrete and continuous properties. Rock and fluid properties, such as porosity and water saturation, are in general not Gaussian, but appear as multimodal and skewed due to the presence of different lithology/fluids classes (Grana and Della Rossa, 2010; Dubreuil-Boisclair et al., 2012; de Figueiredo et al., 2017). Grana and Della Rossa (2010) propose a Gaussian mixture model, in which the rock properties and elastic attributes are Gaussian conditional on the lithology/fluids classes; i.e., each lithology/fluid class is detectable as a single component of the Gaussian mixture. Hidden Markov models (Cappé et al., 2005) have been applied to seismic inversion (Larsen et al., 2006) to honor the vertical geologic ordering of lithologic facies and fluids (Krumbein and Dacey, 1969). A Markov random field prior model (Besag, 1974) for geologic sequences is presented in Ulvmoen and Omre (2010) and Rimstad and Omre (2010) to account for the lateral continuity of the subsurface. These methods are based on Gaussian approximations of Gaussian mixture densities for the rock and fluid properties, together with the efficient recursive forward-backward algorithm (Reeves and Pettitt, 2004) to assess an approximate posterior model for the lithology/fluids classes given the geophysical observations. Rimstad and Omre (2013) and Fjeldstad and Omre (2017) propose various approximate posterior models and empirically evaluated their similarities with the correct posterior model.

Our method provides a rigorous mathematical formulation for the estimation of discrete (facies) and continuous (rock properties and elastic attributes) variables including the spatial coupling (vertical and lateral) of the properties and avoids simplistic assumptions related to the Gaussian distribution of the model parameters. In this work, we define Gaussian mixture prior models for the rock properties and elastic attributes inspired by Grana et al. (2017). The mixing weights are defined according to a first-order Markov chain prior model for the lithology/fluids classes. Combined with Gaussian likelihood functions for the forward geophysical operators, the posterior models of rock properties and elastic attributes are Gaussian mixture densities. We present an efficient MCMC algorithm for assessing the correct posterior models for the lithology/fluids classes, rock properties, and elastic attributes in a Bayesian framework, and we discuss the joint assessment. A case study from the Norwegian Sea based on a seismic cross section and well logs is presented.

MODEL SPECIFICATION

The focus of this work is on the joint assessment of petrophysical and elastic properties such as lithology/fluid classes, porosity, and P-wave velocity from seismic data, often referred to as petrophysical and seismic inversion. Indeed, these problems can be formulated as separate inverse problems (Grana et al., 2017) or as a joint inverse problem (Rimstad and Omre, 2010). We consider the latter and distinguish between the continuous-valued inverse problems (prediction of rock and fluid properties and elastic properties) and the categorical inverse problem (lithology/fluid classification). In the following section, we discuss these inverse problems and specify the probabilistic model of interest.

We present the methodology for a 1D target zone of interest, i.e., a profile through a reservoir unit discretized onto a regular grid $\mathcal{L}_T = \{1, \ldots, T\}$. We denote by $\mathbf{r} \in \mathbb{R}^T$ the continuous-valued

rock properties, e.g., porosity ϕ, and by $\mathbf{m} \in \mathbb{R}^T$ the elastic properties, e.g., P-wave velocity V_P. The seismic signal $\mathbf{d}$ is assumed to be observed on $\mathcal{L}_T$; however, our methodology extends to cases where $\mathbf{r}$, $\mathbf{m}$, and $\mathbf{d}$ are of different dimensions. For example, the seismic signal $\mathbf{d}$ could be defined on a coarser grid than the model variables $\mathbf{r}$ and $\mathbf{m}$.

The objective is to assess the posterior models $p(\mathbf{r}|\mathbf{d})$ and $p(\mathbf{m}|\mathbf{d})$ in a Bayesian framework. We define the continuous-valued petrophysical and seismic inverse problems according to Bayes' theorem:

$$\begin{aligned} p(\mathbf{r}|\mathbf{d}) &= \text{const}_{\mathbf{r}} \times p(\mathbf{d}|\mathbf{r})p(\mathbf{r}) \\ p(\mathbf{m}|\mathbf{d}) &= \text{const}_{\mathbf{m}} \times p(\mathbf{d}|\mathbf{m})p(\mathbf{m}), \end{aligned} \tag{1}$$

where $\text{const}_{\mathbf{r}}$ and $\text{const}_{\mathbf{m}}$ are the normalizing constants, $p(\mathbf{d}|\mathbf{r})$ and $p(\mathbf{d}|\mathbf{m})$ are the likelihood functions, and $p(\mathbf{r})$ and $p(\mathbf{m})$ are the prior models. In practice, the analytical assessment of the posterior models in equation 1 is only feasible for simpler models; i.e., the posterior models are often assessed by MCMC sampling in real case applications (Mukerji et al., 2001).

Let the lithology/fluids classes be labeled by $\boldsymbol{\kappa} \in \Omega_{\boldsymbol{\kappa}} = \Omega_{\kappa}^T$, where $\kappa_t \in \Omega_{\kappa} = \{1, \ldots, L\}$ is an indicator variable taking on exactly one of the L values for each $t \in \mathcal{L}_T$. We also predict the lithology/fluids classes $\boldsymbol{\kappa}$ given the seismic signal $\mathbf{d}$, being a categorical inverse problem:

$$p(\boldsymbol{\kappa}|\mathbf{d}) = \text{const}_{\boldsymbol{\kappa}} \times p(\mathbf{d}|\boldsymbol{\kappa})p(\boldsymbol{\kappa}), \tag{2}$$

where $\text{const}_{\boldsymbol{\kappa}}$ is the normalizing constant, $p(\mathbf{d}|\boldsymbol{\kappa})$ is the likelihood function, and $p(\boldsymbol{\kappa})$ is the prior model.

Seismic data $\mathbf{d}$ are approximated as a convolution of the reflection coefficients of $\mathbf{m}$ due to the propagation of seismic waves. The reflection coefficients are given by the nonlinear Zoeppritz equation (Sheriff and Geldart, 1995). The lithology/fluid classes $\boldsymbol{\kappa}$ and the rock and fluid properties $\mathbf{r}$ affect the velocity of seismic waves $\mathbf{m}$ propagating through the subsurface through a nonlinear relationship, usually referred to as the rock-physics model (Avseth et al., 2005). Such models include, among the others, Dvorkin's cemented sand model, Raymer's equation, and the Kuster-Toksoz models, and they are in general known for conventional reservoirs (Mavko et al., 2009). Indeed, the rock and fluid properties are dependent on the lithology/fluids classes themselves, which we denote by the petrophysics model. Figure 1 displays the conditional independence structure defined by the geophysical models. Indeed, $\mathbf{m}$ is a canonical variable for $\mathbf{d}$; i.e., $(\boldsymbol{\kappa}, \mathbf{r})$ and $\mathbf{d}$ are conditionally independent given $\mathbf{m}$. We observe that the effect of $\boldsymbol{\kappa}$ propagates directly into $\mathbf{m}$ and indirectly through the effect of the rock properties $\mathbf{r}$, as desired.

We focus on the assessment of equation 2 and discuss how it relates to equation 1. Because $\mathbf{m}$ is a canonical variable for $\mathbf{d}$ and is related to the seismic model, we present the seismic likelihood model $p(\mathbf{d}|\mathbf{m})$ first. Afterward, we discuss the various prior models $p(\mathbf{r})$, $p(\mathbf{m})$, and $p(\boldsymbol{\kappa})$. Finally, we discuss the posterior models $p(\boldsymbol{\kappa}|\mathbf{d})$, $p(\mathbf{r}|\mathbf{d})$, and $p(\mathbf{m}|\mathbf{d})$.

Likelihood model

We consider the class of so-called Gauss-linear likelihood functions, i.e., likelihood models that are linear in the conditioning variable together with additive Gaussian errors.

The observed seismic signal for $t \in \mathcal{L}_T$ is approximated as a convolution of the reflection coefficients and a wavelet for each incident angle. The exact reflection coefficients are given by the nonlinear Zoeppritz equation; however, we consider a linearized approximation based on the Aki-Richards formulation for weak contrasts (Aki and Richards, 1980). We define the seismic likelihood model as

$$[\mathbf{d}|\mathbf{m}] = \mathbf{W}\mathbf{m} + \boldsymbol{\epsilon}_{\mathbf{d}|\mathbf{m}}, \tag{3}$$

where $\mathbf{W} \in \mathbb{R}^{T \times T}$ is a matrix including the convolution and Aki-Richards linearized approximation, and $\boldsymbol{\epsilon}_{\mathbf{d}|\mathbf{m}}$ is a zero mean Gaussian error term having covariance matrix $\boldsymbol{\Sigma}_{\mathbf{d}|\mathbf{m}} \in \mathbb{R}^{T \times T}$. Thus, the seismic likelihood function $p(\mathbf{d}|\mathbf{m})$ is a Gaussian density. As in Buland and Omre (2003), we write $\mathbf{W}$ as a matrix product: $\mathbf{W} = \mathbf{CAD}$, where $\mathbf{C}$ is the convolution matrix, $\mathbf{A}$ is the matrix with weak contrast Aki-Richards reflection coefficients, and $\mathbf{D}$ is a first-order differential matrix.

Prior models

Continuous-valued rock and fluid properties $\mathbf{r}$, such as porosity and clay volume, appear in general as skewed and multimodal due to the presence of various lithology/fluids classes (Grana and Della Rossa, 2010); i.e., the conditional mean value for porosity given sandstone is usually larger than the conditional mean value given shale. Buland and Omre (2003) consider a Gaussian prior model that is unimodal and symmetric, in their Bayesian linearized elastic inversion methodology. However, because the rock properties $\mathbf{r}$ of interest are usually skewed and multimodal, we seek a prior model $p(\mathbf{r})$ that honors these characteristics. Because rock properties are constrained to $[0, 1]$, we discuss later a one-to-one transformation from $[0, 1]$ to $\mathbb{R}$.

We define the following lithology/fluid dependent petrophysical model:

$$[\mathbf{r}|\boldsymbol{\kappa}] = \boldsymbol{\mu}_{\mathbf{r}|\boldsymbol{\kappa}} + \boldsymbol{\epsilon}_{\mathbf{r}|\boldsymbol{\kappa}}, \tag{4}$$

where $\boldsymbol{\mu}_{\mathbf{r}|\boldsymbol{\kappa}} = (\mu_{r_1|\kappa_1}, \ldots, \mu_{r_T|\kappa_T})^{\intercal}$ is a vector of length T including the pointwise conditional mean values for the rock properties switching according to the lithology/fluids classes. We assume $\boldsymbol{\epsilon}_{\mathbf{r}|\boldsymbol{\kappa}} \in \mathbb{R}^T$ to be a centered Gaussian error term having the covariance matrix $\boldsymbol{\Sigma}_{\mathbf{r}|\boldsymbol{\kappa}} \in \mathbb{R}^{T \times T}$. Due to the geologic continuity, the rock and fluid properties are observed to be spatially correlated; therefore, we assume $\boldsymbol{\Sigma}_{\mathbf{r}|\boldsymbol{\kappa}}$ to be defined from a vertical spatial correlation function $\rho_{\mathbf{r}}(h)$ for $h = 0, \ldots, T-1$ and the set of marginal standard deviations $\sigma_{r_t|\kappa_t}$ for all $t \in \mathcal{L}_T$ (Grana et al., 2017). Thus,

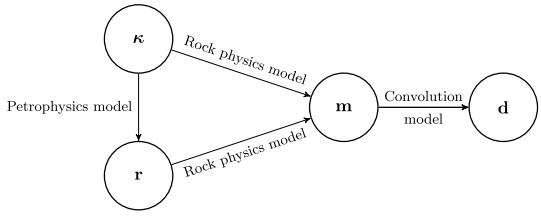


Figure 1. Directed graph of conditional independence structure based on geophysical models.

$p(\mathbf{r}|\boldsymbol{\kappa})$ is a Gaussian density, which we denote as the petrophysics model.

Because the petrophysics model is Gaussian, the multivariate marginal density

$$p(\mathbf{r}) = \sum_{\boldsymbol{\kappa}} p(\mathbf{r}|\boldsymbol{\kappa}) p(\boldsymbol{\kappa}) \tag{5}$$

is a Gaussian mixture prior density. Thus, we have constructed a multimodal and skewed prior for the rock properties. Indeed, $\mathbf{r}$ is a mixture of Gaussian variables with at most L^T unique modes, where the mixing proportions are defined by the lithology/fluids classes' prior model $p(\boldsymbol{\kappa})$. Note that $p(\mathbf{r}|\boldsymbol{\kappa})$ is a Gaussian density, and $\boldsymbol{\kappa}$ can be thought of as an indicator for a set of Gaussian densities indexed by $\boldsymbol{\kappa}$. Because there are L^T unique configurations of $\boldsymbol{\kappa}$, evaluating $p(\mathbf{r})$ is computationally unfeasible even for small L and T. The marginal densities $p(r_t)$ are univariate Gaussian mixture densities for all $t \in \mathcal{L}_T$.

The lithology/fluids classes $\boldsymbol{\kappa}$, and rock and fluid properties $\mathbf{r}$ affect the seismic velocities $\mathbf{m}$ propagating along a vertical profile. For a given set of lithology/fluids classes $\boldsymbol{\kappa}$ and rock properties $\mathbf{r}$, we define the following rock-physics model:

$$\mathbf{m} = \mathbf{f}(\mathbf{r}, \boldsymbol{\kappa}) + \boldsymbol{\epsilon}_{\mathbf{m}}, \tag{6}$$

where $\mathbf{f}$ is the rock-physics model and $\boldsymbol{\epsilon}_{\mathbf{m}} \in \mathbb{R}^T$ is an error term. In general, $\mathbf{f}$ is a complex nonlinear function that allows for computing compressional and shear velocities when the porosities, lithology/fluids, and fluid contents are known. These models are often nonlinear but can be approximated by a first-order Taylor approximation or a linear empirical relation, at least locally.

We define a linearized version of equation 6:

$$[\mathbf{m}|\mathbf{r}, \boldsymbol{\kappa}] = \boldsymbol{\mu}_{\mathbf{m}|\boldsymbol{\kappa}} + \mathbf{B}_{\boldsymbol{\kappa}}(\mathbf{r} - \boldsymbol{\mu}_{\mathbf{r}|\boldsymbol{\kappa}}) + \boldsymbol{\epsilon}_{\mathbf{m}|\boldsymbol{\kappa}}, \tag{7}$$

where $\boldsymbol{\mu}_{\mathbf{m}|\boldsymbol{\kappa}} = (\mu_{m_1|\kappa_1}, \ldots, \mu_{m_T|\kappa_T})^{\intercal}$ is the vector of length T with the pointwise conditional mean values dependent on the lithology/fluids $\boldsymbol{\kappa}$. The matrix with regression coefficients $\mathbf{B}_{\boldsymbol{\kappa}} \in \mathbb{R}^{T\times T}$ is assumed to be dependent on $\boldsymbol{\kappa}$, and for simplicity and to avoid overfitting, we assume $\mathbf{B}_{\boldsymbol{\kappa}}$ to be diagonal having the regression coefficients on the diagonals. These coefficients are obtained by least-squares from a set of training data, e.g., from a well nearby. We assume $\boldsymbol{\epsilon}_{\mathbf{m}|\boldsymbol{\kappa}} \in \mathbb{R}^T$ to be a centered Gaussian error term having covariance matrix $\boldsymbol{\Sigma}_{\mathbf{m}|\boldsymbol{\kappa}} \in \mathbb{R}^{T\times T}$ dependent on a vertical spatial correlation function $\rho_{\mathbf{m}}(h)$ and the set of conditional standard deviations $\sigma_{m_t|\kappa_t}$ for $t \in \mathcal{L}_T$, similar to $\boldsymbol{\Sigma}_{\mathbf{r}|\boldsymbol{\kappa}}$. Indeed, $p(\mathbf{m}|\mathbf{r}, \boldsymbol{\kappa})$ is a Gaussian density, and we refer to it as the rock-physics model. It follows that the multivariate marginal density for the $\mathbf{m}$:

$$p(\mathbf{m}) = \sum_{\boldsymbol{\kappa}} \left[\int p(\mathbf{m}|\mathbf{r}, \boldsymbol{\kappa}) p(\mathbf{r}|\boldsymbol{\kappa}) \mathrm{d}\mathbf{r} \times p(\boldsymbol{\kappa}) \right], \tag{8}$$

is a Gaussian mixture prior density because the inner integral is a Gaussian density.

If our linearization in equation 7 is not adequate, the conditional $p(\mathbf{m}|\mathbf{r}, \boldsymbol{\kappa})$ can be assessed by, e.g., a kernel density estimator (Silverman, 1986) or regression splines (Hastie et al., 2009). Such methods allow for a more flexible rock-physics model at the cost of a more complex prior model. Note that analytical expressions for the inner integral in equation 8 do not exist in this case, except for very particular cases. Therefore, we consider only Gaussian likelihood functions, and we refer to Jullum and Kolbjørnsen (2016) for more complex nonlinear rock-physics likelihood models.

We define a first-order stationary Markov chain prior model for the vector of lithology/fluids classes $\boldsymbol{\kappa}$:

$$p(\boldsymbol{\kappa}) = p(\kappa_1) \times \prod_{t=2}^{T} p(\kappa_t|\kappa_{t-1}), \tag{9}$$

inspired by Eidsvik et al. (2004). Here, $p(\kappa_1)$ is the stationary distribution and the prior marginal probability for each class. For $t \in \mathcal{L}_T$, we assume each κ_t to take on exactly one of L different categories, e.g., shale, gas sandstone, or brine sandstone. A Markov chain prior model is well-suited for geologic constraints, e.g., gravitational sorting of fluids (Krumbein and Dacey, 1969). We consider only a stationary first-order Markov chain; however, the methodology extends to kth order nonstationary Markov chains at essentially no additional computational cost. Indeed, a higher order Markov chain prior model can enforce specific layer constraints such as minimum thickness and sorting in the posterior model.

Becaue the prior model $p(\boldsymbol{\kappa})$ is included in the definition of $p(\mathbf{r})$ and $p(\mathbf{m})$ (equation 5 and equation 8), we focus on the assessment of the posterior model $p(\boldsymbol{\kappa}|\mathbf{d})$. We define the gross likelihood:

$$p(\mathbf{d}|\boldsymbol{\kappa}) = \int p(\mathbf{d}|\mathbf{m}) p(\mathbf{m}|\mathbf{r}, \boldsymbol{\kappa}) p(\mathbf{r}|\boldsymbol{\kappa}) \mathrm{d}(\mathbf{r}, \mathbf{m}), \tag{10}$$

where $p(\mathbf{d}|\mathbf{m})$, $p(\mathbf{m}|\mathbf{r}, \boldsymbol{\kappa})$, and $p(\mathbf{r}|\boldsymbol{\kappa})$ are, respectively, the seismic, rock physics, and petrophysics models. Indeed, by construction, $p(\mathbf{d}|\boldsymbol{\kappa})$ is a Gaussian density, and it can be assessed analytically for a given $\boldsymbol{\kappa}$. Note that the covariance matrix in $p(\mathbf{d}|\boldsymbol{\kappa})$ is dependent on $\boldsymbol{\kappa}$; thus, we have to compute its inverse for each $\boldsymbol{\kappa}$, which is computationally expensive.

Posterior models

Recall that the posterior models of interest are $p(\mathbf{r}|\mathbf{d})$, $p(\mathbf{m}|\mathbf{d})$, and $p(\boldsymbol{\kappa}|\mathbf{d})$, as given in equation 1 and equation 2. Combining equation 2 and equation 10, it can be verified that the normalizing constant in $p(\boldsymbol{\kappa}|\mathbf{d})$ is

$$\mathrm{const}_{\boldsymbol{\kappa}} = \left\{ \sum\nolimits_{\boldsymbol{\kappa}} \int p(\mathbf{d}|\mathbf{m}) p(\mathbf{m}|\mathbf{r}, \boldsymbol{\kappa}) p(\mathbf{r}|\boldsymbol{\kappa}) \mathrm{d}(\mathbf{r}, \mathbf{m}) \times p(\boldsymbol{\kappa}) \right\}^{-1}, \tag{11}$$

and that the posterior $p(\boldsymbol{\kappa}|\mathbf{d})$ is a nonhomogeneous Markov chain with long-ranged vertical spatial dependence. Straightforward assessment of the posterior model $p(\boldsymbol{\kappa}|\mathbf{d})$ is computationally unfeasible because of the vertical spatial coupling in the gross likelihood. The posterior density $p(\boldsymbol{\kappa}|\mathbf{d})$ is assessed by MCMC sampling, and a set of B realizations $\boldsymbol{\kappa}^1, \ldots, \boldsymbol{\kappa}^B$ can be generated from $p(\boldsymbol{\kappa}|\mathbf{d})$. For $\kappa \in \{1, \ldots, L\}$, we define the marginal probability profiles:

$$\hat{\mathbf{p}}^{\kappa} = \left\{ \hat{p}_t^{\kappa} = \frac{1}{B} \sum_{b=1}^{B} \mathrm{I}(\kappa_t^b = \kappa); \quad t \in \mathcal{L}_T \right\} \tag{12}$$

and the marginal maximum posterior (MMAP) predictor:

$$\hat{\boldsymbol{\kappa}} = \{\hat{\kappa}_t = \underset{\kappa}{\operatorname{argmax}}\, \hat{p}_t^{\kappa}; \quad t \in \mathcal{L}_T\}. \tag{13}$$

As for linear transformations together with Gaussian errors of a Gaussian variables, it can be verified that any linear transformation together with Gaussian errors of a Gaussian mixture variable is also a Gaussian mixture variable (Grana et al., 2017). Extending Grana et al. (2017), it can be verified that the posterior densities for $\mathbf{r}$ and $\mathbf{m}$ given $\mathbf{d}$ are Gaussian mixture densities:

$$\begin{aligned} p(\mathbf{r}|\mathbf{d}) &= \sum_{\boldsymbol{\kappa}} p(\mathbf{r}|\mathbf{d}, \boldsymbol{\kappa}) p(\boldsymbol{\kappa}|\mathbf{d}) \\ p(\mathbf{m}|\mathbf{d}) &= \sum_{\boldsymbol{\kappa}} p(\mathbf{m}|\mathbf{d}, \boldsymbol{\kappa}) p(\boldsymbol{\kappa}|\mathbf{d}), \end{aligned} \tag{14}$$

where $p(\mathbf{r}|\mathbf{d}, \boldsymbol{\kappa})$ and $p(\mathbf{m}|\mathbf{d}, \boldsymbol{\kappa})$ are analytically tractable Gaussian densities. Note that the weights $p(\boldsymbol{\kappa}|\mathbf{d})$ are dependent on the data, being the posterior model $p(\boldsymbol{\kappa}|\mathbf{d})$. We note that $p(\mathbf{r}|\mathbf{d})$ and $p(\mathbf{m}|\mathbf{d})$ belong to the same class of densities as their prior models $p(\mathbf{r})$ and $p(\mathbf{m})$; i.e., $p(\mathbf{r})$ and $p(\mathbf{m})$ are conjugate prior models. It follows that $p(r_t|\mathbf{d})$ and $p(m_t|\mathbf{d})$ are univariate Gaussian mixture densities, and we define the following MMAP predictors:

$$\begin{aligned} \hat{\mathbf{r}} &= \{\hat{r}_t = \underset{r_t}{\operatorname{argmax}}\, p(r_t|\mathbf{d}); \quad t \in \mathcal{L}_T\} \\ \hat{\mathbf{m}} &= \{\hat{m}_t = \underset{m_t}{\operatorname{argmax}}\, p(m_t|\mathbf{d}); \quad t \in \mathcal{L}_T\}. \end{aligned} \tag{15}$$

The corresponding prediction intervals are obtained by a series of univariate optimizations.

ASSESSMENT OF POSTERIOR MODELS

Recall that the straightforward assessment of $p(\boldsymbol{\kappa}|\mathbf{d})$ is unfeasible and straightforward, single-site MCMC sampling is unfeasible due to the vertical spatial coupling in $p(\mathbf{d}|\boldsymbol{\kappa})$. We present a class of kth order approximate posterior models $p^{(k)}(\boldsymbol{\kappa}|\mathbf{d})$ for $k = 1, 2, \ldots$ extending the projection approximation proposed in Fjeldstad and Omre (2017). The idea is to replace the Tth order Markov chain model $p(\boldsymbol{\kappa}|\mathbf{d})$ with a kth order Markov chain model $p^{(k)}(\boldsymbol{\kappa}|\mathbf{d})$ where $k \ll T$. For a fixed k, the approximate posterior model $p^{(k)}(\boldsymbol{\kappa}|\mathbf{d})$ is used as proposal density in an independent proposal MCMC Metropolis-Hastings (MH) algorithm (Robert and Casella, 2005) to assess the correct posterior model $p(\boldsymbol{\kappa}|\mathbf{d})$.

Consider the subvectors $\boldsymbol{\kappa}_t^{(k)} = (\kappa_{t-k+1}, \ldots, \kappa_t)^{\intercal}$, $\mathbf{r}_t^{(k)} = (r_{t-k+1}, \ldots, r_t)^{\intercal}$, and $\mathbf{m}_t^{(k)} = (m_{t-k+1}, \ldots, m_t)^{\intercal}$ for $k = 1, 2, \ldots$ of length k.

Because zero probabilities in the prior model $p(\boldsymbol{\kappa})$ enforces zero probabilities in the posterior model $p(\boldsymbol{\kappa}|\mathbf{d})$, we approximate only the gross likelihood model $p(\mathbf{d}|\boldsymbol{\kappa})$. By a trivial extension of the sample space, we observe that the prior density can be rephrased as a kth order Markov chain:

$$p(\boldsymbol{\kappa}) = p^{(k)}(\boldsymbol{\kappa}_k^{(k)}) \prod_{t=k+1}^{T} p^{(k)}(\boldsymbol{\kappa}_t^{(k)}|\boldsymbol{\kappa}_{t-1}^{(k)}), \tag{16}$$

because the first $k-1$ elements in $\boldsymbol{\kappa}_t^{(k)}$ are a member of the conditioning set $\boldsymbol{\kappa}_{t-1}^{(k)}$.

We define a Gaussian approximation $p^*(\mathbf{m})$ to the Gaussian mixture density $p(\mathbf{m})$ by computing the theoretical mean vector and covariance matrix as in Fjeldstad and Omre (2017). Empirically calibrated Gaussian approximations to $p(\mathbf{m})$ are indeed valid; however, we experienced slightly lower acceptance rates in the MCMC algorithm using this approximation. Note that $p^*(\mathbf{m})$ is a valid prior model in the Buland and Omre (2003) approach. The joint approximate density

$$p^*(\mathbf{d}, \mathbf{m}) = p(\mathbf{d}|\mathbf{m}) p^*(\mathbf{m}) \tag{17}$$

is Gaussian, and it follows that the conditional densities $p^*(\mathbf{d}|\mathbf{m}_t^{(k)})$ for $t = k, \ldots, T$ are also Gaussian densities. The conditional densities $p^*(\mathbf{d}|\mathbf{m}_t^{(k)})$ serve as a kth order approximation of the seismic likelihood function.

Recall the definition of the petrophysics and rock-physics models. It can be verified that $p(\mathbf{r}_t^{(k)}|\boldsymbol{\kappa}_t^{(k)})$ and $p(\mathbf{m}_t^{(k)}|\mathbf{r}_t^{(k)}, \boldsymbol{\kappa}_t^{(k)})$ are Gaussian densities for $t = k, \ldots, T$. Thus,

$$\begin{aligned} p^*(\mathbf{d}, \mathbf{m}_t^{(k)}, \mathbf{r}_t^{(k)}|\boldsymbol{\kappa}_t^{(k)}) &= p^*(\mathbf{d}|\mathbf{m}_t^{(k)}) p(\mathbf{m}_t^{(k)}|\mathbf{r}_t^{(k)}, \boldsymbol{\kappa}_t^{(k)}) \\ &\quad \times p(\mathbf{r}_t^{(k)}|\boldsymbol{\kappa}_t^{(k)}) \end{aligned} \tag{18}$$

are Gaussian densities that are assessed analytically. We define the following kth order likelihood approximation for $t = k, \ldots, T$:

$$\begin{aligned} p^{(k)}(\mathbf{d}|\boldsymbol{\kappa}_t^{(k)}) = \Bigg[\int & p^*(\mathbf{d}|\mathbf{m}_t^{(k)}) p(\mathbf{m}_t^{(k)}|\mathbf{r}_t^{(k)}, \boldsymbol{\kappa}_t^{(k)}) \\ & \times p(\mathbf{r}_t^{(k)}|\boldsymbol{\kappa}_t^{(k)}) \mathrm{d}(\mathbf{m}_t^{(k)}, \mathbf{r}_t^{(k)}) \Bigg]^{1/k}, \end{aligned} \tag{19}$$

where the kth root ensures that each κ_t is used exactly once in the likelihood.

