Doctoral theses at NTNU, 2021:395

Ole Bernhard Forberg

Bayesian inversion of seismic data using multimodal selection Gaussian prior models

Doctoral thesis

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

NTNU
Norwegian University of Science and Technology

Ole Bernhard Forberg

Bayesian inversion of seismic data using multimodal selection Gaussian prior models

Thesis for the Degree of Philosophiae Doctor

Trondheim, November 2021

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 978-82-326-6231-9 (printed ver.)
ISBN 978-82-326-5505-2 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2021:395

Printed by NTNU Grafisk senter

Bayesian inversion of seismic data using multimodal selection Gaussian prior models

Ole Bernhard Forberg

August 22, 2021

Preface

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

I would like to thank my supervisor at the Department of Mathematical Sciences at NTNU, Professor Henning Omre, for his guidance throughout the last four years. Thank you for the knowledge you have imparted to me, which includes but is not limited to statistical topics. I appreciate that we could converse about various topics and that you always were willing to share your point of view.

I would also like to thank my co-supervisor Øyvind Kjøsnes at Aker BP for his guidance and for the collaboration. Thank you for sharing your expertise on geophysics and seismic reservoir characterization, it helped better my understanding of the subjects. I would also like to extend a special thanks for your added availability while working on our most recent paper.

Finally, I would like to thank Assistant Professor Dario Grana at the University of Wyoming (UW) for the collaboration and the opportunity to visit UW. I really appreciate the hospitality you showed during my visit, it made my stay at UW a delight. I also appreciate your willingness to share your knowledge and for being there whenever I had a question.

Trondheim, August 2021 Ole Bernhard Forberg

Thesis Outline

Background

Introduction

The success of a mathematical description of the world has led to an abundance of inverse problems in science and engineering. An inverse problem can be associated with any mathematical equation that describes the causal influence of a variable $\mathbf{r} \in \mathcal{R}$ on an observable variable $\mathbf{d} \in \mathcal{D}$, and this type of equation is termed a forward model. Here, the sets $\mathcal{R}$ and $\mathcal{D}$ are taken to be Hilbert spaces. In many applications a forward model can be theoretically based and captured by a function $f : \mathcal{R} \to \mathcal{D}$; hence, the forward model can be expressed

$$\left[\mathbf{d}|\mathbf{r}\right] = f(\mathbf{r}). \tag{1}$$

Computing the observation $\mathbf{d}$ for a given configuration of the causal variable $\mathbf{r}$ is the forward problem associated with the forward model, whereas computing $\mathbf{r}$ for a given $\mathbf{d}$ is the inverse problem.

An inverse problem is said to be well-posed if a unique solution exists and the solution changes continuously with variations in the observations (Hadamard, 1902, 1923). Well-posed problems are desirable because they in principle can be solved exactly by stable algorithms. Problems that are not well-posed are said to be ill-posed, and inverse problems of this type are often approached by regularization techniques. Regularization techniques impose restrictions on the solution that make the regularized problem well-posed. Tikhonov regularization is a classical regularization technique (Tikhonov, 1963) that offers a solution $\hat{\mathbf{r}}$ to ill-posed problems through minimization of an objective function,

$$\hat{\mathbf{r}} = \underset{\mathbf{r}}{\operatorname{argmin}} \left\{ \left|\left| f(\mathbf{r}) - \mathbf{d} \right|\right| + \alpha \left|\left| \mathbf{r} - \mathbf{r}^* \right|\right| \right\}. \tag{2}$$

Here, $\alpha > 0$ is a small parameter and $\mathbf{r}^*$ is a reasonable guess for the solution. The regularization and associated minimization is most commonly carried out with respect to the L_2 norm, but other norms may be used. Nevertheless, the minimization problem is in general