Combining equation 16 and equation 19, we obtain an approximate posterior model:

$$\begin{aligned} p^{(k)}(\boldsymbol{\kappa}|\mathbf{d}) = \mathrm{const}_{\boldsymbol{\kappa}} &\times p^{(k)}(\mathbf{d}|\boldsymbol{\kappa}_k^{(k)}) p(\boldsymbol{\kappa}_k^{(k)}) \\ &\times \prod_{t=k+1}^{T} p^{(k)}(\mathbf{d}|\boldsymbol{\kappa}_t^{(k)}) p(\boldsymbol{\kappa}_t^{(k)}|\boldsymbol{\kappa}_{t-1}^{(k)}), \end{aligned} \tag{20}$$

in factorial form. The approximate posterior model is assessed analytically in $\mathcal{O}((T-k+1) \times L^{k+1})$ operations by the recursive forward-backward algorithm (Reeves and Pettitt, 2004). Maximum posterior (MAP) predictors are obtained by the Viterbi algorithm (Viterbi, 1967) for the approximate posterior model.

We assess the correct posterior model $p(\boldsymbol{\kappa}|\mathbf{d})$ using the approximate posterior model $p^{(k)}(\boldsymbol{\kappa}|\mathbf{d})$ as the proposal density in an independent proposal MCMC MH algorithm. Note that the troublesome normalizing constant in equation 2 cancels in the acceptance probability α. Rimstad and Omre (2013) and Fjeldstad and Omre (2017) empirically verify in synthetic case studies that a higher order k results in a higher acceptance rate, at the cost of increased computational demands in the forward-backward algorithm.

We extend the proposed methodology, and consider a cross section $\mathcal{D}$ in two dimensions discretized onto a lattice $\mathcal{L}_{\mathcal{D}}$. The reservoir is discretized vertically in time domain onto $\mathcal{L}_T$, as defined earlier, and in the horizontal direction onto $\mathcal{L}_n = \{1, \ldots, n\}$; thus, $\mathcal{L}_{\mathcal{D}} = \mathcal{L}_n \times \mathcal{L}_T$. The variables of interest are $\boldsymbol{\kappa} = \{\boldsymbol{\kappa}_x : x \in \mathcal{L}_n\} =$

$\{\kappa_{x,t}: x \in \mathcal{L}_n, t \in \mathcal{L}_T\}$, $\mathbf{r} = \{\mathbf{r}_x : x \in \mathcal{L}_n\} = \{r_{x,t} : x \in \mathcal{L}_n, t \in \mathcal{L}_T\}$, and $\mathbf{m} = \{\mathbf{m}_x : x \in \mathcal{L}_n\} = \{m_{x,t} : x \in \mathcal{L}_n, t \in \mathcal{L}_T\}$, and the objective is to assess them given $\mathbf{d} = \{\mathbf{d}_x : x \in \mathcal{L}_n\} = \{d_{x,t} : x \in \mathcal{L}_n, t \in \mathcal{L}_T\}$.

We assume the gross likelihood function to be in factorial form horizontally, extending Ulvmoen and Omre (2010):

$$p(\mathbf{d}|\boldsymbol{\kappa}) = \prod_{x=1}^{n}\left[\int p(\mathbf{d}_x|\mathbf{m}_x)p(\mathbf{m}_x|\mathbf{r}_x,\boldsymbol{\kappa}_x)p(\mathbf{r}_x|\boldsymbol{\kappa}_x)\mathrm{d}(\mathbf{r}_x,\mathbf{m}_x)\right]. \tag{21}$$

Indeed, this is the extension of equation 10, where the observations are assumed to be collected independently in the horizontal direction.

We define a profile Markov random field prior model, defined by $p(\boldsymbol{\kappa}_x|\boldsymbol{\kappa}_{-x})$ where $\boldsymbol{\kappa}_{-x}$ is every trace except trace x, for the lithology/fluids classes, extending Ulvmoen and Omre (2010). By the Hammersley-Clifford theorem (Besag, 1974), the set of conditional prior densities fully specifies the prior model for the Markov random field. Because the characteristics of the lithology/fluids classes are a result of sedimentary processes in the cross section, the conditional prior for each trace should be nonstationary in the vertical direction. For a given trace x and depth t, we denote, respectively, the sedimentary and fluid neighborhood by $\delta_s(x)$ and $\delta_f(t)$, and the full neighborhood by $\delta(x,t)$.

We assume a nearest horizontal neighborhood; however, higher order horizontal neighborhoods are of course valid:

$$p(\boldsymbol{\kappa}_x|\boldsymbol{\kappa}_{-x}) = p(\kappa_{x,1}|\kappa_{y,s};(y,s) \in \delta(x,1)) \times \prod_{t=2}^{T} p(\kappa_{x,t}|\kappa_{x,t-1},\kappa_{y,s};(y,s) \in \delta(x,t)). \tag{22}$$

We refer to Rimstad and Omre (2010) for details of the parametrization. Gravitational sorting and ordering are controlled by the vertical transition matrix P, and lateral dependence structure in sedimentary direction for the lithologies and that in horizontal direction for the fluids are described by, respectively, β_l and β_f.

By combining equation 21 and equation 22, we obtain a posterior model on conditional form:

$$p(\boldsymbol{\kappa}_x|\boldsymbol{\kappa}_{-x},\mathbf{d}) = \mathrm{const}_{\boldsymbol{\kappa}} \times \int p(\mathbf{d}_x|\mathbf{m}_x)p(\mathbf{m}_x|\mathbf{r}_x,\boldsymbol{\kappa}_x)p(\mathbf{r}_x|\boldsymbol{\kappa}_x)\mathrm{d}(\mathbf{r}_x,\mathbf{m}_x) \times p(\boldsymbol{\kappa}_x|\boldsymbol{\kappa}_{-x}), \tag{23}$$

which we can efficiently assess by a block Gibbs algorithm (Algorithm 1). Note that we assess $p^{(k)}(\boldsymbol{\kappa}_x|\boldsymbol{\kappa}_{-x},\mathbf{d})$ by the recursive forward-backward algorithm at each iteration based on the current neighboring traces; i.e., the proposal density is adaptive with respect to its neighboring traces. Without adaptive proposal densities, we experienced unsatisfactory acceptance rates because the proposal density does not necessarily honor the continuity in the lithology/fluids classes. Indeed, the innermost loop can be parallelized because each trace x is only dependent on its two neighboring traces. Posterior models $p(\mathbf{r}|\mathbf{d})$ and $p(\mathbf{m}|\mathbf{d})$ are similarly assessed for the vertical profile of interest (equation 14).

Our methodology extends trivially to 3D by extending the neighborhood; however, to ease notation, it is only presented for a 2D cross section.

Algorithm 1. Block Gibbs MCMC algorithm.

1 Initialize $\boldsymbol{\kappa}^{(0)}$ with $p(\boldsymbol{\kappa}|\mathbf{d}) > 0$.
2 **for** $b = 1$ **to** B **do**
3 Set $\boldsymbol{\kappa}^{(b)} = \boldsymbol{\kappa}^{(b-1)}$.
4 **for** $x \in \mathcal{L}_n$ in random order **do**
5 Assess $p^{(k)}(\boldsymbol{\kappa}_x|\boldsymbol{\kappa}_{-x}^{(b)},\mathbf{d})$ by the forward-backward algorithm.
6 Propose $\tilde{\boldsymbol{\kappa}}_x \sim p^{(k)}(\boldsymbol{\kappa}_x|\boldsymbol{\kappa}_{-x}^{(b)},\mathbf{d})$.
7 Set $\boldsymbol{\kappa}_x^{(b)} = \tilde{\boldsymbol{\kappa}}_x$ with probability
8 $\alpha = \min\left\{1, \frac{p(\tilde{\boldsymbol{\kappa}}_x|\boldsymbol{\kappa}_{-x}^{(b)},\mathbf{d})p^{(k)}(\boldsymbol{\kappa}_x^{(b)}|\boldsymbol{\kappa}_{-x}^{(b)},\mathbf{d})}{p(\boldsymbol{\kappa}_x^{(b)}|\boldsymbol{\kappa}_{-x}^{(b)},\mathbf{d})p^{(k)}(\tilde{\boldsymbol{\kappa}}_x|\boldsymbol{\kappa}_{-x}^{(b)},\mathbf{d})}\right\}$.
9 **end**
10 **end**

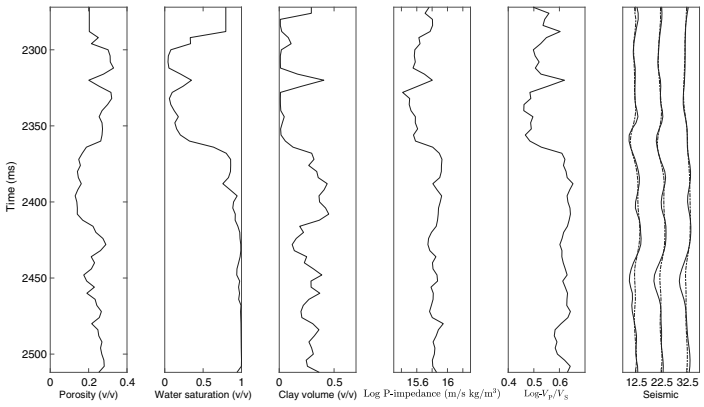


Figure 2. Observed petrophysical properties, elastic attributes, observed and synthetic seismic AVO at well location. From left to right: porosity, water saturation, clay volume, $\log(\rho V_P)$, $\log(V_P/V_S)$, and observed (solid lines) and synthetic (dashed lines) AVO observations.

CASE STUDY

We demonstrate our methodology on a 2D cross section from the Norwegian Sea. The reservoirs zones are characterized by a relatively high value of porosity, approximately 0.25–0.35. The interval of interest shows a stratigraphic sequence of sand and shale. Gas accumulated at the top of the structure in the high porosity sand layer, and it is mixed with a small percentage of irreducible water saturation. Porosity decreases in the lower part of the interval due to the increasing clay volume that reduces the volume of the pore space in the sandstone. The layers with lower clay content in the bottom part of the interval are filled by water. We refer to Avseth et al. (2016) for further

details of the target zone of interest. For this case study, a set of well logs and interpreted petrophysical parameters is available at the well location (Figure 2). The comparison between the measured seismic data and the synthetic AVO predictions computed from the well logs (Figure 2) shows an acceptable well-to-seismic tie, and some thin layers are not resolved, especially in the mid-section of the well-log interval. A 3D seismic survey was also acquired and was processed to derive three partial angle stacks.

We consider geophysical data from a seismic 2D survey including AVO observations based on near, mid, and far angles (Figure 3). The cross section includes one well, which we use for model validation, but not for conditioning. The region of interest is discretized onto a lattice $\mathcal{L}_{\mathcal{D}}$ and spans a domain of 61 vertical samples times 100 traces. The seismic sampling rate is 4 ms, and it starts at approximately 2390 m below sea level, corresponding to an initial traveltime of 2272 ms.

Three distinct lithology/fluids classes are of interest in the target zone, namely, shale, gas sandstone, and brine sandstone. We assume a first-order Markov chain prior model for the lithology/fluids $\boldsymbol{\kappa}$:

$$\mathbf{P} = \begin{pmatrix} 0.75 & 0.05 & 0.20 \\ 0.15 & 0.85 & 0 \\ 0.20 & 0.10 & 0.70 \end{pmatrix}, \tag{24}$$

and we note the prior zero-transition between the gas sandstone and brine sandstone due to gravitational sorting. The prior transition matrix P is calibrated from a well outside the target zone.

The continuous-valued rock properties of interest are porosity, water saturation, and clay volume: $\mathbf{r} = (\boldsymbol{\phi}, \mathbf{s}_w, \mathbf{c})^{\intercal}$, and the elastic properties of interest are the logarithm of P-impedance and the $V_{\mathrm{P}}/V_{\mathrm{S}}$ ratio: $\mathbf{m} = (\log(\boldsymbol{\rho}\mathbf{V_P}), \log(\mathbf{V_P}/\mathbf{V_S}))^{\intercal}$. The use of the logarithm is convenient for the model linearization as in Buland and Omre (2003). We assume $\boldsymbol{\epsilon}_{\mathbf{d}|\mathbf{m}}$ to be white noise and identical for all traces.

Because rock properties are constrained to $[0, 1]$, we apply a modified logit transform (Hastie et al., 2009) to the rock properties of interest to ensure support on $\mathbb{R}$; i.e., we define a Gaussian mixture prior model (equation 5) for $\mathrm{logit}(r)$. In Figure 4, we display crossplots of the empirical linear models $[\exp(m)|r]$ and $[m|\mathrm{logit}(r)]$ for upscaled data extracted from the reference well to justify a simplified linear rock-physics model $[m|\mathrm{logit}(r)]$ without lithology/fluid class dependency. The fitted linear models have adjusted R^2 values of 0.88 for $[\log(\rho V_{\mathrm{P}})|\mathrm{logit}(r)]$ and 0.76 for $[\log(V_{\mathrm{P}}/V_{\mathrm{S}})|r]$, which is satisfactory. The assumptions for a linear model appear to be satisfied. Therefore, the linear approximation in equation 7 appears to be acceptable because we expect the lithology/fluid class dependence in the rock-physics model to capture some of the nonlinearities in $[\log(\rho V_{\mathrm{P}})|\mathrm{logit}(r)]$ and $[\log(V_{\mathrm{P}}/V_{\mathrm{S}})|r]$. In other words, we expect $[m|\mathrm{logit}(r), \kappa]$ to capture some of the nonlinearities in the rock-physics model. The estimated marginal variances (equation 4 and equation 7) are increased by factors in the range of 1.5–5 for numerical stability.

A set of empirical rock-physics models $[m|\mathrm{logit}(r), \kappa]$ is calibrated in Figure 5 from the upscaled well logs outside the target zone. We display three subsections of the rock-physics models $[\log(\rho V_{\mathrm{P}})|\mathrm{logit}(\phi), \kappa]$ for three fixed values of (s_w, c). Each subplot includes the linear models $[\log(\rho V_{\mathrm{P}})|\mathrm{logit}(\phi), \kappa]$ for the three distinct lithology/fluid classes of interest displayed in colored lines. For example, the red line displays the linear model $[\log(\rho V_{\mathrm{P}})|\mathrm{logit}(\phi)]$ in which we have fixed the lithology/fluid class to gas sandstone. Note that the rock-physics models are hyperplanes defined everywhere; however, only a small subset of them is of geologic interest. The choices of (s_w, c) are such that the rock physics models are centered approximately on the mean values for, respectively, shale, gas sandstone, and brine sandstone in the respective subfigures. In the leftmost plot, the values of (s_w, c) are chosen such that shale is the most likely lithology/fluid class, gas sandstone second in the middle one, and brine sandstone in the rightmost one; i.e., even though all rock physics models are defined, they are not necessarily geologically meaningful.

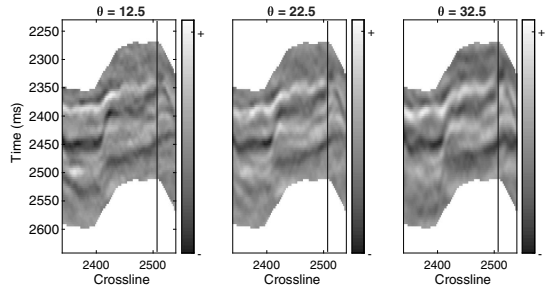


Figure 3. Two-dimensional seismic cross sections for near-, mid-, and far-angle stacks. Location of exploration well displayed by a solid black line.

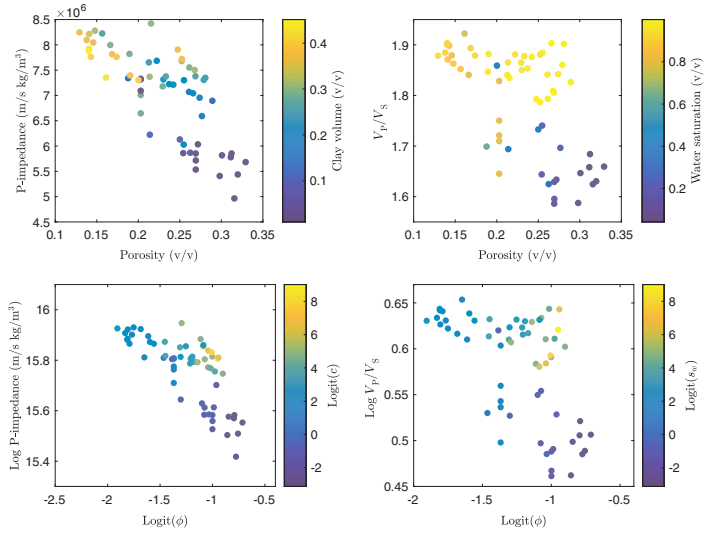


Figure 4. Cross plots of rock physics observations at well location. Top row: $[\rho V_{\mathrm{P}}|\phi, c]$ and $[V_{\mathrm{P}}/V_{\mathrm{S}}|\phi, s_w]$. Bottom row: $[\log(\rho V_{\mathrm{P}})|\mathrm{logit}(\phi, c)]$ and $[\log(V_{\mathrm{P}}/V_{\mathrm{S}})|\mathrm{logit}(\phi, s_w)]$.

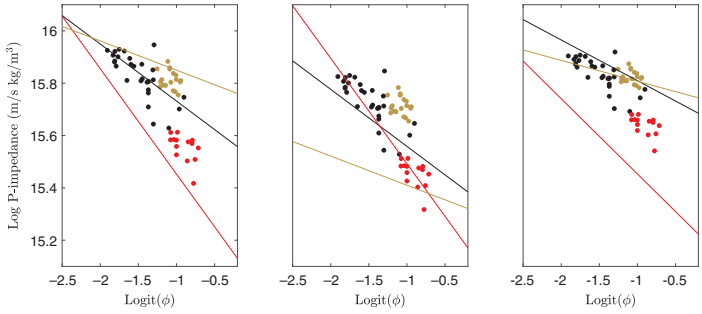


Figure 5. Subset of the fitted rock physics models $[m|r,\kappa]$. Each subplot includes the three linear models for $[\log(\rho V_P)|\text{logit}(\phi),\kappa]$ (solid lines) colored by the respective lithology/fluid classes (shale in black, gas sandstone in red, and brine sandstone in brown) for a fixed pair of (s_w, c). Observed $\log(\rho V_P)$ values at the well location are displayed and colored based on the reference lithology/fluid classification. From left to right, we have fixed $(s_w, c) \in \{(0.90, 0.50), (0.05, 0.01), (0.99, 0.31)\}$.

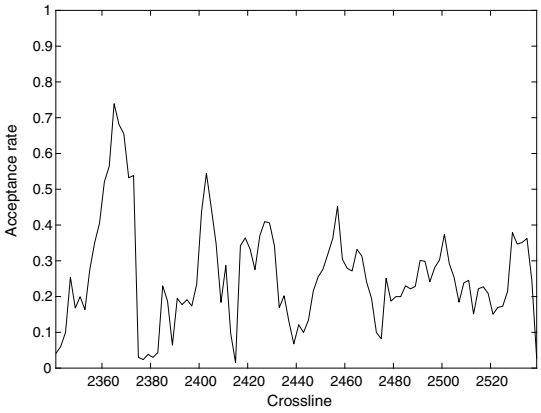


Figure 6. Acceptance rates in the MCMC algorithm at each trace location.

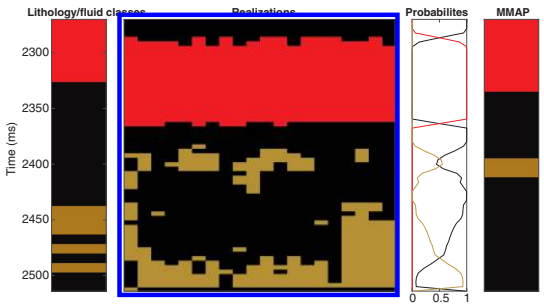


Figure 7. Posterior results at well location for the lithology/fluids classes. From left to right: geologically interpreted reference lithology/fluid classification, conditional realizations, predicted marginal probabilities, and MMAP predictor.

The defined rock physics models are assumed to be identical for all traces.

The horizontal coupling parameters β_f and β_l are set to, respectively, 1.5 and 0.75; i.e., the fluid couplings are chosen to be stronger in the sand than in the shale (Rimstad and Omre, 2010).

We generate in total 2,000,000 realizations from the posterior model $p(\boldsymbol{\kappa}|\mathbf{d})$, and we extract a subsample of every tenth realization to obtain approximately independent realizations. Mixing is observed to be satisfactory; however, it is not included here. We obtain a mean acceptance rate of 0.26 and a median acceptance rate of 0.24. The acceptance rates vary between 0.02 and 0.74 (Figure 6). Indeed, it is reasonable that the acceptance rate for each trace varies within the cross section because of possible misalignment in the seismic, strong local couplings, and poorly calibrated wavelets. There can also be lateral drift trends in the petrophysics and rock-physics models, which we have not included in our model. Also, the quality of our approximation, and thus the approximate posterior model $p^{(k)}(\boldsymbol{\kappa}_x|\mathbf{d})$, is strongly dependent on the position of each trace.

In Figure 7, posterior results for the lithology/fluids classes at the well location are displayed. We observe that the thin shale layer between the gas reservoirs is not fully captured in the posterior realizations; however, the uncertainty at the lowermost part is captured. We note that the MMAP predictor captures the main characteristics of the reference lithology/fluid classification, but the small-scale variability is lost, probably due to the low resolution and low signal-to-noise ratio of seismic compared with the actual well-log data. The mismatch between the predicted values and the observed values approximately 2450 ms is noted, which might be a result of a poorly calibrated wavelet (Figure 2).

Observed upscaled logs for rock and fluid properties and elastic attributes at the well location are displayed in Figure 8. Posterior mean and MMAP predictors are displayed, and they show similar main characteristics; however, the mean predictor tends to be slightly smoothed toward the local mean values.

In Figure 9, we display the fitted posterior densities for porosity at six randomly selected two-way times at the well location. At 2276 ms, we observe the posterior to be centered on the porosity value for shale and slightly skewed. At 2324 ms, the posterior density has been shifted to be centered on the porosity value for gas sandstone, and the density appears to be almost Gaussian. At the last four locations, the posterior are observed to be skewed and multimodal. At 2460 ms, we observe the MMAP predictor to predict the incorrect mode, whereas the mean predictor is located in between the two modes; however, at 2468, the MMAP predictor captures the correct mode resulting in a smaller pointwise prediction error than the mean predictor. In Table 1, we display the root-mean-square (rms) errors for the mean and MMAP predictors, and observe that the values based on the mean predictor are in general slightly smaller. This is true because the mean predictor is often in between the posterior modes, whereas the MMAP predictor can predict the incorrect mode, which results in a large pointwise discrepancy. Similar characteristics were observed also for water saturation, clay volume, $\log i_p$, and $\log V_P/V_S$. The coverage ratios for the 80% prediction intervals are also given in Table 1, and they appear to

be satisfactory, except for the coverage ratio of $\log(\rho V_P)$ which is smaller than expected.

In Figure 10, we display the posterior marginal probabilities for shale, gas sandstone, and brine sandstone together with the MMAP predictor for the cross section. We observe that we only predict gas sandstone in the uppermost part of the reservoir, and that the marginal probabilities appear with some lateral continuity in agreement with the expected geology of the field. Our methodology is observed to separate shale and gas sandstone with more distinct boundaries than shale and brine sandstone; however, the exact positions of the boundaries are uncertain. This is as we expected given the similar petroelastic features of shale and brine sandstone. In the MMAP predictor of the lithology/fluid classes, we observe larger bed thicknesses. We speculate that this feature could be due to the low resolution of the measured seismic data as well as to the high values of the lateral coupling parameters in the Markov random field, which might lead to regression toward the dominating class and loss of thin bed thicknesses. Such limitation could be overcome assuming a variable transition matrix P across the 2D section. For a discussion on the variability of the transition matrix, and consequently the variability in the facies proportions and average thicknesses, we refer to Connolly and Hughes (2016).

Figure 11 displays the MMAP predictors for porosity, water saturation, and clay volume. The MMAP predictors show some lateral and vertical continuity, and the gas layer is observed at the uppermost part of the target zone. The MMAP predictors show rapid transitions; thus, the predictor is fairly nonsmooth vertically compared with the predictor based on the pointwise posterior mean values. We display also the pointwise width of the 80% prediction intervals, and we observe the pointwise errors to be largest around the gas reservoir, as expected. Indeed, identifying the exact boundaries of the gas reservoirs is a challenging problem and is strongly dependent on the discretization of the grid.

Figure 12 displays the posterior results for the elastic attributes in a similar format as Figure 11. Both predictors for the elastic attributes share similarities with the predictors for the rock and fluid properties, such as lateral smoothness and disrupted vertical transitions.

DISCUSSION

Inspired by applications in lithology/fluid prediction, we have presented a convolved hidden Markov model extending Larsen et al. (2006) for a vertical section. The lithology/fluids classes are assumed to follow a Markov chain vertically to honor geologic constraints such as fluid sorting. We have defined a lithology/fluid-dependent Gaussian petrophysics model and a Gauss-linear lithology/fluid-dependent rock physics model. The linearity assumption is valid in many applications, with the exception of the porosity effect in unconsolidated sandstones and the fluid effect in homogeneous

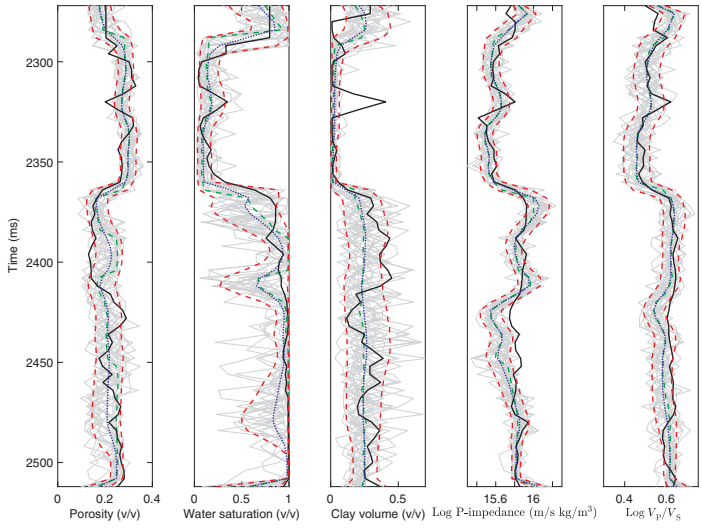


Figure 8. Posterior results at well location for the rock and fluid properties and elastic attributes. From left to right: observed well log (solid black), MMAP predictor (dashed green), mean predictor (dotted blue), 80% prediction interval (dashed red), and conditional realizations (gray) for porosity, water saturation, clay volume, $\log(\rho V_P)$, and $\log(V_P/V_S)$.

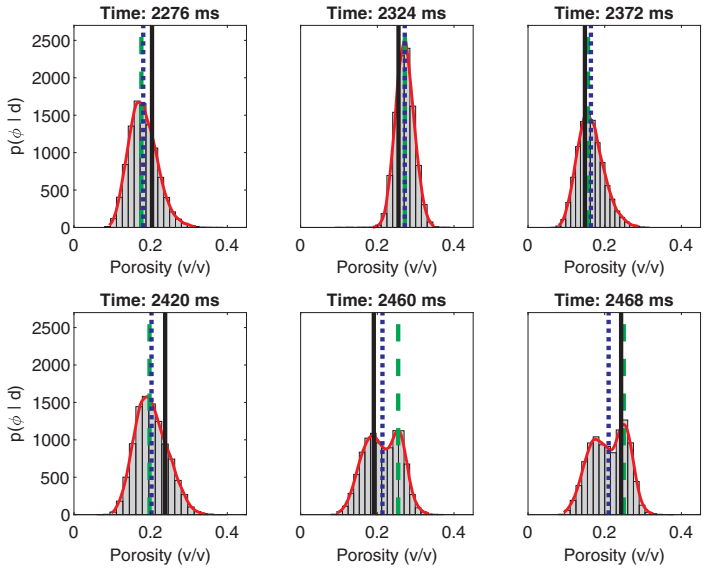


Figure 9. Fitted posterior densities $p(\phi_t|\mathbf{d})$ at the well location for six random locations. Histograms of posterior densities with theoretical posterior density together with well measurements indicated (black lines), MMAP predictors (dashed green lines), and posterior mean predictors (dotted blue lines).

Table 1. Root-mean-square errors at the well location for posterior mean and MMAP predictors, and coverage ratio for the 80% prediction intervals.

	Mean	MMAP	Coverage ratio (%)
ϕ	0.0403	0.0475	75.4
s_w	0.1279	0.1557	78.7
c	0.1044	0.1087	80.3
$\log(\rho V_P)$	0.1009	0.1017	67.2
$\log V_P/V_S$	0.0362	0.0354	77.1

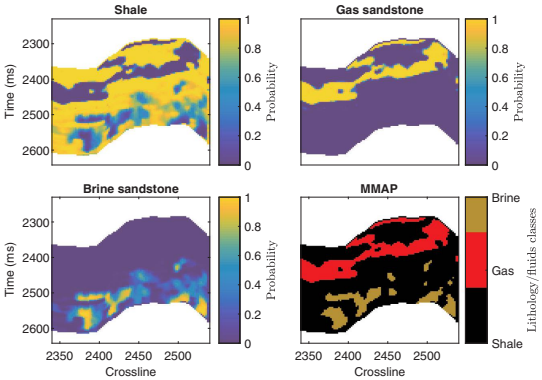


Figure 10. Posterior results lithology/fluids classes in 2D. Marginal probabilities for the various lithology/fluids classes and MMAP predictor for the cross section.

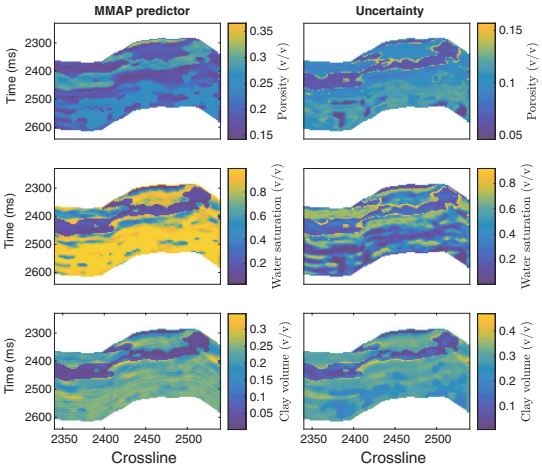


Figure 11. Posterior results rock and fluid properties in two dimensions. Left column: cross-section posterior MMAP predictors. Right column: cross-section pointwise posterior 80% prediction interval width.

mixtures (Dvorkin et al., 2014). Indeed, the fluid effect in homogeneous fluid mixtures and the porosity effect in unconsolidated rocks can introduce nonlinear effects in the model, resulting in unsatisfactory local linearizations. Seismic AVO data are approximated as a convolution of the reflection coefficients with additive Gaussian errors. The advantage of our method is the joint assessment of lithology/fluids classes, rock and fluid properties, and elastic attributes given the seismic AVO data.

In the vertical direction, we define a first-order Markov chain prior model for the lithology/fluids classes. Spatially coupled Gaussian mixture prior densities are defined for the rock and fluid properties and the elastic attributes, and they are verified to be conjugate prior models for Gauss-linear likelihood models. For a detailed comparison between Gaussian and Gaussian mixture prior models, we refer to Grana et al. (2017).

The fitted rock physics model is a first-order linear approximation. The methodology could be extended to include nonlinear rock physics models, such as granular media or inclusion models (Mavko et al., 2009). If the rock physics model is close to being linear, e.g., the stiff sand model or the Kuster-Toksoz model (Grana, 2016), the model could be linearized using Taylor's series approximation. If the model is nonlinear, such as the soft sand model or saturation models for homogeneous fluid mixtures, our method is still applicable, but the assessment of the posterior distribution requires the evaluation of a nonlinear likelihood function. Note that rock and fluid properties and the elastic attributes are no longer conjugate prior models for a nonlinear likelihood function.

We propose a kth order likelihood approximation to obtain an approximate posterior model in factorial form which is efficiently assessed by the forward-backward algorithm. The correct posterior density is assessed using the approximate posterior model as a proposal density in a MCMC MH algorithm. Note that the realizations are sampled conditional on the observations — not from a prior model as in most of the stochastic sampling techniques. These methods tend to underpredict the uncertainty (González et al., 2008; Grana et al., 2012).

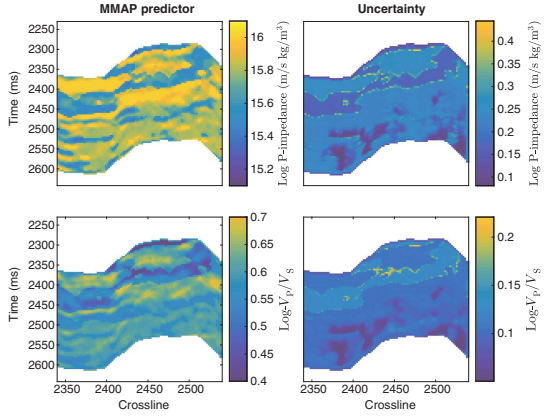


Figure 12. Posterior results elastic attributes in two dimensions. Left column: cross-section posterior MMAP predictors. Right column: cross-section pointwise posterior 80% prediction interval width.

Our methodology is extended to 2D by assuming a Markov random field prior model for the lithology/fluids classes (Ulvmoen and Omre, 2010). The correct posterior model is assessed by a block-Gibbs MCMC algorithm in 2D, in which each trace is updated given its neighboring traces and seismic observations. Because of the first-order neighborhood assumption in the lateral direction, the block-Gibbs algorithm can be run in parallel as every pair of two traces separated by at least one trace are conditionally independent given all other traces.

Posterior densities are observed to capture the skewness and multimodality. The defined MMAP predictors are in general observed to be slightly discontinuous compared with the mean predictor which is regressed toward the local mean values.

Our methodology has been empirically verified on a cross section from the Norwegian Sea. Lithology/fluids classes, rock, and fluids properties and elastic attributes are observed to be predicted realistically and are consistent with the data at the well location. We observe MMAP predictors to capture the lateral extent of the lithology/fluids classes. The acceptance rates are found to be satisfactory.

Future research might include the lateral variability of the transition matrices and the rock physics models, and joint petrophysical and seismic inversion together with model parameter inference.

CONCLUSION

We presented a methodology for the joint prediction of lithology/fluid classes, rock and fluid properties, and elastic attributes in a Bayesian framework. Our methodology allows for efficient MCMC assessment of the posterior models in 2D with satisfactory acceptance rates. The posterior model for the rock and fluid properties and elastic attributes are Gaussian mixture densities that show skewness and multimodality. The application to a real data set for the validation part of the method showed promising results consistent with the well-log data.

ACKNOWLEDGMENTS

The authors acknowledge the Uncertainty in Reservoir Evaluation (URE) activity at the Norwegian University of Science and Technology (NTNU), and the School of Energy Resources at the University of Wyoming for the support. We thank PGS and Tullow Oil for providing the data.

REFERENCES

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

Aster, R., C. Thurber, and B. Borchers, 2005, Parameter estimation and inverse problems: Elsevier Academic Press, International Geophysics series 1.

Avseth, P., A. Janke, and F. Horn, 2016, AVO inversion in exploration — Key learnings from a Norwegian Sea prospect: The Leading Edge, **35**, 405–414, doi: 10.1190/tle35050405.1.

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

Besag, J., 1974, Spatial Interaction and the Statistical Analysis of Lattice Systems: Journal of the Royal Statistical Society. Series B (Methodological), **36**, 192–236.

Bosch, M., T. Mukerji, and E. F. Gonzalez, 2010, Seismic inversion for reservoir properties combining statistical rock physics and geostatistics: A review: Geophysics, **75**, no. 5, 75A165–75A176, doi: 10.1190/1.3478209.

Buland, A., and H. Omre, 2003, Bayesian linearized AVO inversion: Geophysics, **68**, 185–198, doi: 10.1190/1.1543206.

Cappé, O., E. Moulines, and T. Ryden, 2005, Inference in hidden markov models: Springer-Verlag New York, Inc., Springer Series in Statistics.

Connolly, P. A., and M. J. Hughes, 2016, Stochastic inversion by matching to large numbers of pseudo-wells: Geophysics, **81**, no. 2, M7–M22, doi: 10.1190/geo2015-0348.1.

de Figueiredo, L. P., D. Grana, M. Santos, W. Figueiredo, M. Roisenberg, and G. S. Neto, 2017, Bayesian seismic inversion based on rock-physics prior modeling for the joint estimation of acoustic impedance, porosity and lithofacies: Journal of Computational Physics, **336**, 128–142.

Dubreuil-Boisclair, C., E. Gloaguen, G. Bellefleur, and D. Marcotte, 2012, Non-Gaussian gas hydrate grade simulation at the Mallik site, Mackenzie Delta, Canada: Marine and Petroleum Geology, **35**, 20–27, doi: 10.1016/j.marpetgeo.2012.02.020.

Dvorkin, J., M. A. Gutierrez, and D. Grana, 2014, Seismic Reflections of Rock Properties: Cambridge University Press.

Eidsvik, J., T. Mukerji, and P. Switzer, 2004, Estimation of geological attributes from a well log: An application of hidden markov chains: Mathematical Geology, **36**, 379–397, doi: 10.1023/B:MATG.0000028443.75501.d9.

Fjeldstad, T., and H. Omre, 2017, Bayesian inversion of convolved hidden Markov models with applications in reservoir prediction: arXiv:1710.06613 [physics.geo-ph].

González, E. F., T. Mukerji, and G. Mavko, 2008, Seismic inversion combining rock physics and multiple-point geostatistics: Geophysics, **73**, no. 1, R11–R21, doi: 10.1190/1.2803748.

Grana, D., 2016, Bayesian linearized rock-physics inversion: Geophysics, **81**, no. 6, D625–D641, doi: 10.1190/geo2016-0161.1.

Grana, D., and E. Della Rossa, 2010, Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion: Geophysics, **75**, no. 3, O21–O37, doi: 10.1190/1.3386676.

Grana, D., T. Fjeldstad, and H. Omre, 2017, Bayesian Gaussian mixture linear inversion for geophysical inverse problems: Mathematical Geosciences, **49**, 493–515, doi: 10.1007/s11004-016-9671-9.

Grana, D., T. Mukerji, J. Dvorkin, and G. Mavko, 2012, Stochastic inversion of facies from seismic data based on sequential simulations and probability perturbation method: Geophysics, **77**, no. 4, M53–M72, doi: 10.1190/geo2011-0417.1.

Gunning, J., and M. E. Glinsky, 2007, Detection of reservoir quality using Bayesian seismic inversion: Geophysics, **72**, no. 3, R37–R49, doi: 10.1190/1.2713043.

Hansen, T. M., A. G. Journel, A. Tarantola, and K. Mosegaard, 2006, Linear inverse Gaussian theory and geostatistics: Geophysics, **71**, no. 6, R101–R111, doi: 10.1190/1.2345195.

Hastie, T., R. Tibshirani, and J. Friedman, 2009, The Elements of Statistical Learning: Data mining, Inference, and Prediction, 2nd ed.: Springer, Springer Series in Statistics.

Houck, R. T., 2002, Quantifying the uncertainty in an AVO interpretation: Geophysics, **67**, 117–125, doi: 10.1190/1.1451395.

Jullum, M., and O. Kolbjørnsen, 2016, A Gaussian-based framework for local Bayesian inversion of geophysical data to rock properties: Geophysics, **81**, no. 3, R75–R87, doi: 10.1190/geo2015-0314.1.

Krumbein, W. C., and M. F. Dacey, 1969, Markov chains and embedded Markov chains in geology: Mathematical Geology, **1**, 79–96, doi: 10.1007/BF02047072.

Larsen, A. L., M. Ulvmoen, H. Omre, and A. Buland, 2006, Bayesian lithology/fluid prediction and simulation on the basis of a Markov-chain prior model: Geophysics, **71**, no. 5, R69–R78, doi: 10.1190/1.2245469.

Mavko, G., T. Mukerji, and J. Dvorkin, 2009, The rock physics handbook, 2nd ed.: Cambridge University Press.

Mukerji, T., A. Jørstad, P. Avseth, G. Mavko, and J. R. Granli, 2001, Mapping lithofacies and pore-fluid probabilities in a North Sea reservoir: Seismic inversions and statistical rock physics: Geophysics, **66**, 988–1001, doi: 10.1190/1.1487078.

Reeves, R., and A. N. Pettitt, 2004, Efficient recursions for general factorisable models: Biometrika, **91**, 751–757, doi: 10.1093/biomet/91.3.751.

Rimstad, K., and H. Omre, 2010, Impact of rock-physics depth trends and Markov random fields on hierarchical Bayesian lithology/fluid prediction: Geophysics, **75**, no. 4, R93–R108, doi: 10.1190/1.3463475.

Rimstad, K., and H. Omre, 2013, Approximate posterior distributions for convolutional two-level hidden Markov models: Computational Statistics & Data Analysis, **58**, 187–200, doi: 10.1016/j.csda.2012.09.001.

Robert, C. P., and G. Casella, 2005, Monte Carlo statistical methods: Springer-Verlag New York, Inc., Springer Texts in Statistics.

Scales, J. A., and L. Tenorio, 2000, Prior information and uncertainty in inverse problems: Geophysics, **66**, 389–397, doi: 10.1190/1.1444930.

Sheriff, R., and L. Geldart, 1995, Exploration seismology: Cambridge University Press.

Silverman, B., 1986, Density estimation for statistics and data analysis: Taylor & Francis, Chapman & Hall/CRC Monographs on Statistics & Applied Probability.
Tarantola, A., 2005, Inverse problem theory and methods for model parameter estimation: SIAM.
Ulrych, T. J., M. D. Sacchi, and A. Woodbury, 2001, A Bayes tour of inversion: A tutorial: Geophysics, **66**, 55–69, doi: 10.1190/1.1444923.
Ulvmoen, M., and H. Omre, 2010, Improved resolution in Baeysian lithology/fluid inversion from seismic prestack data and well observations: Part I — Methodology: Geophysics, **75**, no. 2, R21–R35, doi: 10.1190/1.3294570.
Viterbi, A., 1967, Error bounds for convolutional codes and an asymptotically optimum decoding Algorithm: IEEE Transactions on Information Theory, **13**, 260–269, doi: 10.1109/TIT.1967.1054010.

Paper IV

Bayesian model for lithology/fluid class prediction using a Markov mesh prior fitted from a training image

---

Håkon Tjelmeland, Xin Luo and Torstein Fjeldstad

Geophysical Prospecting vol. 67 (3) (2019)

Geophysical Prospecting

EAGE EUROPEAN ASSOCIATION OF GEOSCIENTISTS & ENGINEERS

Geophysical Prospecting, 2019, **67**, 609–623 doi: 10.1111/1365-2478.12753

A Bayesian model for lithology/fluid class prediction using a Markov mesh prior fitted from a training image

Håkon Tjelmeland*, Xin Luo and Torstein Fjeldstad

Department of Mathematical Sciences, Norwegian University of Science and Technology, Alfred Getz vei 1, 7491 Trondheim, Norway

Received October 2018, revision accepted January 2019

ABSTRACT

We consider a Bayesian model for inversion of observed amplitude variation with offset data into lithology/fluid classes, and study in particular how the choice of prior distribution for the lithology/fluid classes influences the inversion results. Two distinct prior distributions are considered, a simple manually specified Markov random field prior with a first-order neighbourhood and a Markov mesh model with a much larger neighbourhood estimated from a training image. They are chosen to model both horizontal connectivity and vertical thickness distribution of the lithology/fluid classes, and are compared on an offshore clastic oil reservoir in the North Sea. We combine both priors with the same linearized Gaussian likelihood function based on a convolved linearized Zoeppritz relation and estimate properties of the resulting two posterior distributions by simulating from these distributions with the Metropolis–Hastings algorithm. The influence of the prior on the marginal posterior probabilities for the lithology/fluid classes is clearly observable, but modest. The importance of the prior on the connectivity properties in the posterior realizations, however, is much stronger. The larger neighbourhood of the Markov mesh prior enables it to identify and model connectivity and curvature much better than what can be done by the first-order neighbourhood Markov random field prior. As a result, we conclude that the posterior realizations based on the Markov mesh prior appear with much higher lateral connectivity, which is geologically plausible.

Key words: Computing aspects, Inverse problem, Inversion, Mathematical formulation, Seismics.

INTRODUCTION

From seismic data one can predict elastic properties and lithology/fluid classes (LFCs) in a reservoir. This is an inverse problem and for a given seismic data set many solutions exist. Different methods have been used for inverting seismic data to elastic properties and LFCs, both deterministic approaches such as optimization-based methods (Aster, Borchers and Thurber 2011; Sen and Stoffa 2013) and probabilistic approaches such as Bayesian inversion (Tarantola 2005). Using the Bayesian framework, a linearized relation between the data and the elastic properties is commonly used and a Gaussian likelihood function is adopted (see, for example Buland and Omre 2003; Gunning and Glinsky 2007, and the discussion in Grana, Fjeldstad and Omre 2017). When inverting to elastic properties, the prior is also often assumed to be Gaussian, in which case the posterior becomes Gaussian with analytically available mean and covariance (see again Buland and Omre 2003). When inverting to LFCs, other priors have to be used. In particular Rimstad and Omre (2010) define a hierarchical prior, where a Markov random field (Kindermann and Snell 1980; Hurn, Husby and Rue 2003) is used to model the LFCs and conditional on these the elastic properties are assumed to be Gaussian with mean and covariance

*E-mail: haakon.tjelmeland@ntnu.no

functions depending on the LFCs. Grana and Della Rossa (2010) consider a Gaussian mixture prior for the elastic attributes to include multi-modality and skewness in the prior model where the effect of the LFCs is summed out. With a non-Gaussian prior, the posterior is no longer analytically available and Markov chain Monte Carlo (Gilks, Richardson and Spiegelhalter 1996; Robert and Casella 1999; Gamerman and Lopes 2006) must typically be used to estimate properties of the resulting posterior distribution. It is also challenging to specify lateral connectivity and spatial dependency laterally for non-Gaussian priors, and often the inverse problem is solved trace by trace before a smoother is applied afterwards (Connolly and Hughes 2016).

To specify a prior that reflects available prior information in a spatial problem, such as inversion of seismic data, can be difficult. The properties of a Gaussian field is analytically well understood, so by adopting a Gaussian prior as in Buland and Omre (2003), the prior specification process is simplified. Eidsvik, Mukerji and Switzer (2004) consider a one-dimensional problem and assume a Markov chain prior to predict geological attributes from well log data, and Fjeldstad and Omre (2019) use a similar model to predict LFCs and elastic attributes from seismic data. The properties of Markov chains are also analytically available, which again simplifies the specification of a reasonable prior. For most non-Gaussian spatial prior models the situation is less favourable. In Rimstad, Avseth and Omre (2012), a discrete Markov random field prior is used for the LFCs. The properties of discrete Markov random fields are analytically not available, which makes it difficult to verify the properties of the chosen prior. To cope with the problem of specifying non-Gaussian spatial prior models, it has in geostatistics become common practice to estimate the prior model from a so-called training image. A training image is from an outcrop or a constructed scene assumed to have the same spatial structure as the phenomenon under study. The idea is to estimate a prior model from one or more training images (see the discussion in Mariethoz and Caers 2014). Various multiple-point statistics models (Guardiano and Srivastava 1993; Strebelle 2002; Journel and Zhang 2006; Zhang *et al.* 2012) have been defined to implement this idea. These models are algorithmically defined. The nodes in a lattice are visited in a random order and when a node is visited, the value in that node is simulated conditional on values in previously visited nodes, where the conditional distribution used is estimated from the training image. There are two serious complications associated with the use of multiple-point statistics models. First, the number of conditional distributions that has to be estimated from the training image is enormous, and the information content in a typical training image is not sufficient to estimate this number of parameters. Emery and Lantuéjoul (2014) are discussing this issue mathematically. Second, the models are only algorithmically defined and no simple-to-evaluate expressions are available for the estimated model. The implication of this is that if we want to use the estimated model as a prior and generate realizations conditional on some observed data, it is in general not clear how to do this. Since we have no analytical formula for the prior, we do not have an expression for the posterior. This issue is also discussed in Toftaker and Tjelmeland (2013). As alternatives to the multiple-point statistics models, Arnesen and Tjelmeland (2017) and Luo and Tjelmeland (2019) introduce procedures for fitting Markov random fields and Markov mesh models, respectively, to a given training image. For these model classes explicit expressions for the distributions are available, so to simulate from a corresponding conditional distribution Markov chain Monte Carlo procedures can, for example be employed.

The purpose of this article is to demonstrate how inversion of seismic data into LFCs can be accomplished in a Bayesian framework by estimating a prior model for the LFCs from a given training image, and combine this with a linearized and Gaussian likelihood function. We fit a Markov mesh prior model to a training image as discussed in Luo and Tjelmeland (2019) and use Markov chain Monte Carlo to simulate from the resulting posterior distribution as discussed in Rimstad *et al.* (2012). A Markov mesh prior is used for lithology/fluid prediction also in Stien and Kolbjørnsen (2011), but they specify manually the neighbourhood and interaction structures and fit only the parameter values to the training image. In our fitting procedure, we fit both neighbourhood and interaction structures and parameter values to the given training image. To focus on the methodological aspects, we consider a situation with only two LFCs, oil sand and shale. In particular, we compare the results from our procedure with what we get by instead using a simpler manually chosen Markov random field prior.

The article has the following layout. First, we present the data set and the associated training image, and we analyse and introduce our Bayesian model formulation. A Gaussian likelihood function is defined, and its properties are discussed. We introduce the Markov mesh and Markov random field priors, and in particular focus on how we fit the Markov mesh prior to the given training image. Next, a sampling algorithm for the posterior distribution is discussed. The two priors are tested on a real two-dimensional section case study in the

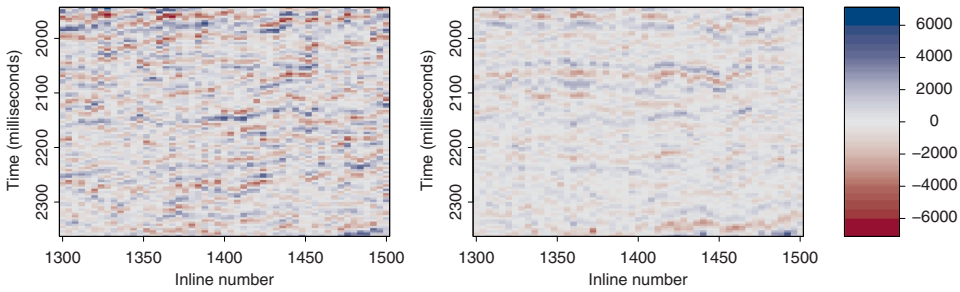


Figure 1 Near (left) and far (right) offset seismic data used for lithology/fluid prediction.

North Sea. Finally, we discuss the posterior properties of the two priors and provide some closing remarks.

METHODOLOGY

The objective is to demonstrate and compare two different prior models in a Bayesian framework to predict lithology/fluid classes (LFCs) in the subsurface. In this section, we introduce the data set and formulate the inverse problem in a Bayesian setting, define a likelihood function and the two priors and discuss posterior simulation.

Data set and Bayesian model formulation

In this article, we consider a seismic section from the Alvheim field in the North Sea, which is a clastic oil reservoir. The Alvheim field is characterized by a complex sand lobe geometry and is buried approximately 2 km below the sea floor. In the analysis, we use one near and one far offset seismic data represented in a 105×51 lattice $G = \{(i, j)|i = 1, \ldots, 105; j = 1, \ldots, 51\}$. The stacked sections were generated from pre-stack time migrated common depth gathers (see Rimstad *et al.* 2012 for further processing details). The seismic data are shown in Fig. 1. The horizontal and vertical sample rates are about 100 m and 4 milliseconds, respectively. We let $\mathbf{d}_{ij}, (i, j) \in G$ denote a vector of size 2 containing the observed near and far offset seismic data in node $(i, j) \in G$, and let $\mathbf{d}$ be a vector where all $\mathbf{d}_{ij}, (i, j) \in G$ are stacked on top of each other. We model two LFCs, oil sand and shale. For each node $(i, j) \in G$ we let $\boldsymbol{\kappa}_{ij} \in \{0, 1\}$ denote the LFC in node (i, j), where $\boldsymbol{\kappa}_{ij} = 0$ and $\boldsymbol{\kappa}_{ij} = 1$ represent shale and oil sand, respectively. We let $\boldsymbol{\kappa}$ be a vector of all $\boldsymbol{\kappa}_{ij}, (i, j) \in G$ stacked on top of each other. To estimate a Markov mesh prior distribution for $\boldsymbol{\kappa}$, we use a training image from Lang and Grana (2017), which is shown in Fig. 2. The training image is from a reservoir with similar geological characteristics as the reservoir under study. It should be noted that it is only used to give prior information about the spatial continuity of the classes, and does not in any way represent a class map of the section of the underground from where the seismic data is coming.

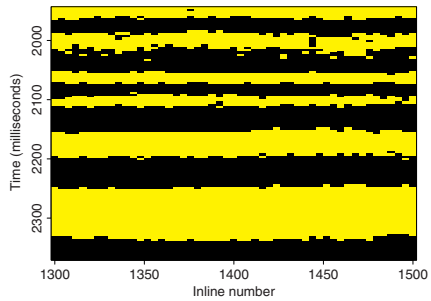


Figure 2 Training image that we use to estimate a Markov mesh prior distribution for the spatial distribution of lithology/fluid classes, $\boldsymbol{\kappa}$. Black and yellow represent shale and oil sand, respectively.

To model the relation between $\boldsymbol{\kappa}$ and $\mathbf{d}$, we first introduce a vector $\mathbf{m} = \{\mathbf{m}_{ij}, (i, j) \in G\}$ of elastic properties, where $\mathbf{m}_{ij}$ is a vector of length 2. We let the first element in $\mathbf{m}_{ij}$ be the impedance, that is, the product of the density ρ and the pressure-wave velocity $\mathbf{v}_p$, in node (i, j) and let the second element be the $\mathbf{v}_p/\mathbf{v}_s$ ratio in the same node, where $\mathbf{v}_s$ is the shear-wave velocity.

For the three variables $\boldsymbol{\kappa}$, $\mathbf{m}$ and $\mathbf{d}$ we adopt a Bayesian model. We let $p(\boldsymbol{\kappa})$ denote a prior distribution for $\boldsymbol{\kappa}$ and let $p(\mathbf{m}|\boldsymbol{\kappa})$ denote the conditional distribution for the elastic parameters $\mathbf{m}$ given the LFCs $\boldsymbol{\kappa}$. Finally, we assume the seismic data $\mathbf{d}$ to be conditionally independent of $\boldsymbol{\kappa}$ when the elastic properties $\mathbf{m}$ are given. We let $p(\mathbf{d}|\mathbf{m})$ denote the conditional distribution for the seismic data $\mathbf{d}$ given elastic properties $\mathbf{m}$.

The $p(\mathbf{d}|\mathbf{m})$ represents a probabilistic formulation of the forward model. Bayes' theorem then gives

$$p(\boldsymbol{\kappa}|\mathbf{d}) \propto p(\boldsymbol{\kappa})p(\mathbf{d}|\boldsymbol{\kappa}), \tag{1}$$

where

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

In the following we first outline the details of $p(\mathbf{m}|\boldsymbol{\kappa})$ and $p(\mathbf{d}|\mathbf{m})$, which is used to specify the likelihood $p(\mathbf{d}|\boldsymbol{\kappa})$, and thereafter specify the prior $p(\boldsymbol{\kappa})$ before we describe the Markov chain Monte Carlo procedure we use to simulate from $p(\boldsymbol{\kappa}|\mathbf{d})$.

Likelihood model

Following Grana *et al.* (2017) and Fjeldstad and Omre (2019), we adopt a linearized and Gaussian likelihood for the forward model for $\mathbf{d}$ given $\boldsymbol{\kappa}$. More specifically, we assume each of $p(\mathbf{m}|\boldsymbol{\kappa})$ and $p(\mathbf{d}|\mathbf{m})$ to be Gaussian, the conditional mean of $\mathbf{d}$ given $\mathbf{m}$ to be a linear function of $\mathbf{m}$ and the conditional covariance matrix of $\mathbf{d}$ given $\mathbf{m}$ not to be a function of $\mathbf{m}$. In the following we outline the distributions $p(\mathbf{m}|\boldsymbol{\kappa})$ and $p(\mathbf{d}|\mathbf{m})$ in more detail, starting with $p(\mathbf{m}|\boldsymbol{\kappa})$.

We assume $\mathbf{m}|\boldsymbol{\kappa}$ to be Gaussian and

$$\mathrm{E}[\mathbf{m}_{ij}|\boldsymbol{\kappa}] = \mu_{\boldsymbol{\kappa}_{ij}} \quad \text{and} \quad \mathrm{cov}[\mathbf{m}_{ij}|\boldsymbol{\kappa}] = \Sigma_{\boldsymbol{\kappa}_{ij}}, \tag{3}$$

where μ_0 and Σ_0 are the conditional mean and covariance for $\mathbf{m}_{ij}$ if node (i, j) contains shale ($\boldsymbol{\kappa}_{ij} = 0$), and μ_1 and Σ_1 are corresponding quantities when node (i, j) contains oil sand ($\boldsymbol{\kappa}_{ij} = 1$). Moreover, we assume a separable correlation function $\rho((i, j), (k, l))$ for $\mathbf{m}|\boldsymbol{\kappa}$, and do not allow the correlations to depend on the LFCs $\boldsymbol{\kappa}$. The validity of the latter assumption can be discussed, but is adopted to make evaluation of the likelihood function computationally more efficient.

For the forward model $p(\mathbf{d}|\mathbf{m})$, we use a convolved linearized approximation of the Zoeppritz equation (Buland and Omre 2003) based on the Aki–Richards formulation that is valid for weak vertical contrasts (Aki and Richards 1980). The vector $\mathbf{d}$ is then formed from $\mathbf{m}$ in several steps. First, all vertical first-order contrasts or differences $\mathbf{m}_{ij} - \mathbf{m}_{i-1,j}$ are formed by pre-multiplying $\mathbf{m}$ with a matrix $\mathbf{D}$. Thereafter, reflection coefficients are formed by pre-multiplying $\mathbf{Dm}$ with a block diagonal matrix $\mathbf{A}$, where all the blocks are identical 2×2 matrices containing coefficients in the Aki–Richards formulation. The mean value of the seismic data are then formed via a convolution of each column of $\mathbf{ADm}$ with wavelets. Different wavelets are used for the near and far offset seismic data as shown in Fig. 3. These wavelets are estimated from data in a well in the same reservoir as the seismic data is coming from. This well is, however, located some distance away from the seismic section we are studying. The effect of the convolutions can be written as pre-multiplying $\mathbf{ADm}$ with a matrix $\mathbf{W}$. Finally, a zero mean Gaussian error term ε with a fixed covariance matrix Σ_ε is added to $\mathbf{WADm}$. Thus, $\mathbf{D}|\mathbf{m}$ is Gaussian with

$$\mathrm{E}[\mathbf{d}|\mathbf{m}] = \mathbf{WADm} \quad \text{and} \quad \mathrm{cov}[\mathbf{d}|\mathbf{m}] = \Sigma_\varepsilon. \tag{4}$$

With the Gaussian distributions specified above for $p(\mathbf{m}|\boldsymbol{\kappa})$ and $p(\mathbf{d}|\mathbf{m})$ it follows from standard properties of the Gaussian distribution that also $p(\boldsymbol{d}|\boldsymbol{\kappa})$ becomes Gaussian. Moreover, expressions are available for the mean $\mathrm{E}[\mathbf{d}|\boldsymbol{\kappa}]$ and the covariance $\mathrm{cov}[\mathbf{d}|\boldsymbol{\kappa}]$ as functions of $\boldsymbol{\kappa}$, μ_0, μ_1, Σ_0, Σ_1, $\rho(\cdot, \cdot)$, $\mathbf{W}$, $\mathbf{A}$, $\mathbf{D}$ and Σ_ε.

Prior models

The main purpose of this article is to demonstrate how a Markov mesh model fitted to a training image can be used as prior in a Bayesian model for lithology/fluid prediction. However, we also want to study how the inversion results change when using such a prior relative to what we get using a simpler manually specified prior. To fit a Markov mesh model to a training image involves extra working and computing time, so there is no reason to do so unless it results in a significant change in the inversion results. In the following, we first specify the class of Markov mesh models and briefly discuss the procedure we use to fit the model to the training image in Fig. 2. Thereafter we describe a simpler manually specified prior we use for comparison, the profile Markov random field introduced in Ulvmoen and Omre (2010).

Markov mesh prior

An introduction to the class of Markov mesh models can be found in Abend, Harley and Kanal (1965) and the more general class of partially ordered Markov models is defined in Cressie and Davidson (1998). In the following description, we limit the attention to binary fields and introduce the necessary notions to define homogeneous Markov mesh models defined on a rectangular lattice.

Let $G = \{(i, j)|i = 1, \ldots, n_1; j = 1, \ldots, n_2\}$ be a rectangular lattice, to each node $(i, j) \in G$ of which we associate a binary variable $\boldsymbol{\kappa}_{ij} \in \{0, 1\}$. We let $\boldsymbol{\kappa} = (\kappa_{ij} : (i, j) \in G)$ denote the collection of all these binary variables and use $\boldsymbol{\kappa}_\lambda = (\boldsymbol{\kappa}_{ij} : (i, j) \in \lambda)$ to denote the collection of binary variables in a set $\lambda \subseteq G$ of nodes. The Markov mesh model is

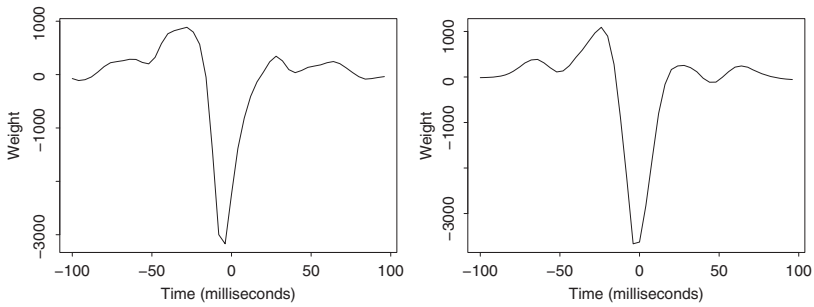


Figure 3 Wavelets used for the near (left) and far (right) offset seismic data. The *x*- and *y*-axes show vertical distance and wavelet values, respectively.

based on numbering the nodes in G from 1 to $n_1 \cdot n_2$ in the lexicographical order. Without loss of generality, the distribution of $\boldsymbol{\kappa}$ can then be expressed as

$$p(\boldsymbol{\kappa}) = \prod_{(i,j)\in G} p(\boldsymbol{\kappa}_{ij}|\boldsymbol{\kappa}_{\rho_{ij}}), \tag{5}$$

where ρ_{ij} is the set of all nodes coming before node (i, j), that is,

$$\rho_{ij} = \{(k,l) \in G : nk + l < ni + j\}. \tag{6}$$

The set ρ_{ij} is called the predecessor set of node (i, j). The central assumption in Markov mesh models is that $p(\boldsymbol{\kappa}_{ij}|\boldsymbol{\kappa}_{\rho_{ij}})$ has a Markov property in that

$$p(\boldsymbol{\kappa}_{ij}|\boldsymbol{\kappa}_{\rho_{ij}}) = p(\boldsymbol{\kappa}_{ij}|\boldsymbol{\kappa}_{v_{ij}}), \tag{7}$$

where $v_{ij} \subseteq \rho_{ij}$ is called the sequential neighbourhood of node (i, j). Following Luo and Tjelmeland (2019), we assume that all the sequential neighbourhoods are generated via a translation of a template sequential neighbourhood τ. The set τ can best be thought of as the sequential neighbourhood of node (0,0) in an infinite lattice. More precisely, τ should contain a finite number of elements and

$$\tau \subset \{(i, j) : i \in \mathbb{Z}^-, j \in \mathbb{Z}\} \cup \{(0, j) : j \in \mathbb{Z}^-\}, \tag{8}$$

where $\mathbb{Z} = \{0, \pm1, \pm2, \ldots\}$ and $\mathbb{Z}^- = \{-1, -2, \ldots\}$ are the sets of all integers and all negative integers, respectively. Given the set τ we assume the sequential neighbourhood v_{ij} to be generated by translating each element in τ a distance (i, j) and, if necessary, dropping elements falling outside the lattice G. Mathematically, v_{ij} is then given as

$$v_{ij} = (\tau \oplus (i, j)) \cap G, \tag{9}$$

where the translation operator $\oplus$ is defined as

$$\tau \oplus (i, j) = \{(k + i, l + j) : (k, l) \in \tau\}. \tag{10}$$

Constructing v_{ij} in this way, the sequential neighbourhoods for all nodes sufficiently far away from the lattice borders will have the same form.

Still following Luo and Tjelmeland (2019), we model $p(\boldsymbol{\kappa}_{ij}|\boldsymbol{\kappa}_{v_{ij}})$ by assuming the logit transformation of $p(\boldsymbol{\kappa}_{ij} = 1|\boldsymbol{\kappa}_{v_{ij}})$ to be given by

$$\begin{aligned}\text{logit}\Big[p(\boldsymbol{\kappa}_{ij} = 1|\boldsymbol{\kappa}_{v_{ij}})\Big] &= \ln\left(\frac{p(\boldsymbol{\kappa}_{ij} = 1|\boldsymbol{\kappa}_{v_{ij}})}{1 - p(\boldsymbol{\kappa}_{ij} = 1|\boldsymbol{\kappa}_{v_{ij}})}\right) \\ &= \theta(\xi(\boldsymbol{\kappa}, \tau, (i, j))),\end{aligned} \tag{11}$$

where $\xi(\boldsymbol{\kappa}, \tau, (i, j)) \subseteq \tau$ is the set of elements $(k, l) \in \tau$ associated to a node with oil sand in the sequential neighbourhood for node (i, j), that is,

$$\begin{aligned}&\xi(\boldsymbol{\kappa}, \tau, (i, j)) \\ &\quad = \{(k, l) \in \tau : (i + k, j + l) \in G \text{ and } \boldsymbol{\kappa}_{i+k,j+l} = 1\},\end{aligned} \tag{12}$$

and $\theta(\cdot)$ is a pseudo-Boolean function (Hammer and Holzman 1992; Grabisch, Marichal and Roubens 2000) to be specified. One should note that, as we assume the same function $\theta(\cdot)$ for all nodes $(i, j) \in G$ we get a homogeneous model. Moreover, one should note that the definition of $\xi(\boldsymbol{\kappa}, \tau, (i, j))$ implies that for nodes (i, j) close to the boundary of the lattice, so that $(\tau \oplus (i, j)) \setminus G \neq \emptyset$, the conditional distribution $p(\boldsymbol{\kappa}_{ij}|\boldsymbol{\kappa}_{v_{ij}})$ becomes as if one had an infinite lattice where all variables associated to nodes outside of G were zero.

The last step in specifying the Markov mesh model is to choose the function $\theta(\cdot)$. This is a real valued function, where the argument is a subset of τ specifying for which sequential neighbours the associated binary variable is equal to 1. Without loss of generality, the $\theta(\cdot)$ can be uniquely expressed in

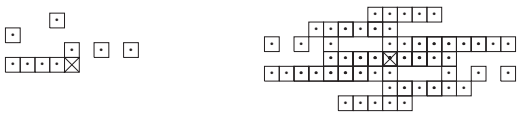


Figure 4 The sequential neighbourhood (left) and the corresponding Markov random field neighbourhood (right) for the fitted prior $p(\boldsymbol{\kappa})$. The nodes marked with a dot are (sequential) neighbours of the node marked with a cross.

terms of a collection of interaction parameters $\{\beta(\lambda) : \lambda \subseteq \tau\}$ by

$$\theta(\lambda) = \sum_{\lambda^* \subseteq \lambda} \beta(\lambda^*) \quad \text{for } \lambda \subseteq \tau. \tag{13}$$

The number of interaction parameters is $2^{|\tau|}$, where $|\tau|$ is the number of elements in τ. Unless $|\tau|$ is very small the number of parameters necessary to specify $\theta(\cdot)$ is thereby very large. We still follow Luo and Tjelmeland (2019) and limit the number of model parameters by restricting many of the interaction parameters to be zero. More specifically, for some $\Lambda \subseteq \Omega(\tau)$, where $\Omega(\tau)$ is the set of all subsets of τ, we assume $\beta(\lambda) = 0$ for all $\lambda \notin \Lambda$. Thus, we specify the Markov mesh model by choosing the sets τ and Λ and the interaction values $\{\beta(\lambda) : \lambda \in \Lambda\}$.

To fit the Markov mesh model specified above to the training image in Fig. 2, we adopt the Bayesian procedure introduced in Luo and Tjelmeland (2019), including the hyperparameter values used in that article. A prior is specified for τ, Λ and $\{\beta(\lambda) : \lambda \in \Lambda\}$ and assuming the training image to be a sample from the specified Markov mesh model, a Metropolis–Hastings algorithm is used to generate samples from the resulting posterior distribution for τ, Λ and $\{\beta(\lambda) : \lambda \in \Lambda\}$. When we conditioned on the training image in Fig. 2, the convergence of the Metropolis–Hastings algorithm was so slow that we were unable to obtain convergence within a reasonable computation time. As a pragmatic approach to obtain a reasonable prior $p(\boldsymbol{\kappa})$ we simply run the Metropolis–Hastings algorithm in Luo and Tjelmeland (2019) for a large number of iterations and used the last values for τ, Λ and $\{\beta(\lambda) : \lambda \in \Lambda\}$ in this run to define the prior $p(\boldsymbol{\kappa})$. The resulting $p(\boldsymbol{\kappa})$ prior has $|\tau| = 9$ sequential neighbours and $|\Lambda| = 31$ interaction parameters that are allowed to differ from zero. The sequential neighbourhood τ is illustrated in the left part of Fig. 4, while the complete specification of τ, Λ and $\{\beta(\lambda) : \lambda \in \Lambda\}$ is given in an Appendix. The Markov mesh prior can also be reformulated as a Markov random field. The resulting Markov random field has then a neighbourhood system where the set of neighbours to node $(i, j) \in G$ is

$$\partial_{ij} = v_{ij} \cup \left(\bigcup_{(k,l) \in G : (i,j) \in v_{kl}} (v_{kl} \cup \{(k, l)\}) \right). \tag{14}$$

For nodes sufficiently far away from the lattice borders the ∂_{ij} becomes as shown in the right part of Fig. 4.

The best way to understand the properties of the prior is perhaps to look at realizations sampled from $p(\boldsymbol{\kappa})$, four of which are shown in Fig. 5. We see that the fitted prior is reproducing large continuous areas of shale and oil sand as seen in the training image, but the boundaries between shale and oil sand is less horizontal in the realizations from the prior than in the training image.

Profile Markov random field prior

The profile Markov random field prior was first defined and used for seismic inversion in Ulvmoen and Omre (2010) (see also Rimstad and Omre 2010). Even though the prior class is defined for categorical variables, in our description of the model we limit the attention to the binary variable case.

Let again $G = \{(i, j) | i = 1, \ldots, n_1; j = 1, \ldots, n_2\}$ be a rectangular lattice, where in each node we associate a binary variable $\kappa_{ij} \in \{0, 1\}$. We let $C_j = \{(i, j) : i = 1, \ldots, n_1\}$ be the set of nodes in profile or column j of the lattice G and let $\boldsymbol{\kappa}_{C_j} = (\kappa_{ij} : (i, j) \in C_j)$ denote the collection of the binary variables associated to this column. The collection of all the binary variables except the ones in column j we denote by $\boldsymbol{\kappa}_{-C_j}$. The profile Markov random field prior is then specified by first adopting the Markov property

$$p(\boldsymbol{\kappa}_{C_j} | \boldsymbol{\kappa}_{-C_j}) = p(\boldsymbol{\kappa}_{C_j} | \boldsymbol{\kappa}_{C_{j-1}}, \boldsymbol{\kappa}_{C_{j+1}}), \tag{15}$$

that is, given the values in columns $j - 1$ and $j + 1$, the values in column j are independent of the values in the remaining columns. Secondly, the profile Markov random field prior assumes $p(\boldsymbol{\kappa}_{C_j} | \boldsymbol{\kappa}_{C_{j-1}}, \boldsymbol{\kappa}_{C_{j+1}})$ to be a Markov chain down along the column,

$$\begin{aligned} p(\boldsymbol{\kappa}_{C_j} | \boldsymbol{\kappa}_{C_{j-1}}, \boldsymbol{\kappa}_{C_{j+1}}) = {} & p(\kappa_{(1,j)} | \boldsymbol{\kappa}_{(1,j-1)}, \boldsymbol{\kappa}_{(1,j+1)}) \\ & \times \prod_{i=2}^{n} p(\kappa_{(i,j)} | \boldsymbol{\kappa}_{(i-1,j)}, \boldsymbol{\kappa}_{(i,j-1)}, \boldsymbol{\kappa}_{(i,j+1)}), \end{aligned} \tag{16}$$

where the conditional distribution $p(\kappa_{(i,j)} | \boldsymbol{\kappa}_{(i-1,j)}, \boldsymbol{\kappa}_{(i,j-1)}, \boldsymbol{\kappa}_{(i,j+1)})$ is the same for all values of i and j. For the transition probabilities, we have used values adapted from

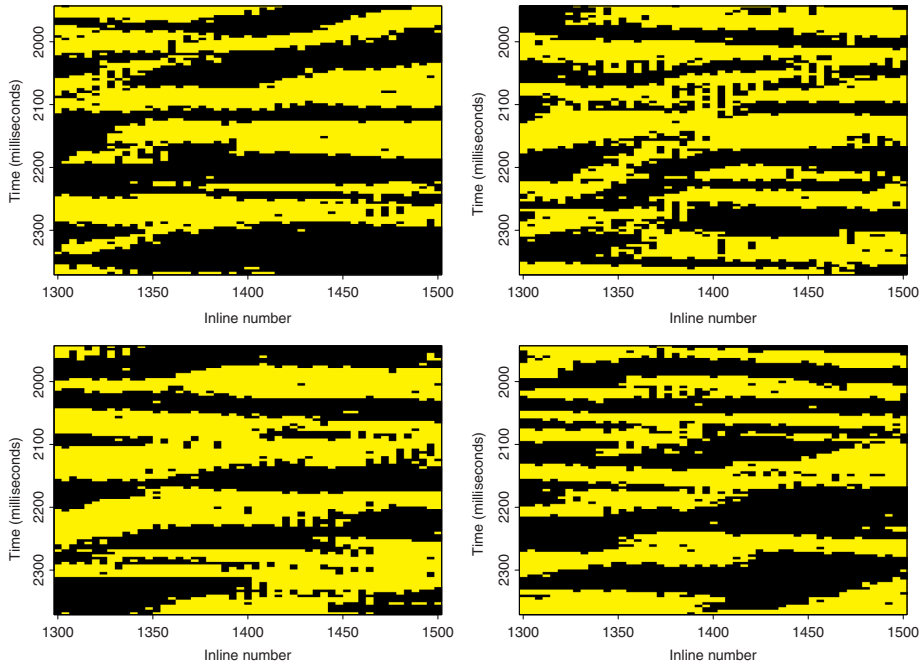


Figure 5 Four independent realizations from the Markov mesh prior fitted to the training image shown in Fig. 2. Black and yellow represent shale and oil sand, respectively.

Table 1 Values used for $p(\kappa_{(i,j)}|\kappa_{(i-1,j)}, \kappa_{(i,j-1)}, \kappa_{(i,j+1)})$ in the specification of the profile Markov random field prior

$\kappa_{i,j-1}=0, \kappa_{i,j+1}=0$			$\kappa_{i,j-1}=0, \kappa_{i,j+1}=1$		
	$\kappa_{ij}=0$	$\kappa_{ij}=1$		$\kappa_{ij}=0$	$\kappa_{ij}=1$
$\kappa_{i-1,j}=0$	0.9877	0.0123	$\kappa_{i-1,j}=0$	0.6539	0.3461
$\kappa_{i-1,j}=1$	0.8339	0.1661	$\kappa_{i-1,j}=1$	0.1056	0.8944
$\kappa_{i,j-1}=1, \kappa_{i,j+1}=0$			$\kappa_{i,j-1}=1, \kappa_{i,j+1}=1$		
	$\kappa_{ij}=0$	$\kappa_{ij}=1$		$\kappa_{ij}=0$	$\kappa_{ij}=1$
$\kappa_{i-1,j}=0$	0.6539	0.3461	$\kappa_{i-1,j}=0$	0.0425	0.9575
$\kappa_{i-1,j}=1$	0.1056	0.8944	$\kappa_{i-1,j}=1$	0.0028	0.9972

corresponding values in Rimstad *et al.* (2012). Table 1 defines the values we have used. Rimstad and Omre (2010) describe the structure used to specify these values. The basic idea is that these values should represent high probability for lateral continuity of oil sand and shale. The initial distribution $p(\kappa_{(1,j)}|\kappa_{(1,j-1)}, \kappa_{(1,j+1)})$ are set equal to $p(\kappa_{(i,j)}|\kappa_{(i-1,j)}=0, \kappa_{(i,j-1)}, \kappa_{(i,j+1)})$, that is, conditioning on shale being present above the lattice. Correspondingly, $p(\kappa_{(i,j)}|\kappa_{(i-1,j)}, \kappa_{(i,j-1)}, \kappa_{(i,j+1)})$ for the left and rightmost columns $j=1$ and $j=n_2$ are defined by conditioning on shale being present outside the lattice.

Posterior model and simulation algorithm

For each of the Markov mesh and profile Markov random field priors we obtain a posterior distribution for the LFCs $\boldsymbol{\kappa}$, given in (1). To explore and estimate properties of the two posterior distributions we adopt the Metropolis–Hastings algorithm (Gilks *et al.* 1996; Robert and Casella 1999; Gamerman and Lopes 2006). Since the wavelets in the likelihood model induce strong dependencies between different κ_{ij}s in the same column, a simple single-site updating scheme would give a Markov chain with a long burn-in and slow mixing. We therefore adopt the proposal scheme previously used in Rimstad and Omre (2010) and propose in each iteration new values for all LFCs in one column. Using notation from the discussion of the profile Markov random field prior, the joint full conditional for the LFCs in column j is

$$p(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j}, \boldsymbol{d}) \propto p(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j})p(\boldsymbol{d}|\boldsymbol{\kappa}). \quad (17)$$

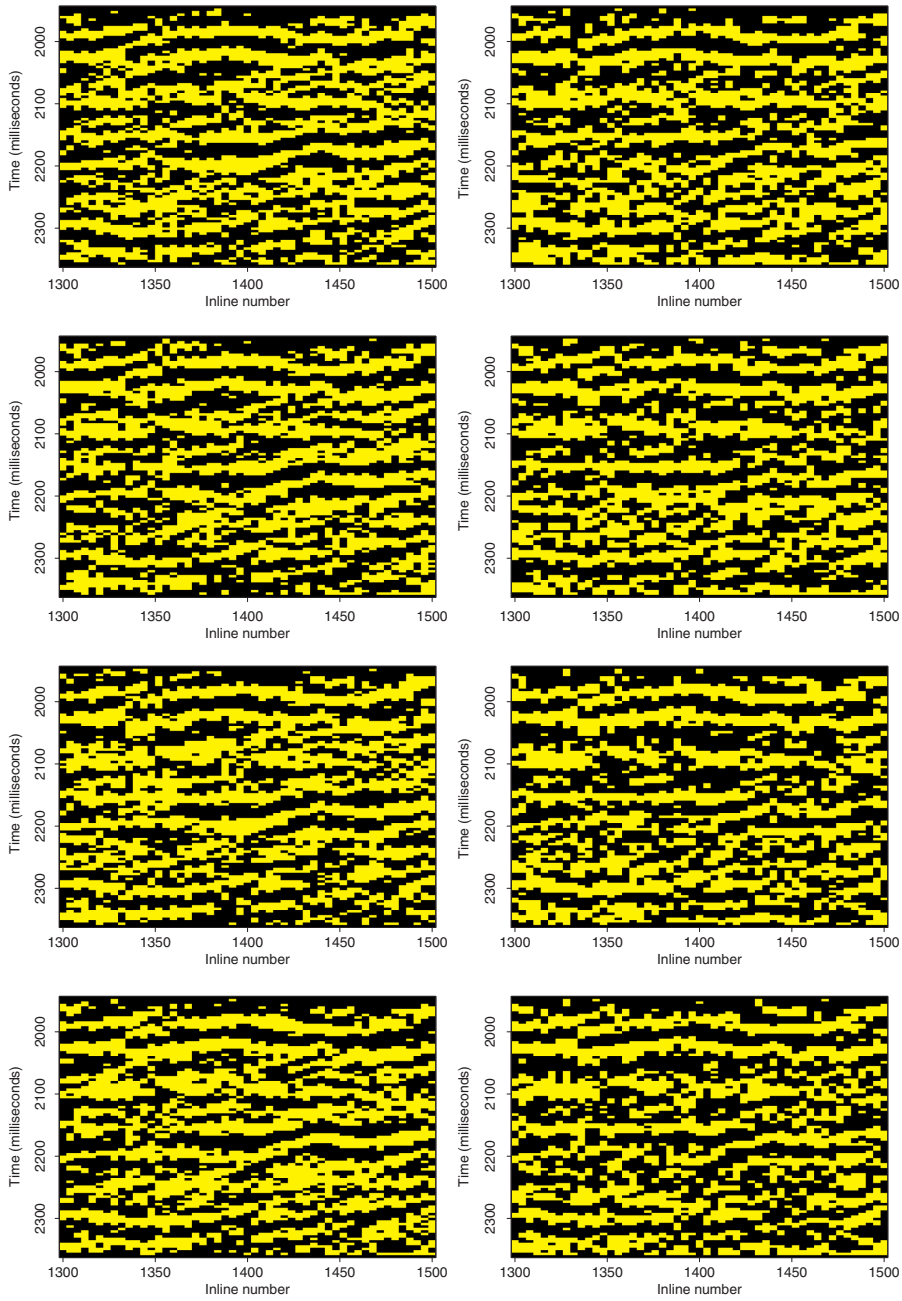


Figure 6 The left column shows four independent realizations from the posterior distribution when using the Markov mesh prior. The right column shows correspondingly four independent realizations from the posterior distribution when using the profile Markov random field prior. Black and yellow represent shale and oil sand, respectively.

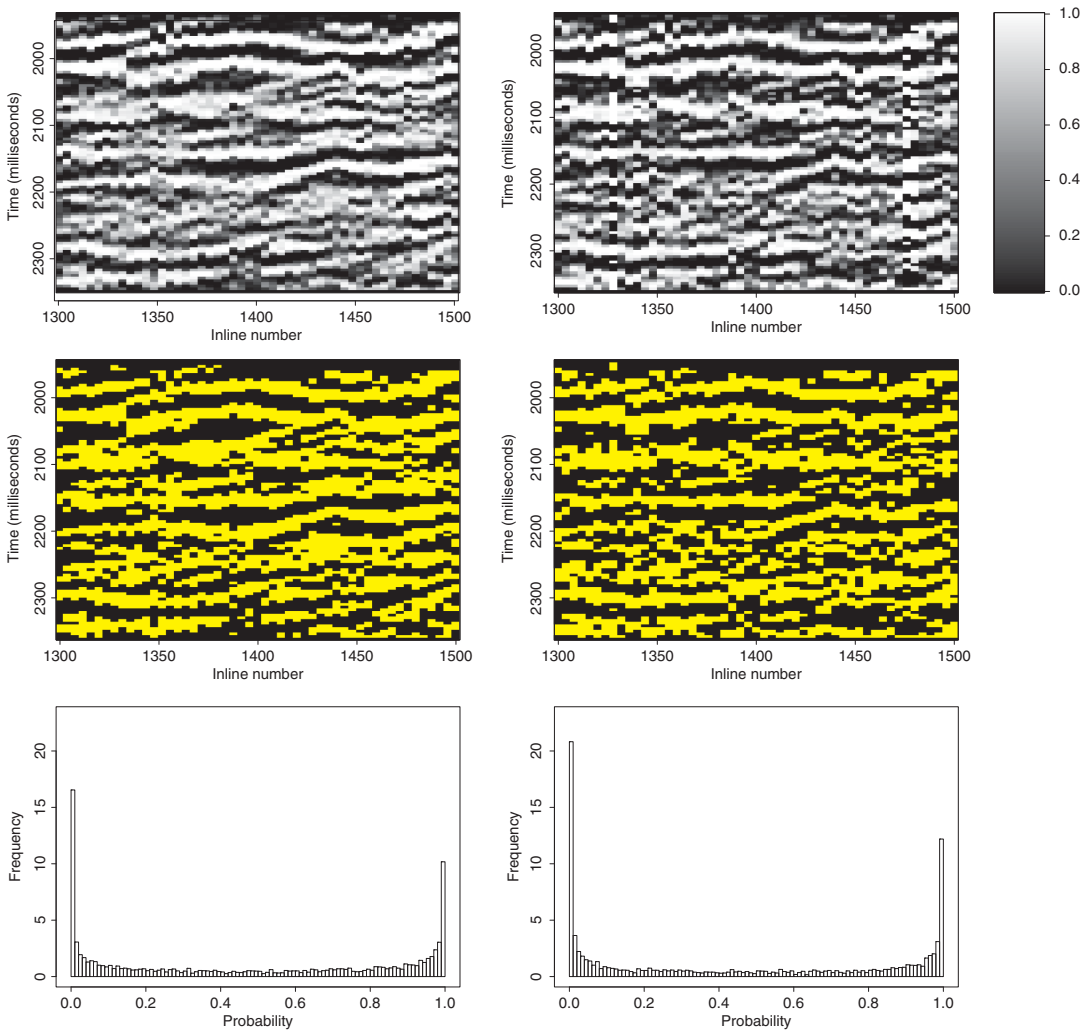


Figure 7 Upper row: Estimated posterior marginal probabilities for oil sand when using the Markov mesh (left) and the profile Markov random field (middle) priors. The colour scale is shown in the rightmost plot. Middle row: Estimated marginal posterior mode for each node. Black and yellow represent shale and oil sand, respectively. Lower row: Probability histograms of estimated posterior marginal probabilities when using the Markov mesh (left) and the profile Markov random field (right) priors, respectively.

To sample from this distribution is, however, computationally very expensive due to the long range dependencies in $\boldsymbol{\kappa}_{C_j}$ induced by the wavelets in the likelihood model. Still following Rimstad and Omre (2010), we therefore adopt the approximation scheme specified in Larsen *et al.* (2006) to construct an approximation $p_\nu^\star(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j}, \mathbf{d})$ to $p(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j}, \mathbf{d})$, where ν is an algorithmic tuning parameter, and generate potential new values for $\boldsymbol{\kappa}_{C_j}$ by sampling from $p_\nu^\star(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j}, \mathbf{d})$. The $p_\nu^\star(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j}, \mathbf{d})$ is a higher order Markov chain and thereby by construction easy to sample from. In general the approximation quality grows with ν, but so does the computation time required for simulating one realization from $p_\nu^\star(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j}, \mathbf{d})$. Based on preliminary runs of the Metropolis–Hastings algorithm we find a value for ν which

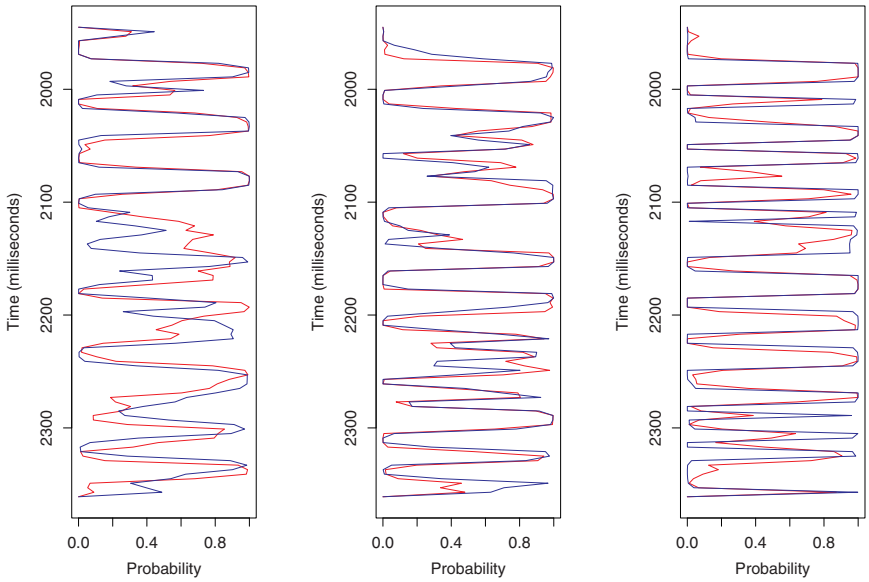


Figure 8 Estimated marginal probabilities for oil sand in three traces. Results when using the Markov mesh and profile Markov random field priors are shown by red and blue lines, respectively. The left, middle and right plots show the marginal probabilities in traces $j = 15$, $j = 30$ and $j = 45$, respectively.

gives reasonable acceptance rates for the Metropolis–Hastings algorithm.

To run the Metropolis–Hastings scheme discussed above we first need to have expressions for $p(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j})$ for each of the two priors. For the profile Markov random field prior this is by construction given by (15) and the values in Table 1. To obtain $p(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j})$ for the Markov mesh prior we first need to reformulate the Markov mesh model as a Markov random field as discussed related to Fig. 4. When the prior $p(x)$ is formulated as a Markov random field it is straightforward to find the corresponding $p(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j})$ by first ignoring potential functions for cliques which do not include any node in C_j and thereafter plugging in values for x_{-C_j} in the remaining potential functions. In particular, for the Markov mesh sequential neighbourhood used in this study, $p(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j})$ becomes a third-order Markov chain.

The second factor in (17) is a high-dimensional multivariate Gaussian density. To be able to evaluate this efficiently, it is essential that we have chosen the correlation structure of $\boldsymbol{d}|\boldsymbol{\kappa}$ to be separable. For each of the two priors we run the Metropolis–Hastings algorithm scheme for the resulting posterior distribution for a large number of iterations. We use standard output analysis to identify and discard a burn-in period. In the next section, we use the $\boldsymbol{\kappa}$ realizations after the burn-in period to estimate and compare properties of the two posterior distributions.

NORTH SEA CASE STUDY

Recall that the objective is to assess the posterior of the lithology/fluid classes (LFCs) $\boldsymbol{\kappa}_{ij} \in G$ given seismic AVO data $\mathbf{d}$ in a clastic oil reservoir in the North Sea. That is, we want to assess the posterior $p(\boldsymbol{\kappa}|\mathbf{d})$ given in (1), for the two prior models discussed earlier. Note that the two posteriors will not be identical since the priors are different.

To study and compare the properties of the two posterior distributions, we can first look at the posterior realizations shown in Fig. 6. The left and right columns show four realizations from each of the two posteriors. Realizations from the posterior when using the Markov mesh prior are shown in the left column, whereas the realizations in the right column is based on a model with the profile Markov random field prior. The eight realizations are quite similar, but when studying them in more detail one can observe that with the Markov mesh prior there seems to be more skewed and curved structures than when using the profile Markov random field prior. Since the Markov mesh prior has much larger neighbourhoods than the profile Markov random field prior, this

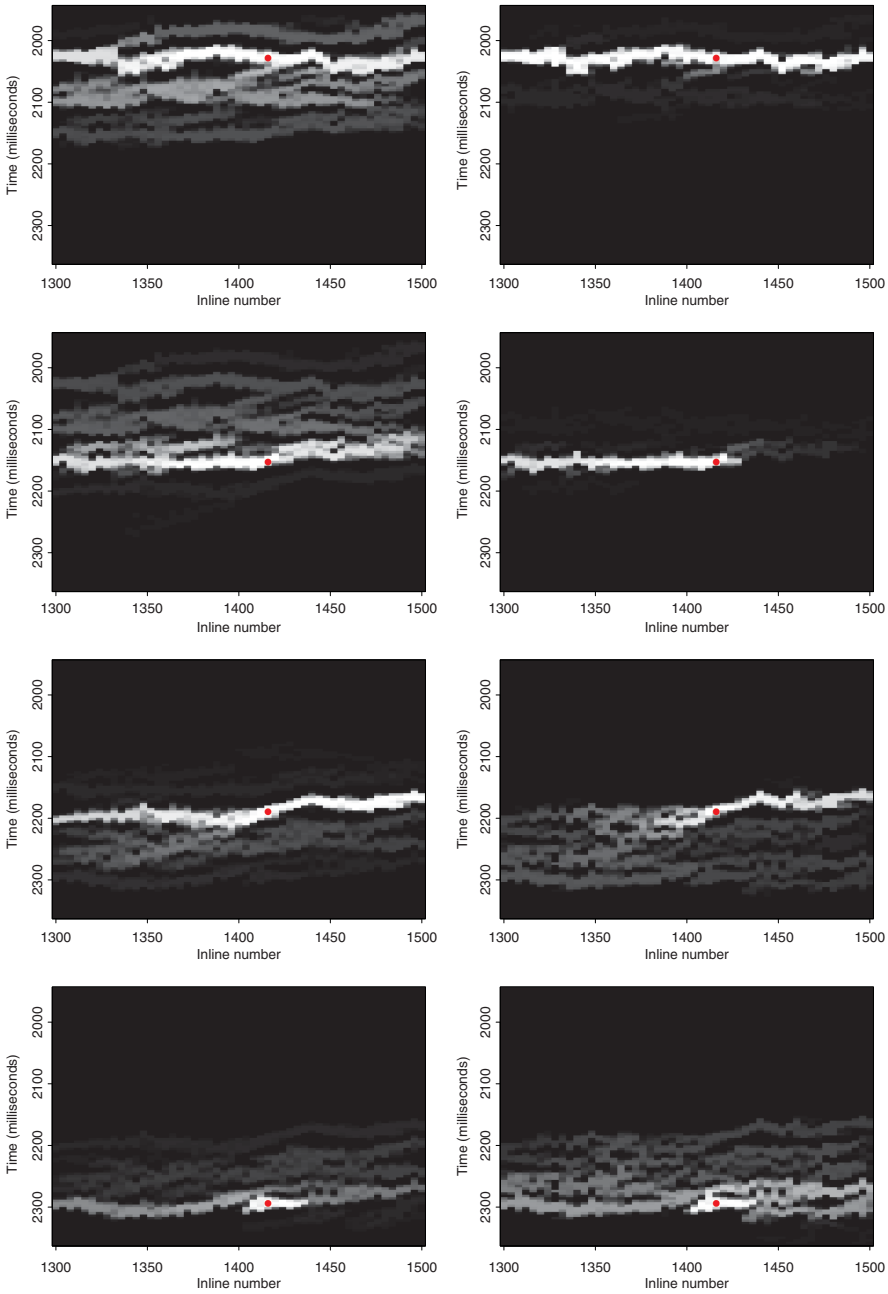


Figure 9 For each node, estimated posterior probabilities that there is contact (via other oil sand nodes) between this node and the node marked with a red filled circle. The plots in the left and right columns are when using the Markov mesh and the profile Markov random field priors, respectively. The colour scale is as defined by the legend in Fig. 7.

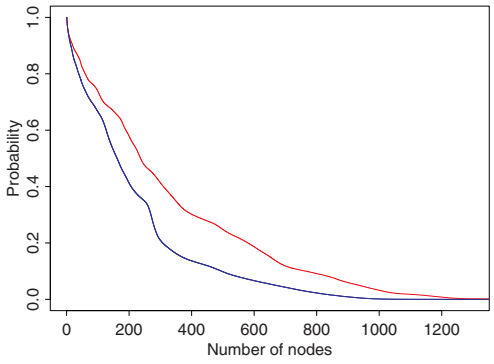


Figure 10 As a function of a number of nodes η, estimated posterior probability for a random oil sand node to be connected (via other oil sand nodes) to at least η other oil sand nodes. The η is along the x-axis and the estimated probability is along the y-axis. Results when using the Markov mesh and profile Markov random field priors are shown in red and blue, respectively.

is not really surprising. With larger neighbourhoods, and corresponding larger cliques, it becomes possible for the model to identify skewed and curved structures.

The upper row in Fig. 7 shows the result of estimating in each node the marginal posterior probability of oil sand. In each node the probability is estimated as the fraction of the realizations where the node has oil sand. Again the left and right images are results when using the Markov mesh and profile Markov random field priors, respectively. The two probability maps are similar, but somewhat more continuity of skewed and curved high probability areas can be observed when using the Markov mesh prior. In the middle row of Fig. 7, the probabilities in the upper row is rounded to the nearest integer to get an estimate of the most probable LFC in each node. Again we can observe somewhat more continuity of skewed and curved oil sand areas when using the Markov mesh prior. The histograms in the lower row of Fig. 7 are simply probability histograms of the estimated marginal posterior probabilities shown in the upper row of the same figure. We can observe that when using the profile Markov random field prior, somewhat more marginal posterior probabilities are close to zero and one than when using the Markov mesh prior.

To study the marginal probabilities a little more we have chosen three traces, or columns, $j = 15$, 30 and 45, and in Fig. 8 plotted the marginal probabilities. The estimated posterior marginal probabilities when using the Markov mesh and profile Markov random field priors are plotted in red and blue, respectively. More than in Fig. 7 we can here see how close the two posterior probabilities are for most of the nodes. In a few of the nodes, however, the difference is quite clear.

The continuity of oil sand is very important for fluid flow in a petroleum reservoir. We can get some understanding of how the prior influences this continuity by studying Fig. 7, but to study the continuity in more detail we need to summarize how this continuity is in each posterior realization. To do this, we have manually picked four nodes with very high posterior probability for oil sand both when using the Markov mesh and the profile Markov random field priors. These four nodes are marked with a red bullet in Fig. 9, one row for each of the four chosen nodes. For each of these four nodes and for each posterior realization we identified all other nodes with oil sand which through other oil sand nodes had contact with the chosen node. Thereby we could estimate the posterior probability that any node was in contact with the chosen node as the fraction of the realizations where this occurred. The resulting estimated probabilities are visualized in Fig. 9. The left and right columns are again the results when using the Markov mesh and the profile Markov random field priors, respectively. In the three upper rows we can see a lot more continuity in the posterior realizations when using the Markov mesh prior than when using the profile Markov random field prior. In the lower row the situation is for some reason reversed. To study this type of continuity more generally, not only for the four hand-picked nodes used in Fig. 9, we finally repeat the exercise of finding all nodes in a realization with oil sand connected to a particular node, but now the particular node is sampled at random among all nodes with oil sand. For each realization and each particular node we find the number of oil sand nodes connected to the particular node. In Fig. 10, we show the resulting estimated posterior probabilities for the randomly chosen particular node to be connected to more than η other oil sand nodes, as a function of η. The red and blue curves are the results when using the Markov mesh and the profile Markov random field priors, respectively. We see that the curve related to the Markov mesh prior lies consistently clearly above the curve related to the profile Markov random field prior, showing that the Markov mesh prior produces more posterior continuity of oil sand than the profile Markov random field prior.

DISCUSSION

In this article we have, for a particular seismic data set in the North Sea and two particular prior models for the lateral connectivity of the lithology/fluid classes (LFCs), studied how the prior influences the posterior properties. When focussing

on the posterior marginal probabilities we found, for most nodes, that the prior had little influence. When focussing on posterior continuity of oil sand, however, we found that the prior had a quite strong influence on the results. Not surprisingly, the prior with the largest neighbourhoods produced the largest posterior continuity. When evaluating whether to use a simple prior with a small neighbourhood or to use a more complicated prior with a larger neighbourhood one should therefore first decide what posterior properties that are of interest. If the focus is only on the posterior marginal probabilities, a simple prior is perhaps sufficient. If the focus is on fluid flow, however, spatial continuity is crucial and it may be beneficial with a more complicated prior which is better able to capture spatial continuity.

When deciding what prior to use one should also take into account the computational resources necessary to simulate from the posterior distribution. A more complicated prior typically gives a posterior which requires more computation time to explore. With our implementations, sampling from the posterior when using the Markov mesh prior required approximately 20 times more computation time compared to when using the profile Markov random field prior. However, our implementation of the sampling when using the Markov mesh prior was partly in Matlab and partly in C++ and a lot of the computation time here was just overhead in the communication between Matlab and C++. Our implementation of the sampling algorithm when using the profile Markov random field prior was entirely in Matlab, so we did not have the same overhead in this case. If we had implemented also the sampling algorithm when using the Markov mesh prior entirely in Matlab we expect this algorithm would have required a factor between 3 and 5 more computation time than that for the profile Markov random field prior. That the sampling when using the Markov mesh prior requires more computation time than when using the profile Markov random field prior should come as no surprise, since the Markov mesh prior has a much larger neighbourhood than the profile Markov random field prior.

The study presented in this article is quite limited. The model includes only two LFCs, we have considered only one seismic section, and we have not studied how sensitive the results are for the parameter values used in the two prior models. It is of interest to study the effect of using a prior with a larger neighbourhood also when the model represent more than two LFCs. The profile Markov random field prior is already defined with more than two LFCs and the Markov mesh construction used here can easily be extended to such a situation. The computational complexity of our Markov mesh fitting procedure will grow with the number of LFCs used, but we do not expect this increase to become a major problem. One should note that the training image we have used in the present study seems to have too much spatial continuity in that the realizations from the posterior shown in Fig. 6 have less spatial continuity than both the training image and the realizations from the Markov mesh prior shown in Fig. 5. If we had used a more realistic training image, for example a training image with spatial continuity similar to the posterior realizations in Fig. 6, we expect the convergence when fitting the Markov mesh model to the training image to have been quicker.

To condition on a seismic cube is also of interest. Then the LFCs need to be represented on a three-dimensional lattice and the prior models need to be defined for such a situation. Again the profile Markov random field prior is already formulated in such a situation. The Markov mesh formulation used here can also be extended to a three-dimensional lattice, but it remains to see whether it is computationally feasible to handle such a model. A computationally cheaper alternative would be to adopt a two-dimensional Markov mesh prior for the lithology/fluid values in each layer of the three-dimensional lattice and assume lithology/fluid values in the different layers to be *a priori* independent.

ACKNOWLEDGEMENTS

The authors acknowledge the Uncertainty in Reservoir Evaluation (URE) activity at the Norwegian University of Science and Technology (NTNU).

ORCID

Håkon Tjelmeland
https://orcid.org/0000-0001-9263-2504

REFERENCES

Abend K., Harley T. and Kanal L. 1965. Classification of binary random patterns. *IEEE Transactions on Information Theory* **11**, 538–544.

Aki K. and Richards P.G. 1980. *Quantitative Seismology: Theory and Methods*. W.H. Freeman, New York.

Arnesen P. and Tjelmeland H. 2017. Prior specification of neighbourhood and interaction structure in binary Markov random fields. *Statistics and Computing* **27**, 737–756.

Aster R., Borchers B. and Thurber C.H. 2011. *Parameter Estimation and Inverse Problems*, Elsevier, Amsterdam.

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

Connolly P. and Hughes M. 2016. Stochastic inversion by matching to large numbers of pseudo-wells. *Geophysics* **81**(2), M7–M22.

Cressie N. and Davidson J. 1998. Image analysis with partially ordered Markov models. *Computational Statistics and Data Analysis* **29**, 1–26.

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

Emery X. and Lantuéjoul C. 2014. Can a training image be a substitute for a random field model? *Mathematical Geosciences* **46**, 133–147.

Fjeldstad T. and Omre H. 2019. Bayesian inversion of convolved hidden Markov models with applications in reservoir prediction. *IEEE Transations on Geoscience and Remote Sensing*, to appear.

Gamerman D. and Lopes H.F. 2006. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, 2nd edn. Chapman & Hall/CRC, London.

Gilks W.R., Richardson S. and Spiegelhalter D.J. 1996. *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London.

Grabisch M., Marichal J.L. and Roubens M. 2000. Equivalent representations of set functions. *Mathematics of Operations Research* **25**, 157–178.

Grana D. and Della Rossa E. 2010. Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion. *Geophysics* **75**(3), O21–O37.

Grana D., Fjeldstad T. and Omre H. 2017. Bayesian Gaussian mixture linear inversion for geophysical inverse problems. *Mathematical Geosciences* **49**, 493–515.

Guardiano F. and Srivastava R. 1993. Multivariate geostatistics: beyond bivariate moments. In: *Geostatistics Tróia'92* (ed. A. Soares), pp. 133–144. Kluwer, Dordrecht.

Gunning J. and Glinsky M.E. 2007. Detection of reservoir quality using Bayesian seismic inversion. *Geophysics* **72**(3), R37–R39.

Hammer P.L. and Holzman R. 1992. Approximations of pseudo-Boolean functions; applications to game theory. *Methods and Models of Operation Research* **36**, 3–21.

Hurn M., Husby O. and Rue H. 2003. A tutorial on image analysis. In: *Spatial Statistics and Computational Methods, Lecture Notes in Statistics*, Vol. **173** (ed. J. Møller), pp. 87–139. Springer.

Journel J. and Zhang T. 2006. The necessity of a multiple-point prior model. *Mathematical Geology* **38**, 591–610.

Kindermann R. and Snell J.L. 1980. *Markov Random Fields and Their Applications*. American Mathematical Society, Providence, RI.

Lang X. and Grana D. 2017. Geostatistical inversion of prestack seismic data for the joint estimation of facies and impedances using stochastic sampling from Gaussian mixture posterior distribution. *Geophysics* **82**(4), M55–M65.

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**(5), R69–R78.

Luo X. and Tjelmeland H. 2019. Prior specification for binary Markov mesh models. *Statistics and Computing*, to appear.

Mariethoz G. and Caers J. 2014. *Multiple-Point Geostatistics: Stochastic Modeling with Training Images*. John Wiley & Sons.

Rimstad K., Avseth P. and Omre H. 2012. Hierarchical Bayesian lithology/fluid prediction: a North Sea case study. *Geophysics* **77**(2), B69–B856.

Rimstad K. and Omre H. 2010. Impact of rock-physics depth trends and Markov random fields on hierarchical Bayesian lithology/fluid prediction. *Geophysics* **75**(4), R93–R108.

Robert C.P. and Casella G. 1999. *Monte Carlo Statistical Methods*. Springer.

Sen M.K. and Stoffa P.L. 2013. *Global Optimization Methods in Geophysical Inversion*. Cambridge University Press.

Stien M. and Kolbjørnsen O. 2011. Facies modeling using a Markov mesh model specification. *Mathematical Geosciences* **43**, 611–624.

Strebelle S. 2002. Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology* **34**, 1–21.

Tarantola A. 2005. *Inverse Problem Theory*. SIAM, Philidelphia.

Toftaker H. and Tjelmeland H. 2013. Construction of binary multi-grid Markov random field prior models from training images. *Mathematical Geosciences* **45**, 383–409.

Ulvmoen M. and Omre H. 2010. Improved resolution in Bayesian lithology/fluid inversion from prestack seismic data and well observations: part 1-methodology. *Geophysics* **75**(2), R21–R35.

Zhang T., Pedersen S.I., Knudby C. and McCormick D. 2012. Memory-efficient categorical multi-point statistics algorithms based on compact search trees. *Mathematical Geosciences* **44**, 863–879.

APPENDIX: FITTED MARKOV MESH PRIOR, $p(\kappa)$

The Markov mesh model fitted to the training image in Fig. 2 has template sequential neighbourhood $\tau = \{(-1,0),\ (0,-1),(-1,2),(0,-2),(-3,-1),(0,-3),(-1,4),(0,-4),(-2,-4)\}$ and Λ and $\{\beta(\lambda) : \lambda \in \Lambda\}$ are as specified in Table A1.

Table A1 The elements λ in the set Λ and the associated interaction parameters $\beta(\lambda)$

$\lambda \in \Lambda$	$\beta(\lambda)$
$\emptyset$	−4.33884
{(−1,0)}	3.27479
{(0,−1)}	2.96595
{(−1,0),(0,−1)}	−0.460735
{(−1,2)}	1.49237
{(−1,2),(0,−1)}	−1.10759
{(0,−2)}	1.99035
{(−3,−1)}	−1.43573
{(0,−3)}	3.06786
{(−1,0),(0,−3)}	−3.44258
{(0,−3),(0,−1)}	−2.03335
{(−1,0),(0,−3),(0,−1)}	1.95605
{(0,−3),(0,−2)}	−1.02729
{(−1,4)}	2.90431
{(−1,0),(−1,4)}	−3.42674
{(−1,4),(0,−1)}	−0.404195
{(−1,2),(−1,4)}	0.268767
{(−1,4),(0,−3)}	−2.73426
{(−1,0),(−1,4),(0,−3)}	2.96929
{(−1,4),(0,−3),(0,−1)}	1.95346
{(0,−4)}	2.1858
{(−1,0),(0,−4)}	−0.355664
{(0,−4),(0,−2)}	−1.61185
{(0,−4),(0,−3)}	−1.23267
{(−1,0),(0,−4),(0,−3)}	0.606075
{(0,−4),(0,−3),(0,−2)}	2.03717
{(−1,4),(0,−4)}	−4.01512
{(−1,0),(−1,4),(0,−4)}	3.80173
{(−1,4),(0,−4),(0,−3)}	2.6053
{(−1,0),(−1,4),(0,−4),(0,−3)}	−1.64379
{(−2,−4)}	−0.717159

Paper V

A one-step Bayesian inversion framework for three-dimensional reservoir characterization based on a Gaussian mixture model – A Norwegian Sea demonstration

---

Torstein Fjeldstad, Per Åge Avseth and Henning Omre

Submitted

A one-step Bayesian inversion framework for three-dimensional reservoir characterization based on a Gaussian mixture model – A Norwegian Sea demonstration

Torstein Fjeldstad, Per Åge Avseth & Henning Omre

February 2020

Abstract

A one-step approach for Bayesian prediction and uncertainty quantification of lithology/fluid classes, petrophysical properties and elastic attributes conditional on prestack 3D seismic amplitude-versus-offset data is presented. A 3D Markov random field prior model is assumed for the lithology/fluid classes to ensure spatially coupled lithology/fluid class predictions in both the lateral and vertical directions. Conditional on the lithology/fluid classes, we consider Gauss-linear petrophysical and rock physics models including depth trends. Then, the marginal prior models for the petrophysical properties and elastic

attributes are multivariate Gaussian mixture models. The likelihood model is assumed to be Gauss-linear to allow for analytic computation. A recursive algorithm that translates the Gibbs formulation of the Markov random field into a set of vertical Markov chains is proposed. This algorithm provides an efficient proposal density in a Markov chain Monte Carlo algorithm such that simulation from the posterior model of interest in three dimensions is feasible. The model is demonstrated on real data from a Norwegian Sea gas reservoir. We evaluate the model at the location of a blind well, and we compare results from the proposed model with results from a set of 1D models, where each vertical trace is inverted independently. At the blind well location, we obtain an improvement of at most a 60 % reduction in the root mean square error for the proposed 3D model compared to the model without lateral spatial coupling.

Introduction

Quantitative interpretation (Avseth et al., 2005) of prestack seismic data is an essential part of reservoir characterization in the exploration phase to predict the proportion of hydrocarbon and to determine the well design for production. The objective is to reduce the technological and economic risk during the development phase. In reservoir characterization, there are three inverse problems (Tarantola, 2005): prediction of elastic attributes, such as P-impedance (seismic inversion); prediction of rock and fluid properties,

such as porosity and water saturation (petrophysical inversion); and prediction of the lithology/fluid classes (lithology/fluid classification). We refer to Grana et al. (2017) and references therein for a discussion of these inverse problems. Assessment of such geophysical inverse problems is a challenging problem due to the uncertainty in the measurements and nonuniqueness of the solution. There are several sources of uncertainty in the workflow, including observation errors, limited bandwidth of seismic data and rock physics modeling errors. We refer to Bosch et al. (2010) for an overview of seismic inversion.

Various deterministic (Aster et al., 2005; Sen and Stoffa, 2013) and probabilistic (Doyen, 1988; Lia et al., 1997; Tarantola, 2005) approaches exist to solve the abovementioned inverse problems. These approaches were first applied to seismic inversion but have lately been extended to also cover petrophysical inversion (Doyen, 2007). Deterministic techniques are often based on optimization of a misfit function including a penalty term for regularization of the solution. In a Bayesian setting, a prior probability model is assigned to the reservoir variables of interest in order to include prior knowledge and experience. Probabilistic approaches have been applied successfully for reservoir characterization; see, e.g., Doyen (1988); Lia et al. (1997); Mukerji et al. (2001); Buland and Omre (2003); Gunning and Glinsky (2007); Rimstad and Omre (2010); Grana and Della Rossa (2010); Rimstad et al. (2012); Jullum and Kolbjørnsen (2016) and Connolly and Hughes (2016).

We consider a Bayesian framework where the solution is not only a point prediction but rather the full posterior model for the variables of interest, which allows for uncertainty and risk quantification. The posterior model is computed by combining the likelihood for the observed data given the reservoir variables and the prior model for the latter. In general, the class of models that can be solved analytically is limited, and sampling-based methods such as Markov chain Monte Carlo have to be applied. In sampling-based methods, an ensemble of realizations represent the posterior model, on which summary statistics and predictions are based. However, construction of a satisfactory proposal density in the simulation algorithm that ensures sufficiently fast convergence is challenging in high-dimensional problems.

Traditionally, Bayesian inversion techniques are based on a stepwise procedure, where one first inverts for the elastic attributes (Doyen, 2007), then for the petrophysical properties and finally for the lithology/fluid classes. These methodologies are often applied at the pointwise level, where lateral spatial continuity in the predictions is inferred only from the spatial continuity of the seismic data. During the last decade, simultaneous (or joint/integrated/one-step) inversion techniques have been developed (Rimstad and Omre, 2010), and they capture the joint model and tend to represent the uncertainty more realistically. We consider the latter approach, where we jointly assess the posterior model of the variables of interest.

For a linear seismic model with an additive Gaussian error term and

a Gaussian prior model for the elastic attributes, the posterior model for the elastic attributes is also Gaussian, with analytic expressions for the mean vector and covariance matrix (Buland and Omre, 2003). Pointwise classification of the lithology/fluid classes based on classification techniques, such as discriminant analysis (Hastie et al., 2009) or other machine learning techniques, may then be applied to obtain a lithology/fluid classification. The prior model has recently been extended to also cover Gaussian mixture prior models (Grana and Della Rossa, 2010; Grana et al., 2017) to model multimodal and skewed marginal characteristics, and we consider the latter class of models.

Spatial histograms of well logs for the petrophysical properties, such as porosity and water saturation, often appear as multimodal and/or skewed due to varying lithology classes and fluid fillings of the subsurface. We therefore include a lithology/fluid class variable to model these variations, as they have an important impact on the petrophysical properties. To honor vertical sorting and ordering of the lithology/fluid classes, Markov chain prior models are frequently used. The usage of Markov chains to model a vertical profile dates back to Krumbein and Dacey (1969). These models are either used only for one-dimensional problems (Eidsvik et al., 2004; Connolly and Hughes, 2016; Fjeldstad and Omre, 2019) or used in coupling of vertical Markov chains in a 2D and 3D random field context (Ulvmoen and Omre, 2010; Rimstad and Omre, 2010; Fjeldstad and Grana, 2018; de Figueiredo et al., 2019). We extend this work by replacing the one-

dimensional Markov chain prior model for the lithology/fluid classes with a three-dimensional Markov random random field prior model, which allows for consistent three-dimensional modeling. We refer to Gunning and Sams (2018) for more information regarding the use of Markov random fields in reservoir characterization. The main advantage of phrasing a full 3D model is that each posterior realization in a vertical trace borrows predictive power from its neighboring traces. Hence, posterior realizations are expected to have larger lateral spatial continuity. If the ultimate objective is to forecast reservoir productions, reproducing lateral connectivity is of the utmost importance. We refer to Tjelmeland et al. (2019) for a discussion of the impact of lateral continuity in seismic inversion related to fluid flow.

Rock physics models relate the rock and fluid properties to the elastic attributes, and these relations are generally known for conventional reservoirs (Avseth et al., 2005; Mavko et al., 2009). Houck (2002) contains a discussion regarding the importance of considering both seismic and rock physics uncertainties, and in Bachrach (2006), stochastic rock physics models for joint prediction of porosity and saturation are presented. We consider the class of Gauss-linear models conditional on the lithology/fluid classes and the petrophysical properties and elastic attributes, where the forward model is assumed to be linear in the reservoir variable together with an additive zero-mean Gaussian error term.

The ultimate objective is to assess the reservoir variables of interest given seismic amplitude-versus-offset (AVO) data. We use a Markov chain Monte

Carlo Metropolis-Hastings algorithm to generate realizations from the correct posterior model of interest. Each iteration in the Metropolis-Hastings algorithm consists of two main steps: proposal of a new realization and accepting or rejecting the proposed sample. In high-dimensional problems, such as in reservoir characterization, the major challenge is to construct a reasonable proposal density in order to obtain satisfactory acceptance and convergence rates. Our focus is on construction of the proposal density in three dimensions, extending Rimstad and Omre (2013) and Fjeldstad and Omre (2019), and hence improving the convergence rate of the algorithm.

The main contribution of this paper is an algorithm that rephrases the complete Markov random field model in three dimensions into the set of corresponding conditional one-dimensional vertical Markov chain models. We present an efficient block-Gibbs algorithm in three dimensions based on analytic evaluations that may be used for sequential simulation. First, we define the probabilistic model in a Bayesian inversion setting and discuss assessment of the posterior model. Second, we demonstrate the methodology on a real Norwegian Sea case study. We consider seismic AVO data and refer to Avseth et al. (2016) for details of the reservoir. The results are validated at the location of a blind well and compared to an alternative model based on a collection of 1D models, where each vertical profile is assumed to be independent of the other profiles.

Model description

In the following section, we define the variables of interest and the probabilistic model, extending Fjeldstad and Grana (2018).

Notation

Denote by $p(\cdot)$ an arbitrary probability density/mass function (pdf). Vectors are given in lowercase bold font and matrices in uppercase bold font. Let $\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denote the (multivariate) Gaussian pdf for a random vector $\mathbf{x}$ having mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. We refer to a likelihood model as Gauss-linear if the modeling variable is linear in the conditioning variable together with an additive Gaussian error term. Let $\mathbb{R}_{[a,b]}$ denote the set of real numbers on the interval $[a, b]$, where $a < b$ are real numbers.

Consider a discretized grid of the subsurface, $\mathcal{L} = \{(x, y, t) : x = 1, \ldots, n_x; y = 1, \ldots, n_y; \quad t = 1, \ldots, n_t\}$, where x and y are the coordinates in the horizontal direction and t is the vertical position (e.g., time) indexed top-down, and let $n = n_x n_y n_t$ denote the total number of grid cells in the cube. Let $v = xyt \in \mathcal{L}$ denote an arbitrary cell and $-v$ denote all cells except v. Moreover, denote by $\mathcal{L}_{xy\cdot} = \{(x, y, t) : x, y, t = 1, \ldots, n_t\} \subset \mathcal{L}$ the set of vertical nodes at horizontal position xy and let $\mathcal{L}_{-xy\cdot} = \mathcal{L} \setminus \mathcal{L}_{xy\cdot}$, where $\setminus$ denotes the set difference. Finally, let $\mathbf{u} = (xy1, \ldots, xyn_t) \subset \mathcal{L}$ denote the indices of a full vertical profile at horizontal position xy.

The variables of interest on the grid $\mathcal{L}$ are the lithology/fluid classes $\boldsymbol{\kappa} = (\kappa_1, \ldots, \kappa_n)$, the petrophysical properties $\mathbf{r} = (r_1, \ldots, r_n)$ and the logarithm of the elastic attributes $\mathbf{m} = (m_1, \ldots, m_n)$. Moreover, let $r_v \in \Omega_r = \mathbb{R}_{[0,1]}$ and $m_v \in \Omega_m = \mathbb{R}_{[0,\infty)}$ denote the petrophysical property, taking on values on $[0, 1]$, and the elastic attribute, which takes positive values, respectively. Since rock properties such as the porosity, water saturation, permeability and net-to-gross are bounded on $\mathbb{R}^n_{[0,1]}$, we use an elementwise logit-transformation to ensure support on $\mathbb{R}^n$. The logit-transform is one-to-one, which ensures that we can transform back to the original domain. To ease notation, we specify in the following the probabilistic model only for one petrophysical property variable (porosity) and one elastic attribute variable (logarithm of P-impedance) at each grid cell $v \in \mathcal{L}$. The proposed methodology is valid for additional petrophysical properties and elastic attributes such as water saturation and/or the logarithm of the Poisson ratio at the expense of a more complex notation. Each node $v \in \mathcal{L}$ is assigned three stochastic variables κ_v, r_v and m_v. Let $\kappa_v \in \Omega_\kappa = \{1, \ldots, L\}$; hence, κ_v takes one out of L categorial values in each grid cell. These categorical values represent the lithology/fluid classes of the subsurface, for example, gas sandstone or shale. We consider prestack seismic amplitude-versus-offset (AVO) data $\mathbf{d} = (d_1, \ldots, d_n)$ and consider only one incidence angle in the following to ease notation.

Bayesian inversion

Recall that the variables of interest are the lithology/fluid classes $\boldsymbol{\kappa}$, the petrophysical properties $\mathbf{r}$ and the logarithm of the elastic attributes $\mathbf{m}$ of the subsurface. The objective is to characterize these variables given seismic AVO data in a joint spatial Bayesian inverse setting. A one-step joint methodology allows for a consistent treatment of the uncertainties, as these uncertainties are often underestimated in a stepwise procedure.

We operate in a Bayesian inversion framework, where the ultimate objective is to assess the joint posterior pdf of the variables of interest; that is,

$$\begin{aligned} p\big(\boldsymbol{\kappa}, \mathbf{r}, \mathbf{m} \mid \mathbf{d}\big) &\propto p\big(\mathbf{d} \mid \boldsymbol{\kappa}, \mathbf{r}, \mathbf{m}\big) p(\boldsymbol{\kappa}, \mathbf{r}, \mathbf{m}) \\ &= p\big(\mathbf{d} \mid \mathbf{m}\big) p\big(\mathbf{m} \mid \boldsymbol{\kappa}, \mathbf{r}\big) p\big(\mathbf{r} \mid \boldsymbol{\kappa}\big) p(\boldsymbol{\kappa}) \end{aligned} \tag{1}$$

since $\mathbf{m}$ is a canonical variable for $\mathbf{d}$. We refer to the likelihood model $p\big(\mathbf{d} \mid \mathbf{m}\big)$ as the seismic model. The joint prior model $p(\boldsymbol{\kappa}, \mathbf{r}, \mathbf{m})$ is sequentially decomposed, and we refer to $p(\boldsymbol{\kappa})$ as the lithology/fluid class model, $p\big(\mathbf{r} \mid \boldsymbol{\kappa}\big)$ as the petrophysical model and $p\big(\mathbf{m} \mid \boldsymbol{\kappa}, \mathbf{r}\big)$ as the rock physics model. The latter two models are defined conditional on the lithology/fluid classes $\boldsymbol{\kappa}$. Obtaining the normalizing constant of Equation (1) may not be computationally feasible since it requires a summation over $\boldsymbol{\kappa} \in \Omega_{\kappa}^{n}$ and evaluation of the high-dimensional integral over $\mathbf{r}$ and $\mathbf{m}$ to obtain the normalizing constant $p(\mathbf{d})$.

We assume the petrophysical model, rock physics model and likelihood

model to have the factorial form

$$p\big(\boldsymbol{\kappa}, \mathbf{r}, \mathbf{m} \mid \mathbf{d}\big) \propto \prod_{\mathbf{u}=1}^{n_x n_y} \Big[p\big(\mathbf{d_u} \mid \mathbf{m_u}\big) p\big(\mathbf{m_u} \mid \boldsymbol{\kappa_u}, \mathbf{r_u}\big) p\big(\mathbf{r_u} \mid \boldsymbol{\kappa_u}\big)\Big] \times p(\boldsymbol{\kappa}), \quad (2)$$

Hence, the observations $\mathbf{d_u}$ for each vertical trace are conditionally independent of the observations for every other vertical trace given $\boldsymbol{\kappa}$. Next, we define the petrophysical model, rock physics model and likelihood model for each vertical profile $\mathbf{u}$.

Likelihood model

The observed seismic signal d_v at grid cell $v = xyt \in \mathcal{L}$ can be represented as a convolution of the reflection coefficients along the vertical profile $\mathcal{L}_{xy\cdot}$ and a wavelet due to the dispersion of the seismic waves in the subsurface.

Seismic model

We consider a linearized approximation of the nonlinear Zoeppritz equations based on the Aki-Richards formulation for weak contrasts (Aki and Richards, 1980) following Buland and Omre (2003). We assume a Gauss-linear likelihood model

$$p\big(\mathbf{d_u} \mid \mathbf{m_u}\big) = \mathcal{N}\left(\mathbf{d_u}; \mathbf{G}\mathbf{m_u}, \boldsymbol{\Sigma}_{\mathbf{d_u}}\right); \quad (3)$$

that is, the model is linear in the conditioning variable $\mathbf{m}$ with additive Gaussian error terms. The linear operator $\mathbf{G}$ is a $(n_t \times n_t)$-matrix assumed

to be the product of three matrices $\mathbf{G} = \mathbf{WAD}$ (Buland and Omre, 2003), where $\mathbf{W}$ is the $(n_t \times n_t)$ convolution matrix, $\mathbf{A}$ is a $(n_t \times n_t)$-matrix containing the Aki-Richards coefficients and $\mathbf{D}$ is a $(n_t \times n_t)$-matrix approximating derivative. We assume that the $(n_t \times n_t)$-covariance matrix $\mathbf{\Sigma_{d_u}}$ includes vertically colored noise.

Prior model

The prior model in Equation (2) is specified sequentially, as in Fjeldstad and Grana (2018), and we discuss the lithology/fluid class, petrophysical property and elastic attributes models separately.

Lithology/fluid class model

To model the spatial connectivity and continuity of the lithology/fluid classes $\boldsymbol{\kappa}$, we consider a Markov random field model (Besag, 1974), which requires some additional notation. We consider the set of cliques $c \subset \mathcal{L}$, which consists of the pairs of closest neighbors. Let $\mathcal{C}$ denote the clique system that is the set of all cliques. Let $n_v \in \mathcal{L}$ be the set of neighbors of each $v \in \mathcal{L}$. Given the clique system of the closest pairwise cliques, it follows that n_v consists of the six closest neighbors, not including v itself, for each $v \in \mathcal{L}$. The methodology presented is also valid for a more complex clique set, at the expense of a more complex notation. We phrase the

lithology/fluid prior model in Gibbs form:

$$p(\boldsymbol{\kappa}) = \text{const}_1^{-1} \times \exp\left(-\sum_{c\in\mathcal{C}} g_c(\boldsymbol{\kappa}_c)\right), \tag{4}$$

where each $g_c(\cdot)$ is an arbitrary real-valued function. The normalizing constant in Equation (4), $\text{const}_1 = \sum_{\boldsymbol{\kappa}'\in\Omega_\kappa^n} \exp\big(-\sum_{c\in\mathcal{C}} g_c(\boldsymbol{\kappa}'_c)\big)$, is in general not analytically tractable, as it requires summation over $\boldsymbol{\kappa} \in \Omega_\kappa^n$, which has L^n elements.

The locationwise (conditional) Markov formulation is given as

$$\begin{aligned} p\big(\kappa_v \mid \boldsymbol{\kappa}_{-v}\big) &\propto \exp\left(-\sum_{\substack{c\in\mathcal{C}\\ v\in c}} g_c(\kappa_v, \boldsymbol{\kappa}_w; w \in c \setminus v)\right) \\ &= \text{const}_2^{-1} \times \exp\big(-h_v(\kappa_v \mid \kappa_w; w \in n_v)\big) \end{aligned} \tag{5}$$

where $h_v(\cdot)$ is a real-valued function defined by the Gibbs formulation and neighborhood system n_v. The normalizing constant is given by $\text{const}_2 = \sum_{\kappa'_v\in\Omega_\kappa} \exp\big(-h_v(\kappa'_v \mid \kappa_w; w \in n_v)\big)$, which is feasible to compute since it only requires a sum over $\kappa_v \in \Omega_\kappa$. The restriction $v \in c$ in Equation (5) implies that we need only consider the set of cliques associated with neighborhood n_v. Hence, the pdf including the normalizing constant is computationally tractable. Simulation from the prior $p(\boldsymbol{\kappa})$ defined in Equation (4) is often performed by Markov chain Monte Carlo simulation using the set of full-conditional pdfs $p\big(\kappa_v \mid \boldsymbol{\kappa}_{-v}\big)$ in Equation (5) for each $v \in \mathcal{L}$.

We consider a block-update Gibbs scheme where a subset of nodes is updated at each iteration. A joint update scheme for each vertical trace $\mathbf{u}$

is proposed. We rephrase the conditional pdf for each trace in sequential form as a vertical Markov chain:

$$\begin{aligned} p\big(\boldsymbol{\kappa}_{xy\cdot} \mid \boldsymbol{\kappa}_{-xy\cdot}\big) &= \prod_{t=1}^{n_t} p\Big(\kappa_{xyt} \mid \boldsymbol{\kappa}_{xy,1:(t-1)}, \boldsymbol{\kappa}_{-xy\cdot}\Big) \\ &= \prod_{t=1}^{n_t} p\Big(\kappa_{xyt} \mid \boldsymbol{\kappa}_{xy,(t-1)}, \boldsymbol{\kappa}_{-xy\cdot}\Big) \end{aligned}, \tag{6}$$

where $p\big(\kappa_{xy1} \mid \kappa_{xy0}, \boldsymbol{\kappa}_{-xy\cdot}\big) = p\big(\kappa_{xy1} \mid \boldsymbol{\kappa}_{-xy\cdot}\big)$ to ease notation. The latter equality of Equation (6) follows from the simpler first order neighborhood n_v. We propose a recursive algorithm to obtain the set of conditional pdfs $p\Big(\kappa_{xyt} \mid \boldsymbol{\kappa}_{xy,(t-1)}, \boldsymbol{\kappa}_{-xy\cdot}\Big)$.

We observe that $p\Big(\kappa_{xy,n_t} \mid \boldsymbol{\kappa}_{xy,(n_t-1)}, \boldsymbol{\kappa}_{-xy\cdot}\Big)$ is available directly from the locationwise Markov formulation in Equation (5). The locationwise Markov formulation for $t = n_t - 1, \ldots, 2$ is given as

$$\begin{aligned} p\big(\kappa_{xyt} \mid \boldsymbol{\kappa}_{-xyt}\big) &\propto p\Big(\kappa_{xyt} \mid \boldsymbol{\kappa}_{xy,1:(t-1)}, \boldsymbol{\kappa}_{-xy\cdot}\Big) \prod_{t'=t}^{n_t-1} p\Big(\kappa_{xy,(t'+1)} \mid \boldsymbol{\kappa}_{xy,1:t'}, \boldsymbol{\kappa}_{-xy\cdot}\Big) \\ &\propto p\Big(\kappa_{xyt} \mid \boldsymbol{\kappa}_{xy,(t-1)}, \boldsymbol{\kappa}_{-xy\cdot}\Big) p\Big(\kappa_{xy,(t+1)} \mid \boldsymbol{\kappa}_{xyt}, \boldsymbol{\kappa}_{-xy\cdot}\Big) \end{aligned}. \tag{7}$$

By rephrasing Equation (7), we obtain

$$p\Big(\kappa_{xyt} \mid \boldsymbol{\kappa}_{xy,(t-1)}, \boldsymbol{\kappa}_{-xy\cdot}\Big) \propto \frac{p\big(\kappa_{xyt} \mid \boldsymbol{\kappa}_{-xyt}\big)}{p\Big(\kappa_{xy,(t+1)} \mid \boldsymbol{\kappa}_{xyt}, \boldsymbol{\kappa}_{-xy\cdot}\Big)}. \tag{8}$$

The normalizing constant is tractable since it is a sum over $\kappa \in \Omega_\kappa$, which is feasible to compute. Since each pdf $p\Big(\kappa_{xyt} \mid \boldsymbol{\kappa}_{xy,(t-1)}, \boldsymbol{\kappa}_{-xy\cdot}\Big)$ depends

only on $p\big(\kappa_{xyt} \mid \boldsymbol{\kappa}_{-xyt}\big)$ and the previous iteration $p\Big(\kappa_{xy,(t+1)} \mid \boldsymbol{\kappa}_{xyt}, \boldsymbol{\kappa}_{-xy\cdot}\Big)$, one may use a recursive algorithm to obtain the transition probabilities in Equation (6) by iterating downwards $t = n_t - 1, \ldots, 2$. Finally, for $t = 1$, we obtain

$$p\big(\kappa_{xy1} \mid \boldsymbol{\kappa}_{-xy\cdot}\big) \propto \frac{p\big(\kappa_{xy1} \mid \boldsymbol{\kappa}_{-xy1}\big)}{p\big(\kappa_{xy2} \mid \kappa_{xy1}, \boldsymbol{\kappa}_{-xy\cdot}\big)}. \tag{9}$$

The resulting recursive algorithm is given in Algorithm 1.

Algorithm 1: Reverse algorithm for transition probabilities.

Result: Transition probabilities $\left\{p\Big(\kappa_{xyt} \mid \boldsymbol{\kappa}_{xy,(t-1)}, \boldsymbol{\kappa}_{-xy\cdot}\Big)\right\}$

$p\Big(\kappa_{xyn_t} \mid \boldsymbol{\kappa}_{xy,(n_t-1)}, \boldsymbol{\kappa}_{-xy\cdot}\Big) = p\big(\kappa_{xyn_t} \mid \boldsymbol{\kappa}_{-xyn_t}\big)$

for $t = n_t - 1$ **to** 2 **do**

  $p\Big(\kappa_{xyt} \mid \boldsymbol{\kappa}_{xy,(t-1)}, \boldsymbol{\kappa}_{-xy\cdot}\Big) = \text{const} \times \frac{p\left(\kappa_{xyt} \mid \boldsymbol{\kappa}_{-xyt}\right)}{p\left(\kappa_{xy,(t+1)} \mid \boldsymbol{\kappa}_{xyt}, \boldsymbol{\kappa}_{-xy\cdot}\right)}$

  $\text{const}^{-1} = \sum_{\kappa'_{xyt} \in \Omega_\kappa} p\Big(\kappa'_{xyt} \mid \boldsymbol{\kappa}_{xy,(t-1)}, \boldsymbol{\kappa}_{-xy\cdot}\Big)$

end

$p\big(\kappa_{xy1} \mid \boldsymbol{\kappa}_{-xy\cdot}\big) = \text{const} \times \frac{p\left(\kappa_{xy1} \mid \boldsymbol{\kappa}_{-xy1}\right)}{p\left(\kappa_{xy2} \mid \kappa_{xy1}, \boldsymbol{\kappa}_{-xy\cdot}\right)}$

$\text{const}^{-1} = \sum_{\kappa'_{xy1}} p\Big(\kappa'_{xy1} \mid \boldsymbol{\kappa}_{-xy\cdot}\Big)$

return $\left\{p\Big(\kappa_{xyt} \mid \boldsymbol{\kappa}_{xy,(t-1)}, \boldsymbol{\kappa}_{-xy\cdot}\Big)\right\}$.

Petrophysical model

We assume the petrophysical model to be Gaussian,

$$p\big(\mathbf{r_u} \mid \boldsymbol{\kappa_u}\big) = \mathcal{N}\left(\mathbf{r_u}; \boldsymbol{\mu}_{\mathbf{r_u}|\boldsymbol{\kappa_u}}, \boldsymbol{\Sigma}_{\mathbf{r_u}|\boldsymbol{\kappa_u}}\right), \tag{10}$$

with conditional n_t-mean-vector $\boldsymbol{\mu}_{\mathbf{r_u}|\boldsymbol{\kappa_u}} = \left(\mu_{r_{xy1}|\kappa_{xy1}}, \ldots, \mu_{r_{xyn_t}|\kappa_{xyn_t}}\right)$ and $(n_t \times n_t)$-covariance matrix $\boldsymbol{\Sigma}_{\mathbf{r_u}|\boldsymbol{\kappa_u}}$. The pointwise expected value $\mu_{r_{xyt}|\kappa_{xyt}}$ takes on one of L distinct values dependent on the value of $\kappa_v \in \Omega_\kappa$ and might depend on v to model fixed lateral and vertical trends such as compaction.

It can be demonstrated that the marginal multivariate pdf for the petrophysical properties is a spatially coupled Gaussian mixture pdf (see Fjeldstad and Omre (2019) and references therein):

$$p(\mathbf{r}) = \sum_{\boldsymbol{\kappa} \in \Omega_\kappa^n} p\big(\mathbf{r} \mid \boldsymbol{\kappa}\big) p(\boldsymbol{\kappa}). \tag{11}$$

That is, each univariate marginal pdf $p(r_v)$ is a Gaussian mixture pdf that can be used to model skewness and multimodality a priori to represent various lithology effects.

Rock physics model

Rock physics models are in general nonlinear but can be locally linearized (Landrø, 2001; Grana, 2016) or empirically fitted. We consider a probabilistic Gaussian lithology/fluid class-dependent rock physics model to represent

various lithology/fluid and petrophysical effects:

$$p\big(\mathbf{m_u} \mid \boldsymbol{\kappa_u}, \mathbf{r_u}\big) = \mathcal{N}\left(\mathbf{m_u}; \boldsymbol{\mu}_{\mathbf{m_u}|\boldsymbol{\kappa_u}} + \mathbf{B}_{\boldsymbol{\kappa_u}}\mathbf{r_u}, \boldsymbol{\Sigma}_{\mathbf{m_u}|\boldsymbol{\kappa_u}}\right), \tag{12}$$

where $\boldsymbol{\mu}_{\mathbf{m_u}|\boldsymbol{\kappa_u}} = \left(\mu_{m_{xy1}|\kappa_{xy1}}, \ldots, \mu_{m_{xyn_t}|\kappa_{xyn_t}}\right)$ is the n_t-vector of pointwise expected values for the elastic attributes similar as $\boldsymbol{\mu}_{\mathbf{r_u}|\boldsymbol{\kappa_u}}$, $\mathbf{B}_{\boldsymbol{\kappa_u}}$ is a $(n_t \times n_t)$-block-diagonal matrix with lithology/fluid class dependent coefficients, and $\boldsymbol{\Sigma}_{\mathbf{m_u}|\boldsymbol{\kappa_u}}$ is an $(n_t \times n_t)$-covariance matrix with colored noise. As for the petrophysical model, it is possible to include fixed depth-trends in the rock physics model for each lithology/fluid class.

The multivariate marginal pdf for the elastic attributes is a Gaussian mixture pdf:

$$p(\mathbf{m}) = \sum_{\boldsymbol{\kappa} \in \Omega_\kappa^n} \int p\big(\mathbf{m} \mid \boldsymbol{\kappa}, \mathbf{r}\big) p\big(\mathbf{r} \mid \boldsymbol{\kappa}\big) \mathrm{d}\mathbf{r} \times p(\boldsymbol{\kappa}) = \sum_{\boldsymbol{\kappa} \in \Omega_\kappa^n} p\big(\mathbf{m} \mid \boldsymbol{\kappa}\big) \times p(\boldsymbol{\kappa}). \tag{13}$$

We interpret the marginal rock physics model as a nonlinear model, where the model itself assigns the rock physics model marginally depending on the corresponding lithology/fluid class.

Posterior model

We present a block-Gibbs simulation algorithm to assess the joint posterior $p\big(\boldsymbol{\kappa}, \mathbf{r}, \mathbf{m} \mid \mathbf{d}\big)$ with a block update of the full vertical trace at horizontal position xy in each iteration. In general, the Markov chain Monte Carlo Metropolis-Hastings algorithm consists of two steps. First, there is a proposal step where a trace of updated variables is proposed, and then, there

is an accept-or-reject step where this trace is accepted with a certain probability. We discuss both parts in detail. Finally, we relate the proposed simulation algorithm to strategies proposed earlier.

First, initialize $\boldsymbol{\kappa}, \mathbf{r}$ and $\mathbf{m}$ with $p\big(\boldsymbol{\kappa}, \mathbf{r}, \mathbf{m} \mid \mathbf{d}\big) > 0$, denote by superscript i the current value of $\boldsymbol{\kappa}, \mathbf{r}$ and $\mathbf{m}$, and consider a random vertical trace $\mathbf{u}$. The proposed workflow consist of the following steps:

Proposal step: In each iteration, we consider the following block-dependent proposal density:

$$\begin{aligned} q\left(\boldsymbol{\kappa}_{\mathbf{u}}, \mathbf{r}_{\mathbf{u}}, \mathbf{m}_{\mathbf{u}} \mid \boldsymbol{\kappa}^i_{-\mathbf{u}}, \mathbf{r}^i_{-\mathbf{u}}, \mathbf{m}^i_{-\mathbf{u}}, \mathbf{d}\right) \propto\ & q\left(\boldsymbol{\kappa}_{\mathbf{u}} \mid \boldsymbol{\kappa}^i_{-\mathbf{u}}, \mathbf{d}_{\mathbf{u}}\right) \\ & \times p\Big(\mathbf{r}_{\mathbf{u}} \mid \boldsymbol{\kappa}_{\mathbf{u}}, \boldsymbol{\kappa}^i_{-\mathbf{u}}, \mathbf{r}^i_{-\mathbf{u}}\Big) \\ & \times p\Big(\mathbf{m}_{\mathbf{u}} \mid \boldsymbol{\kappa}_{\mathbf{u}}, \mathbf{r}_{\mathbf{u}}, \boldsymbol{\kappa}^i_{-\mathbf{u}}, \mathbf{r}^i_{-\mathbf{u}}, \mathbf{m}^i_{-\mathbf{u}}\Big) \end{aligned} . \tag{14}$$

To simulate from the block proposal density, we perform the following steps. First, we construct the Markov chain $p\big(\boldsymbol{\kappa}_{\mathbf{u}} \mid \boldsymbol{\kappa}^i_{-\mathbf{u}}\big)$ using the recursion defined in Algorithm 1. The approximate pdf $q\left(\boldsymbol{\kappa}_{\mathbf{u}} \mid \boldsymbol{\kappa}_{-\mathbf{u}}, \mathbf{d}_{\mathbf{u}}\right) \propto \widetilde{p}\left(\mathbf{d}_{\mathbf{u}} \mid \boldsymbol{\kappa}_{\mathbf{u}}\right) p\big(\boldsymbol{\kappa}_{\mathbf{u}} \mid \boldsymbol{\kappa}^i_{-\mathbf{u}}\big)$ is then exactly assessed by Algorithm 1. Here, $\widetilde{p}\left(\mathbf{d}_{\mathbf{u}} \mid \boldsymbol{\kappa}_{\mathbf{u}}\right)$ is an approximation to the exact likelihood model $p\big(\mathbf{d}_{\mathbf{u}} \mid \boldsymbol{\kappa}_{\mathbf{u}}\big)$ (Fjeldstad and Omre, 2019). Note that the algorithm presented in Algorithm 1 can be used directly to assess $q\left(\boldsymbol{\kappa}_{\mathbf{u}} \mid \boldsymbol{\kappa}_{-\mathbf{u}}, \mathbf{d}_{\mathbf{u}}\right)$ by including a term dependent on $\mathbf{d}$, $\widetilde{p}\left(\mathbf{d}_{\mathbf{u}} \mid \boldsymbol{\kappa}_{\mathbf{u}}\right)$, in Equation (4). Because of the convolution, the simpler first-order neighborhood system is extended to a higher-order neighborhood sys-

tem, and the algorithm has to be modified accordingly. Finally, we propose $\boldsymbol{\kappa}_{\mathbf{u}} \sim q\left(\boldsymbol{\kappa}_{\mathbf{u}} \mid \boldsymbol{\kappa}^i_{-\mathbf{u}}, \mathbf{d}_{\mathbf{u}}\right)$, $\mathbf{r}_{\mathbf{u}} \sim p\big(\mathbf{r}_{\mathbf{u}} \mid \boldsymbol{\kappa}_{\mathbf{u}}, \boldsymbol{\kappa}^i_{-\mathbf{u}}, \mathbf{r}^i_{-\mathbf{u}}\big)$ and $\mathbf{m}_{\mathbf{u}} \sim p\big(\mathbf{m}_{\mathbf{u}} \mid \boldsymbol{\kappa}_{\mathbf{u}}, \mathbf{r}_{\mathbf{u}}, \boldsymbol{\kappa}^i_{-\mathbf{u}}, \mathbf{r}^i_{-u}, \mathbf{m}^i_{-u}\big)$. Here, laterally smooth realizations of $\mathbf{r}_u$ are constructed by assuming a joint multivariate Gaussian pdf for $\mathbf{r}_u$ and its four neighboring traces in $\mathbf{r}^i_{-\mathbf{u}}$ together with a spatial correlation coefficient ϱ_r, which acts as a spatial smoother or regularizer in the horizontal direction. Note that the marginal expectations and variances are specified from the petrophysical model; this implies that an analytic expression for $p\big(\mathbf{r}_{\mathbf{u}} \mid \boldsymbol{\kappa}_{\mathbf{u}}, \boldsymbol{\kappa}^i_{-\mathbf{u}}, \mathbf{r}^i_{-\mathbf{u}}\big)$ is available. We construct $p\big(\mathbf{m}_{\mathbf{u}} \mid \boldsymbol{\kappa}_{\mathbf{u}}, \mathbf{r}_{\mathbf{u}}, \boldsymbol{\kappa}^i_{-\mathbf{u}}, \mathbf{r}^i_{-\mathbf{u}}, \mathbf{m}^i_{-\mathbf{u}}\big)$ similarly by correlating the vertical profile $\mathbf{m}_{\mathbf{u}}$ with its four nearest neighboring traces in $\mathbf{m}^i_{-\mathbf{u}}$ by assuming a spatial correlation coefficient ϱ_m between neighboring vertical traces to construct a joint pdf.

Accept/reject step: The proposed values of $\boldsymbol{\kappa}_{\mathbf{u}}, \mathbf{r}_{\mathbf{u}}$ and $\mathbf{m}_{\mathbf{u}}$ are accepted with probability

$$\alpha = \min\left\{1, \frac{p\big(\mathbf{d}_{\mathbf{u}} \mid \mathbf{m}_{\mathbf{u}}\big) p\big(\mathbf{m}_{\mathbf{u}} \mid \boldsymbol{\kappa}_{\mathbf{u}}, \mathbf{r}_{\mathbf{u}}\big) p\big(\mathbf{r}_{\mathbf{u}} \mid \boldsymbol{\kappa}_{\mathbf{u}}\big)}{p\big(\mathbf{d}_{\mathbf{u}} \mid \mathbf{m}^i_{\mathbf{u}}\big) p\big(\mathbf{m}^i_{\mathbf{u}} \mid \boldsymbol{\kappa}^i_{\mathbf{u}}, \mathbf{r}^i_{\mathbf{u}}\big) p\big(\mathbf{r}^i_{\mathbf{u}} \mid \boldsymbol{\kappa}^i_{\mathbf{u}}\big)} \times \frac{\widetilde{p}\left(\mathbf{d}_{\mathbf{u}} \mid \boldsymbol{\kappa}^i_{\mathbf{u}}\right) p\big(\mathbf{r}^i_{\mathbf{u}} \mid \boldsymbol{\kappa}^i_{\mathbf{u}}, \boldsymbol{\kappa}^i_{-\mathbf{u}}, \mathbf{r}^i_{-\mathbf{u}}\big) p\big(\mathbf{m}^i_{\mathbf{u}} \mid \boldsymbol{\kappa}^i_{\mathbf{u}}, \mathbf{r}^i_{\mathbf{u}}, \boldsymbol{\kappa}^i_{-\mathbf{u}}, \mathbf{r}^i_{-\mathbf{u}}, \mathbf{m}^i_{-\mathbf{u}}\big)}{\widetilde{p}\left(\mathbf{d}_{\mathbf{u}} \mid \boldsymbol{\kappa}_{\mathbf{u}}\right) p\big(\mathbf{r}_{\mathbf{u}} \mid \boldsymbol{\kappa}_{\mathbf{u}}, \boldsymbol{\kappa}^i_{-\mathbf{u}}, \mathbf{r}^i_{-\mathbf{u}}\big) p\big(\mathbf{m}_{\mathbf{u}} \mid \boldsymbol{\kappa}_{\mathbf{u}}, \mathbf{r}_{\mathbf{u}}, \boldsymbol{\kappa}^i_{-\mathbf{u}}, \mathbf{r}^i_{-\mathbf{u}}, \mathbf{m}^i_{-\mathbf{u}}\big)}\right\}. \tag{15}$$

These steps are performed until convergence, and the result is an ensemble of realizations from the posterior $p\big(\boldsymbol{\kappa}, \mathbf{r}, \mathbf{m} \mid \mathbf{d}\big)$. After convergence, the realizations are combined to construct marginal summary statistics such as

marginal probabilities for the lithology/fluid classes and marginal maximum a-posteriori (MMAP) predictors for the continuous-valued properties.

Alternative sampling strategies

The accept-or-reject step increases the computational complexity because it requires evaluation of a high-dimensional Gaussian pdf $p\big(\mathbf{d_u} \mid \boldsymbol{\kappa_u}\big)$. To reduce the computational complexity, one alternative is to omit the accept-or-reject step and only consider an approximate posterior solution $\tilde{p}\left(\boldsymbol{\kappa}, \mathbf{r}, \mathbf{m} \mid \mathbf{d}\right)$. If we propose $\mathbf{r_u} \sim p\big(\mathbf{r_u} \mid \boldsymbol{\kappa_u}\big)$ and $\mathbf{m_u} \sim p\big(\mathbf{m_u} \mid \boldsymbol{\kappa_u}, \mathbf{r_u}\big)$, the acceptance rate defined in Equation (15) can be further simplified to reduce the computational complexity (Rimstad and Omre, 2010). Another alternative is to also condition on $\mathbf{d_u}$ when proposing to update $\mathbf{r_u}$ and $\mathbf{m_u}$, that is, to simulate $\mathbf{r_u} \sim p\big(\mathbf{r_u} \mid \boldsymbol{\kappa_u}, \boldsymbol{\kappa}^i_{-\mathbf{u}}, \mathbf{r}^i_{-\mathbf{u}}, \mathbf{d_u}\big)$ and $\mathbf{m_u} \sim p\big(\mathbf{m_u} \mid \boldsymbol{\kappa_u}, \mathbf{r_u}, \boldsymbol{\kappa}^i_{-\mathbf{u}}, \mathbf{r}^i_{-\mathbf{u}}, \mathbf{m}^i_{-\mathbf{u}}, \mathbf{d_u}\big)$, which also increases the computational demand. In our experience, the mode indicator $\boldsymbol{\kappa_u}$ appears to be far more important for the mixing and convergence rates in the Markov chain Monte Carlo algorithm. There are other sampling strategies such as a moving-window update (see de Figueiredo et al. (2017)) or considering a model where each vertical profile is treated independently of all other traces (Connolly and Hughes, 2016). Finally, an extension of Fjeldstad and Omre (2019) is to first generate an ensemble of realizations from $p\big(\boldsymbol{\kappa} \mid \mathbf{d}\big)$ by iteratively sampling $\boldsymbol{\kappa_u} \sim q\left(\boldsymbol{\kappa_u} \mid \boldsymbol{\kappa}^i_{-\mathbf{u}}, \mathbf{d_u}\right)$ and accepting each realization with

probability $\min\left\{1, \frac{p(\mathbf{d_u}|\boldsymbol{\kappa_u})}{p(\mathbf{d_u}|\boldsymbol{\kappa}_\mathbf{u}^i)} \times \frac{\widetilde{p}(\mathbf{d_u}|\boldsymbol{\kappa}_\mathbf{u}^i)}{\widetilde{p}(\mathbf{d_u}|\boldsymbol{\kappa_u})}\right\}$ based on the current value $\boldsymbol{\kappa}_\mathbf{u}^i$. Afterwards, exact realizations from $p(\mathbf{r} \mid \mathbf{d})$ and $p(\mathbf{m} \mid \mathbf{d})$ are obtained since the Gaussian mixture prior pdfs defined in Equations (11) and (13) are conjugate priors for a Gauss-linear likelihood model (Grana et al., 2017).

Results from a Norwegian Sea case study

We demonstrate the proposed methodology on a Norwegian Sea gas discovery; see Avseth et al. (2016) for details. The 3D seismic data consist of broadband prestack time-migrated and normal-moveout-corrected gathers from a survey covering the target area. We condition on near- and far-angle prestack AVO data and invert for the three distinct lithology/fluid classes: brine sandstone, gas sandstone and shale. Moreover, we include the porosity ϕ, the water saturation s_w, the clay volume/proportion c, and the elastic attributes $\log \rho V_P$ (log P-impedance) and $\log V_P/V_S$ (log Poisson ratio). The data cover a domain discretized onto a grid with $98 \times 75 \times 100 = 735,000$ cells. The observations are sampled regularly in the depth domain at every 4 ms and cover a lateral domain of approximately 3 km $\times$ 3 km. Note that the dimension of the variable space, being the number of spatially coupled univariate posterior pdfs to assess, is $6 \times 735,000 = 4,410,000$, and the dimension of the data space is $2 \times 735,000 = 1,470,000$. The domain contains one well, which we use for blind well validation.

Figure 1 displays the near- and far-angle prestack seismic AVO data

for a 2D section containing the blind well. We observe the seismic AVO measurements to have a fairly long ranged horizontal spatial dependency. In Figure 2, we display the variables of interest observed in the blind well together with the observed seismic signal and the set of synthetic seismic signals. We display spatial histograms of the continuous variables of interest and observe the variables to appear as either unimodal, multimodal or skewed due to varying lithology effects. The well-tie between the observed seismic and synthetic seismic signal, based on the likelihood parameters specified below, is reasonable. The corresponding correlations are 0.70 and 0.89 for the near and far angles, respectively. In Figure 3, we display the near- and far-angle wavelets. The signal-to-noise ratio is set to 2.5 for the near stack and 1.5 for the far stack, where most of the colored noise depends on the convolutional model.

The prior model parameters defined earlier are empirically calibrated based on a well outside the target area. Figure 3 displays a subset of the empirically calibrated rock physics model: $\log \rho V_P$ against porosity and lithology/fluid class for a fixed value of water saturation and clay volume. The rock physics model is assumed to be a linear model dependent on both the porosity and the lithology/fluid class. We assume vertical squared exponential spatial correlation functions for the petrophysical properties and elastic properties, both having a range parameter equal to 2. The marginal prior pdfs defined in Equation (11) and Equation (13) will be discussed later. The correlation of $\log \rho V_p$ and $\log V_P/V_S$ is set to 0.8, to -0.5 for

porosity and water saturation, to -0.65 for porosity and clay volume, and to 0.65 for water saturation and clay volume. The prior Markov random field is specified such that the marginal probabilities are $(0.5, 0.15, 0.35)$ for shale, gas and brine, respectively, together with an anisotropic spatial interaction in the depth and horizontal directions. Finally, we set $\varrho_r = \varrho_m = 0.5$ in the simulation algorithm.

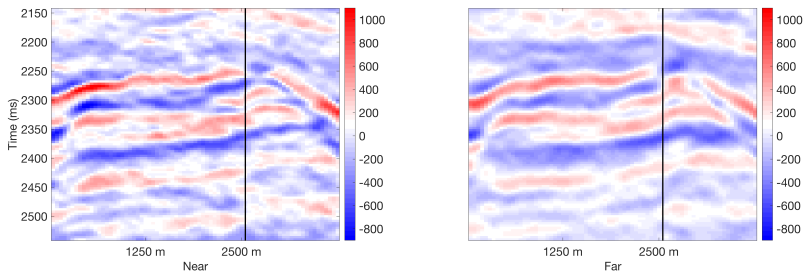


Figure 1: 2D vertical cross-section with near- and far-angle AVO observations. Here, red indicates a positive amplitude (hard event) and blue indicates a negative amplitude (soft event). The location of the blind well is indicated by the solid vertical line.

We compare the results based on the proposed methodology to the results based on a set of trace-independent models, with each vertical profile being independent of every other vertical profile; see Connolly and Hughes (2016). The lithology/fluid class prior model in this trace-independent model is assumed to follow a vertical Markov chain downwards with tran-

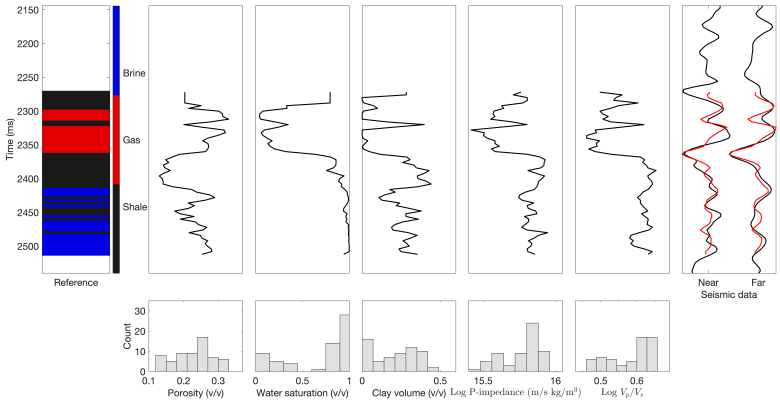


Figure 2: 1D blind well observations together with observed and synthetic seismic observations. Top row from left to right: reference classification (shale in black, gas in red and brine in blue), porosity, water saturation, clay volume, $\log \rho V_P$ and $\log V_P/V_S$, and observed (black) and synthetic (red) seismic observations. Bottom row: marginal histograms of the observed variables of interest.

sition matrix

$$\mathbf{P} = \begin{pmatrix} 0.7 & 0.1 & 0.2 \\ 0.2 & 0.7 & 0.1 \\ 0.3 & 0 & 0.7 \end{pmatrix}, \tag{16}$$

which has marginal distribution $(0.47, 0.16, 0.37)^\intercal$. This marginal distribution is comparable to the one of the Markov random field prior in the proposed model. We refer to the two models as the 3D model for the proposed spatially-coupled model and the 1D model for the trace-independent

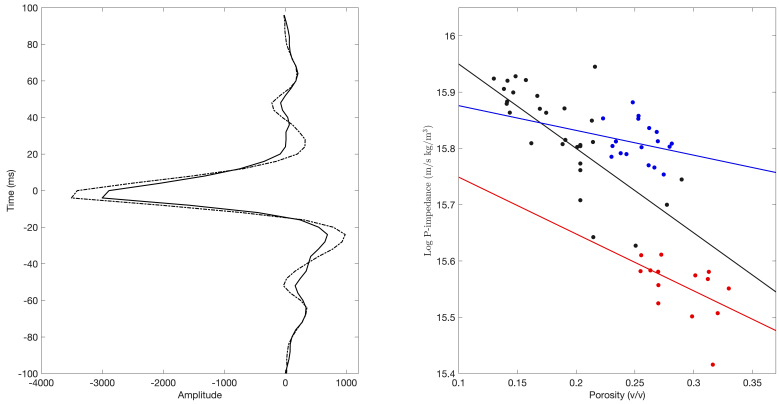


Figure 3: 1D wavelet and rock physics model. From left to right: near-angle (solid) and far-angle (dashed) wavelets and $\log \rho V_P$ against porosity and lithology/fluid class (shale in black, gas in red and brine in blue).

model. We generated $10,000$ approximately independent posterior realizations after thinning ($100,000$ before thinning) for both models. For the 3D model, we obtain an average acceptance rate of 27.8 % across all vertical traces (min. 6.3 % and max. 64.5 %); correspondingly, we obtain 14.2 % (min. 2.0 % and max. 69.5 %) for the 1D model. The trace-wise acceptance rate is dependent on the similarity of the proposal density in the simulation algorithm and the correct posterior model (Fjeldstad and Omre, 2019). On a 12-core shared university workstation from mid-2016, the computational requirement for the 3D model is 24 hours, and it is 7 hours for the 1D model. The mixing and convergence properties appear to be satisfactory

based on trace plots, not presented here, with convergence after only a few full updates of the full 3D grid.

To validate and compare the two models, we display the resulting posterior pdfs for the variables of interest at the blind well location. Figure 4 contains the reference lithology/fluid classification together with a set of posterior realizations, the marginal posterior pdf for the lithology/fluid classes and the MMAP predictor for the lithology/fluid classes. Overall, the two models have similar marginal posterior characteristics; however, the 1D model fails to identify the thin top gas reservoir in the MMAP predictor. Additionally, the MMAP predictor based on the 3D model predicts the bottom gas zone to be thicker than that for the 1D model. Indeed, the thin shale-layers around depth 2450 ms are not identified in the MMAP predictors, which may not be surprising since predictions are known to be more homogeneous than in reality. Note that these shale layers appear in the posterior realizations. In summary, introducing a lateral spatial dependence into the probabilistic model appears to help to identify thin layers.

Figure 5 displays the posterior pdfs for the porosity, water saturation and clay volume for the two models at the blind well location. For each property, we display the log-prior pdf together with the log-posterior models based on the 3D model and the 1D model to ease interpretation. The prior models for the petrophysical properties and elastic attributes need not be identical in the two models because of the different lateral coupling in the lithology/fluid class model, but we expect them to be very similar. For

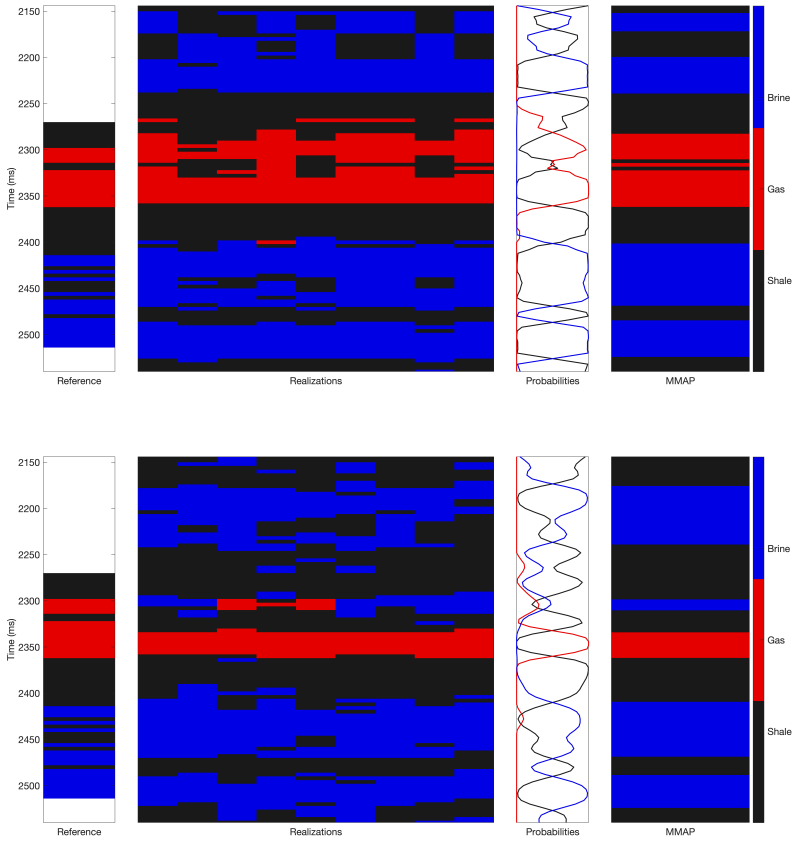


Figure 4: 1D posterior results lithology/fluid classes at the blind well location. The top row is based on results for the 3D model, while the bottom row is for the 1D model. Each row consists of, from left to right, the reference classification based on the blind well, a set of posterior realizations, the marginal probability profiles, and the marginal maximum posterior predictor.

simplicity, we display the prior based on the 1D model. The marginal prior at each depth point for porosity is close to unimodal but skewed, while the marginal prior at each depth point is bimodal, with two modes close to 0 and 1 for water saturation. In general, the MMAP predictors based on the 3D model and the 1D model are similar; however, the latter fails to capture the high-porosity zone and low-water-saturation zone around depth 2315 ms, where the top gas reservoir is located. The MMAP predictors are generally observed to be more homogeneous than the true profile. In general, we observe the posterior models based on the 3D model to have greater variability than those based on the 1D model, i.e., they have a larger pointwise uncertainty. Both models are able to satisfactorily capture the rapid transition from low to high water saturation at approximately 2360 ms, which corresponds to the boundary of the lower gas reservoir.

Figure 6 displays in a similar format as Figure 5 the posterior pdfs for $\log \rho V_P$ and $\log V_P/V_S$. We have included a depth trend in the prior for the elastic attributes with a higher expected response for $\log \rho V_P$ at the bottom of the target zone than on top. Again, we observe the 3D model to have a more realistic level of uncertainty compared to the observed well measurements.

The marginal posterior pdfs need not be unimodal but rather may be both multimodal and skewed; see Figure 7. The pdfs in the figure are chosen at time depths to represent a variety of the posterior models, which are observed to be unimodal, skewed and multimodal. Note the large dis-

crepancy between the posterior models at locations 2300 ms and 2324 ms, which corresponds to the top reservoir and the upper part of the lowermost reservoir at the blind well location. Note that the marginal modes for the posterior pdfs for the petrophysical properties and elastic attributes are strongly dependent on the corresponding marginal posterior pdf for the lithology/fluid classes, which acts as the weights in the Gaussian mixture model. We interpret the smaller marginal variances for the 1D model to be the result of a bias-variance trade-off, where we have obtained a smaller marginal variance at the cost of a biased predictor.

In Table 1, we present the mean absolute error (MAE) and root mean square error (RMSE) for the two models at the blind well location. Except for the clay volume, we obtain an improvement of up to 59 % for the MAE and 62 % for the RMSE for the variables of interest.

Table 1: Mean absolute error and root mean square error for the petrophysical properties and elastic attributes for the two models at the blind well location.

	MAE		RMSE	
	3D	1D	3D	1D
Porosity	**0.0326**	0.0376	**0.0402**	0.0484
Water saturation	**0.0667**	0.1629	**0.1124**	0.2986
Clay volume	0.1014	**0.0819**	0.1353	**0.1130**
Log ρV_P	**0.0498**	0.0822	**0.0636**	0.0822
Log V_P/V_S	**0.0244**	0.0325	**0.0351**	0.0438

In Figure 8 through Figure 10, we display posterior results for the 2D section corresponding to the seismic data presented in Figure 1. In Figure 8, the marginal probabilities for the three lithology/fluid classes and the corresponding MMAP predictor for the two models are presented. In general, the marginal posterior characteristics are observed to be similar; however, the lateral connectivity is larger for the 3D model. Since the MMAP predictor is a marginal property, the lateral connectivity does not need to be preserved in the predictor. Figures 9 and 10 display the MMAP predictor for the petrophysical properties and elastic attributes, respectively. For the 3D model, we observe the gas zones to be thicker and the predictions to be more homogeneous.

Next, we present the results for the variables of interest in a three-dimensional view based on the two models. In Figure 11, we present the MMAP predictor for the lithology/fluid classes based on the two different models. In general, we observe the MMAP predictor based on the 3D model to have a larger spatial connectivity in the lateral directions, as expected. The main characteristics of the sand bodies are observed in both models, but the MMAP predictor based on the 1D model appears to be less smooth and include abrupt invalid geophysical transitions. One advantage of the more complex 3D model is that each posterior realization appears with stronger spatial connectivity, which resembles more geologically plausible scenarios such as elongated sand bodies and channels that allow fluid flow (Figure 12). In Figure 13, we display the ISO-50 cube for gas (more than 50 % probability

of gas) based on the two models. The 3D model is observed to have a far greater lateral extent of the gas reservoir, while the 1D model is observed to have less connectivity. A similar conclusion is drawn from Figure 14, where we display a set of horizontal slices for the marginal probability of gas for three consecutive horizontal layers. The main characteristics are shared across the two different models; however, the marginal probabilities based on the 3D model appear smoother, as expected. Note that the marginal probabilities are smooth across the time slices.

In Figure 15 and Figure 16, we display the MMAP predictor for the petrophysical properties and elastic attributes in a 3D perspective based on the two models. They share the main characteristics, but the predictors based on the 3D model appear more laterally connected and smoother than those based on the 1D model.

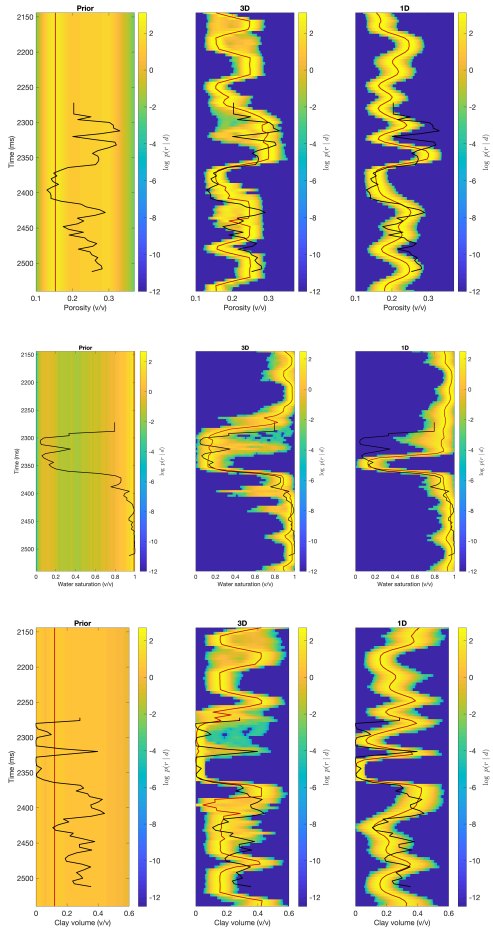


Figure 5: 1D posterior results petrophysical properties at the blind well location. Each row consists of, from left to right, the marginal log-prior pdf, marginal log-posterior pdf based on the 3D model, and marginal log-posterior pdf based on the 1D model. In each plot, the observed well log (black line) is displayed together with the marginal mode (red line).

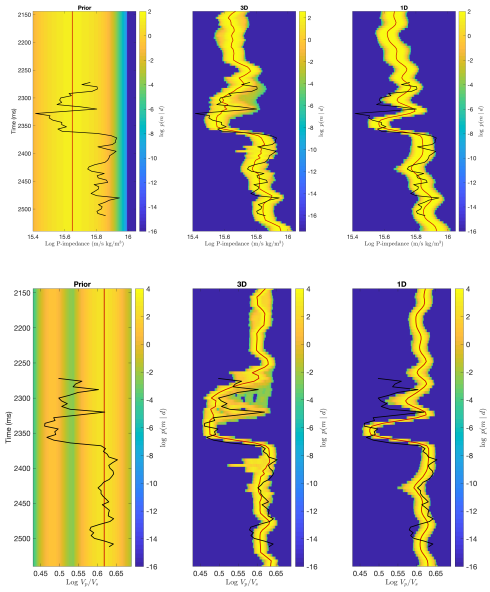


Figure 6: 1D posterior results elastic attributes at the blind well location. Each row consists of, from left to right, the marginal log-prior pdf, marginal log-posterior pdf based on the 3D model, and marginal log-posterior pdf based on the 1D model. In each plot, the observed well log (black line) is displayed together with the marginal mode (red line).

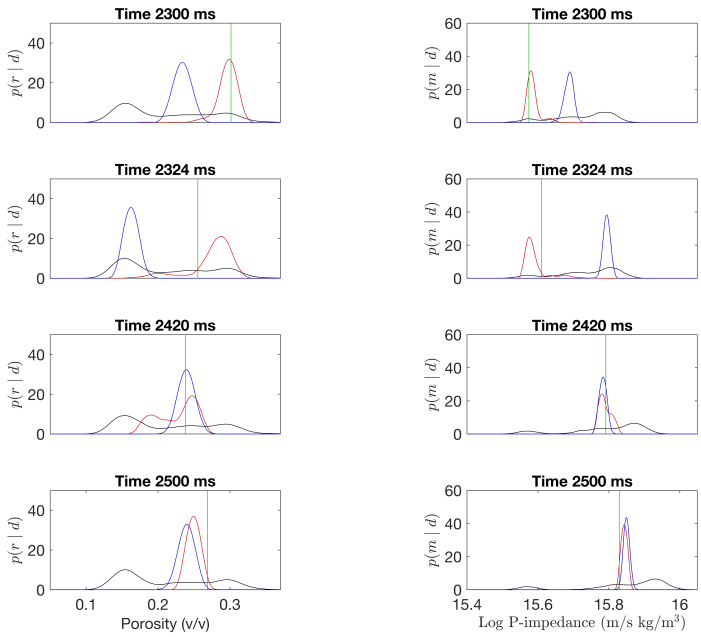


Figure 7: 1D posterior densities at the blind well location for porosity (left column) and log ρV_P (right column) for four chosen time points. Each plot consists of the prior (solid black), 1D model posterior pdf (blue line), and 3D model posterior pdf (red line), together with the observed value at the blind well location (vertical green line).

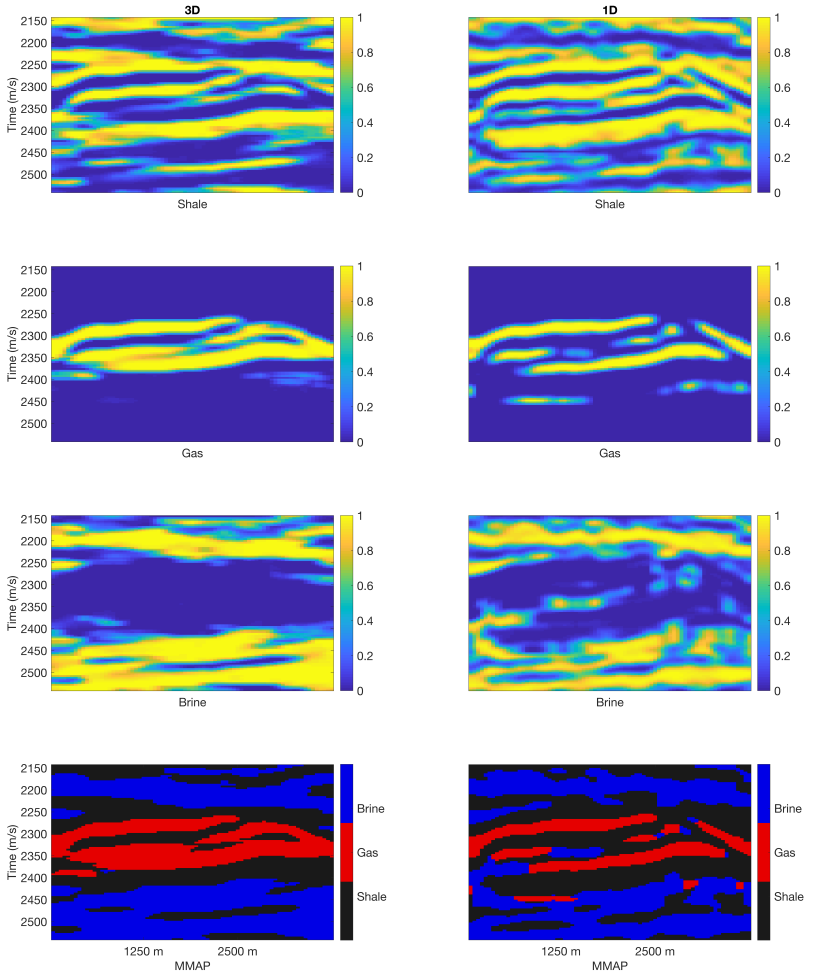


Figure 8: 2D posterior marginal densities for the lithology/fluid classes. From top to bottom: marginal probability shale, marginal probability gas, marginal probability brine and MMAP predictor. The results based on the 3D model are given in the leftmost column, and the results based on the 1D model are given on the right.

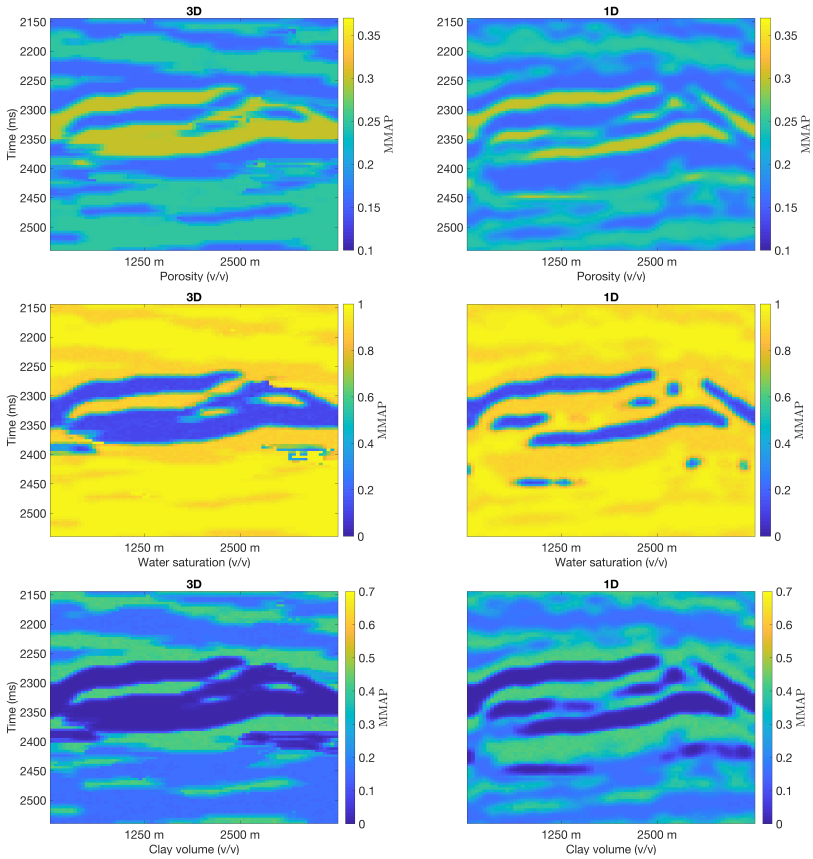


Figure 9: 2D posterior densities for the petrophysical properties. From top to bottom: MMAP predictor porosity, MMAP predictor water saturation and MMAP predictor clay volume. The results based on the 3D model are given in the leftmost column, and the results based on the 1D model are given on the right.

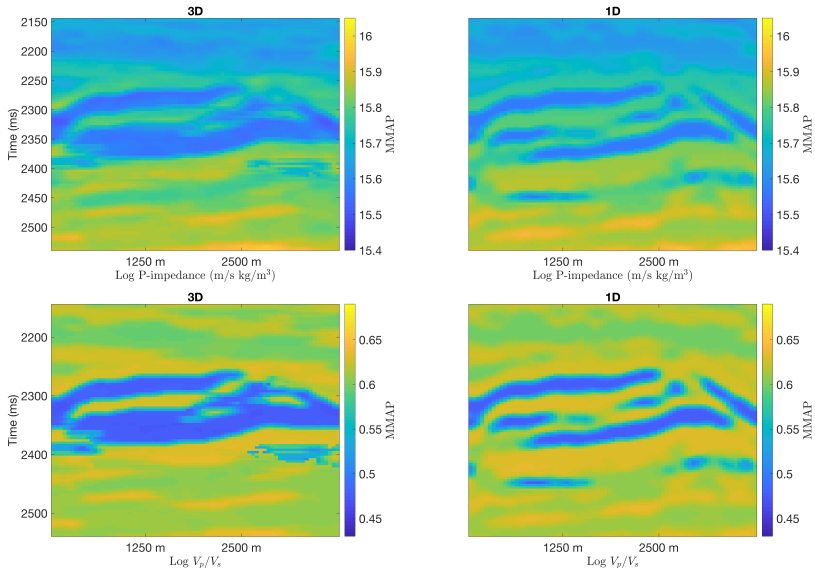


Figure 10: 2D posterior densities for the elastic attributes. From top to bottom: MMAP predictor $\log \rho V_P$ and MMAP predictor $\log V_P/V_S$. The results based on the 3D model are given in the leftmost column, and the results based on the 1D model are given on the right.

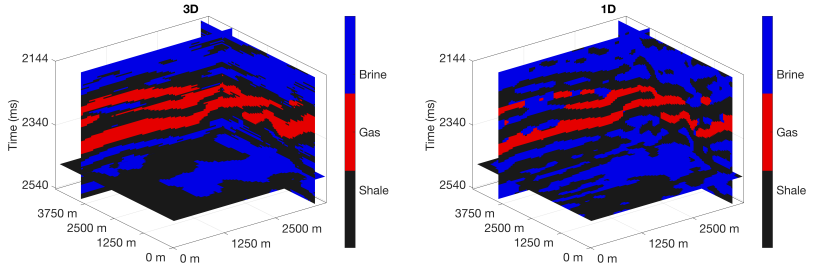


Figure 11: 3D MMAP predictor for the lithology/fluid classes based the 3D model (right) and 1D model (left).

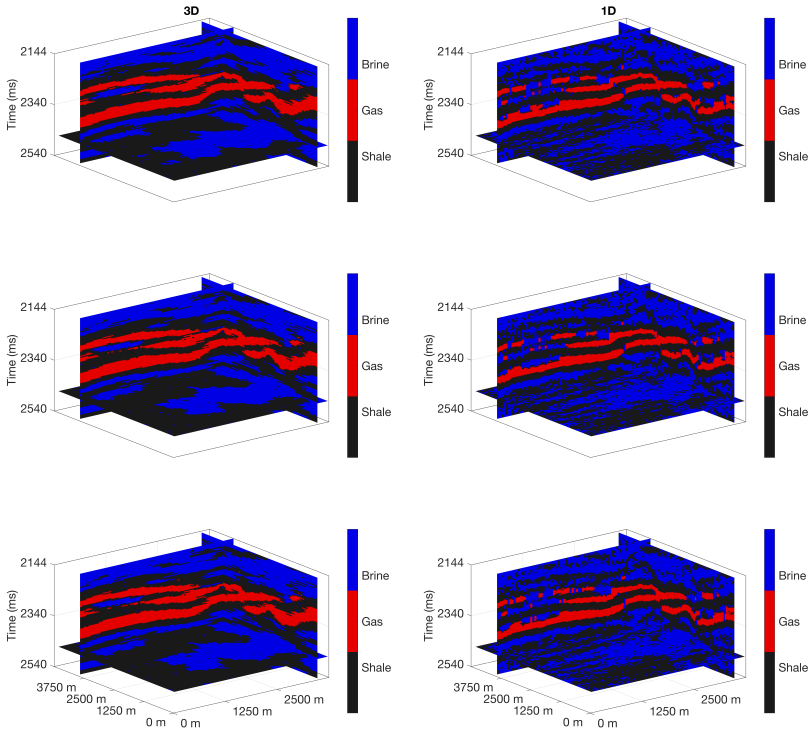


Figure 12: 3D posterior realizations for the lithology/fluid classes for the 3D model (right) and 1D model (left).

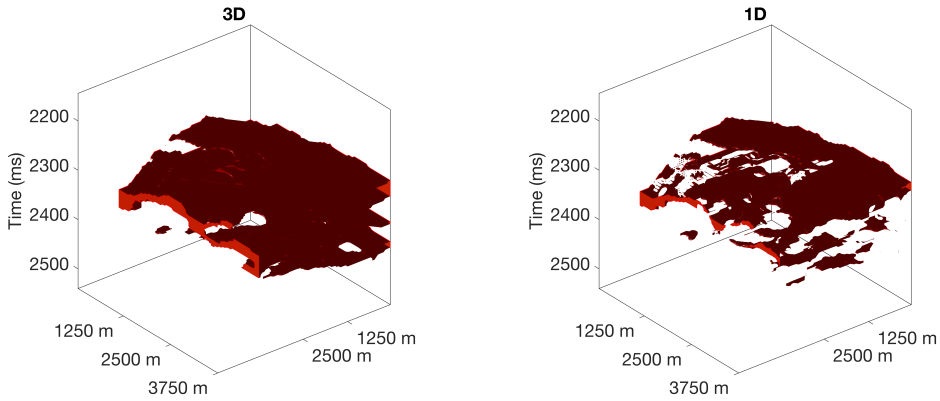


Figure 13: 3D ISO-50 probability for gas. From left to right: the results based on the 3D model and the 1D model.

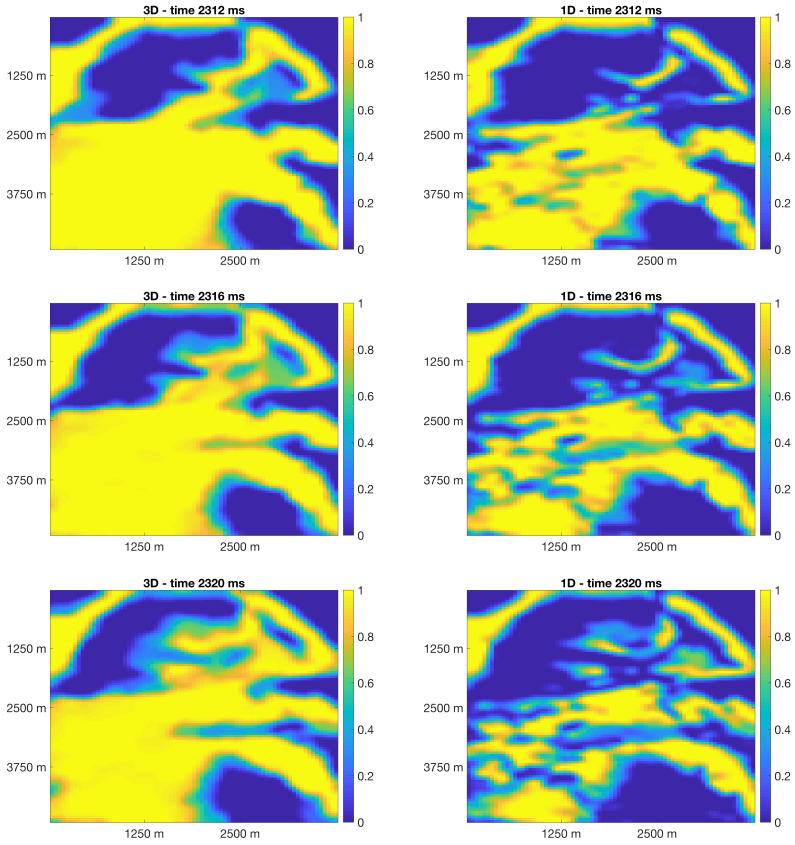


Figure 14: 3D posterior marginal probability of gas for three consecutive horizontal slices. From left to right: the results based on the 3D model and the 1D model.

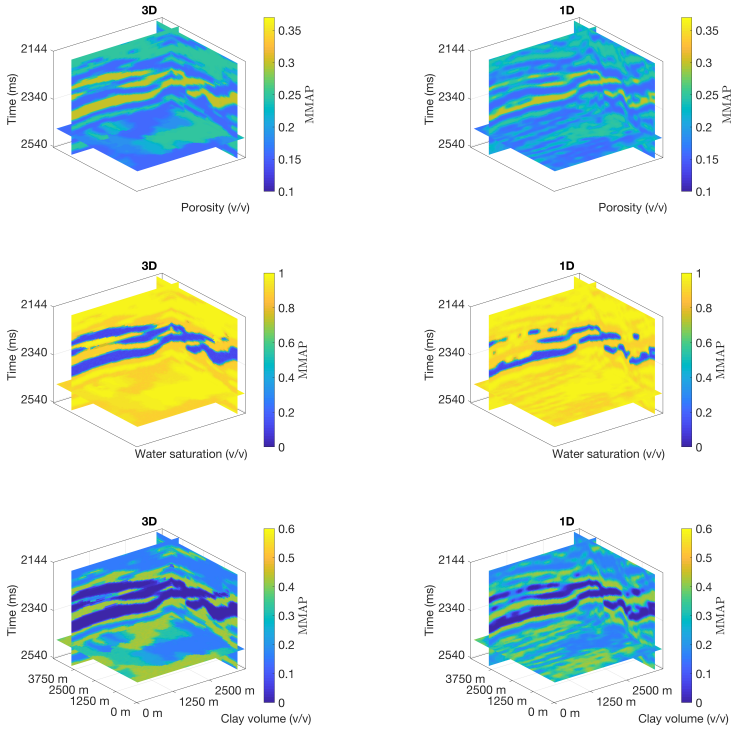


Figure 15: 3D MMAP predictor for the petrophysical properties. From left to right: the results based on the 3D model and the 1D model. From top to bottom: the results given for the porosity, water saturation and clay volume.

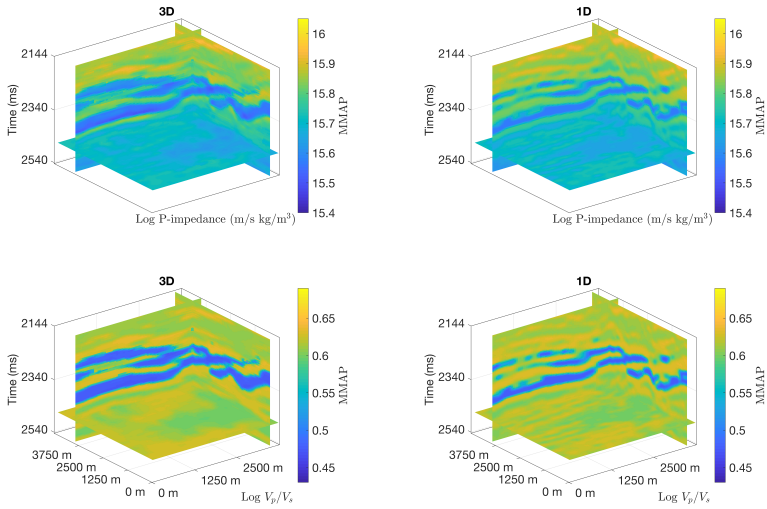


Figure 16: 3D MMAP predictor for the elastic attributes. From left to right: the results based on the 3D model and the 1D model. The top row includes $\log \rho V_P$ and the bottom row includes $\log V_P/V_S$.

Discussion

We have evaluated the proposed model using a real 3D case study and compared it with a model consisting of a set of 1D models without spatial lateral dependency. The lithology/fluid classes, petrophysical properties and elastic attributes are observed to be predicted with realistic spatial continuity in three dimensions. The marginal posterior densities for the continuous-valued properties capture multimodality and skewness, a property observed in the well logs of the blind well. The mean square error predictions in the blind well are reduced by at most 60 % by using the proposed 3D model compared to a set of 1D models.

In practice, seismic images are heavily processed images of the subsurface, where the vertical traces already have been migrated to match its neighboring traces by some criterion. Since prestack AVO data are strongly laterally connected, the 1D model will also tend to give posterior predictions that are laterally smooth to some extent; however, each realization does not need to satisfy this property. Since the 3D model has predictive power in neighboring traces, thin laterally extending layers are more easily identified than those in the 1D model. However, this advantage comes at the potential cost of oversmoothing.

In our experience, it is challenging to specify the model parameters in the Gibbs formulation in such a way that it does not become too dominant relative to the likelihood. Finally, the 3D model is less prone to misaligned

vertical traces and/or misspecification of the likelihood model because it acts as a spatial smoother.

However, the computational demand for the 3D model is severe compared to that of the 1D model since each Markov chain Monte Carlo iteration requires a rerun of the reverse algorithm conditional on the current value of the neighboring traces. Higher memory usage is also required because larger parts of the 3D cube need to be stored in memory in each iteration. We note that the 1D model is easily computed in parallel, while the 3D model requires a larger computational overhead to be parallelized. In our experience, the 1D model is feasible on a regular laptop, while the 3D model requires a workstation or cluster to be feasible.

Conclusions

We propose a one-step block-Gibbs update scheme for joint probabilistic prediction of lithology/fluid classes, petrophysical properties and elastic attributes in three dimensions. The proposed methodology is demonstrated on seismic AVO data from a Norwegian Sea discovery and is validated at a blind well position. Realistic lateral spatial connectivity is obtained in both realizations and predictions for the variables of interest for the 3D model. Compared to a 1D model based on inverting the set of vertical profiles independently, we obtain an increased average acceptance rate in the simulation algorithm and a significant reduction in the mean absolute error and root

mean square error at the blind well location.

Future research should include joint statistical inference of the model parameters in the Gibbs formulation together with the variable prediction to avoid oversmoothing.

Acknowledgments

The authors acknowledge the Uncertainty in Reservoir Evaluation (URE) initiative at the Norwegian University of Science and Technology (NTNU) for funding this research and PGS for providing the seismic data set.

References

Aki, K. and Richards, P. (1980). *Quantitative Seismology: Theory and Methods*. W. H. Freeman and Co., New York.

Aster, R., Thurber, C., and Borchers, B. (2005). *Parameter Estimation and Inverse Problems*. Number v. 1 in International Geophysics series. Elsevier Academic Press.

Avseth, P., Janke, A., and Horn, F. (2016). AVO inversion in exploration — Key learnings from a Norwegian Sea prospect. *The Leading Edge*, 35(5):405–414.

Avseth, P., Mukerji, T., and Mavko, G. (2005). *Quantitative Seismic In-*

terpretation: Applying Rock Physics Tools to Reduce Interpretation Risk. Cambridge University Press.

Bachrach, R. (2006). Joint estimation of porosity and saturation using stochastic rock-physics modeling. *Geophysics*, 71(5):O53–O63.

Besag, J. (1974). Spatial Interaction and the Statistical Analysis of Lattice Systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, 36(2):pp. 192–236.

Bosch, M., Mukerji, T., and Gonzalez, E. F. (2010). Seismic inversion for reservoir properties combining statistical rock physics and geostatistics: A review. *Geophysics*, 75:75A165–75A176.

Buland, A. and Omre, H. (2003). Bayesian linearized AVO inversion. *Geophysics*, 68(1):185–198.

Connolly, P. A. and Hughes, M. J. (2016). Stochastic inversion by matching to large numbers of pseudo-wells. *Geophysics*, 81:M7–M22.

de Figueiredo, L. P., Grana, D., Roisenberg, M., and Rodrigues, B. B. (2019). Gaussian mixture markov chain monte carlo method for linear seismic inversion. *Geophysics*, 84(3):R463–R476.

de Figueiredo, L. P., Grana, D., Santos, M., Figueiredo, W., Roisenberg, M., and Neto, G. S. (2017). Bayesian seismic inversion based on rock-physics

prior modeling for the joint estimation of acoustic impedance, porosity and lithofacies. *Journal of Computational Physics*, 336(C):128–142.

Doyen, P. (2007). *Seismic reservoir characterization: an earth modelling perspective*. Education tour series. EAGE publications.

Doyen, P. M. (1988). Porosity from seismic data: A geostatistical approach. *Geophysics*, 53(10):1263–1275.

Eidsvik, J., Mukerji, T., and Switzer, P. (2004). Estimation of Geological Attributes from a Well Log: An Application of Hidden Markov Chains. *Mathematical Geology*, 36(3):379–397.

Fjeldstad, T. and Grana, D. (2018). Joint probabilistic petrophysics-seismic inversion based on gaussian mixture and markov chain prior models. *Geophysics*, 83(1):R31–R42.

Fjeldstad, T. and Omre, H. (2019). Bayesian inversion of convolved hidden markov models with applications in reservoir prediction. *IEEE Transactions on Geoscience and Remote Sensing*, pages 1–12.

Grana, D. (2016). Bayesian linearized rock-physics inversion. *Geophysics*, 81:D625–D641.

Grana, D. and Della Rossa, E. (2010). Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion. *Geophysics*, 75:O21–O37.

Grana, D., Fjeldstad, T., and Omre, H. (2017). Bayesian Gaussian Mixture Linear Inversion for Geophysical Inverse Problems". *Mathematical Geosciences*, 49(4):493–515.

Gunning, J. and Glinsky, M. E. (2007). Detection of reservoir quality using Bayesian seismic inversion. *Geophysics*, 72:R37–R49.

Gunning, J. and Sams, M. (2018). Joint facies and rock properties bayesian amplitude-versus-offset inversion using markov random fields. *Geophysical Prospecting*, 66(5):904–919.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data mining, Inference, and Prediction*. Springer Series in Statistics. Springer, New York, 2nd edition.

Houck, R. T. (2002). Quantifying the uncertainty in an AVO interpretation. *Geophysics*, 67(1):117–125.

Jullum, M. and Kolbjørnsen, O. (2016). A Gaussian-based framework for local Bayesian inversion of geophysical data to rock properties. *Geophysics*, 81:R75–R87.

Krumbein, W. C. and Dacey, M. F. (1969). Markov chains and embedded Markov chains in geology. *Mathematical Geology*, 1:79–96.

Landrø, M. (2001). Discrimination between pressure and fluid saturation changes from time-lapse seismic data. *Geophysics*, 66(3):836–844.

Lia, O., Omre, H., Tjelmeland, H., Holden, L., and Egeland, T. (1997). Uncertainty in reservoir production forecasts. *AAPG Bulletin*, 81(5).

Mavko, G., Mukerji, T., and Dvorkin, J. (2009). *The Rock Physics Handbook*. Cambridge University Press, second edition.

Mukerji, T., Jørstad, 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.

Rimstad, K., Avseth, P., and Omre, H. (2012). Hierarchical bayesian lithology/fluid prediction: A north sea case study. *GEOPHYSICS*, 77(2):B69–B85.

Rimstad, K. and Omre, H. (2010). Impact of rock-physics depth trends and Markov random fields on hierarchical Bayesian lithology/fluid prediction. *Geophysics*, 75:R93–R108.

Rimstad, K. and Omre, H. (2013). Approximate posterior distributions for convolutional two-level hidden Markov models. *Computational Statistics & Data Analysis*, 58:187–200.

Sen, M. K. and Stoffa, P. L. (2013). *Global Optimization Methods in Geophysical Inversion*. Cambridge University Press.

Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics.

Tjelmeland, H., Luo, X., and Fjeldstad, T. (2019). A bayesian model for lithology/fluid class prediction using a markov mesh prior fitted from a training image. *Geophysical Prospecting*, 67(3):609–623.

Ulvmoen, M. and Omre, H. (2010). Improved resolution in Baeysian lithology/fluid inversion from seismic prestack data and well observations: Part I - Methodology. *Geophysics*, 75:R21–R35.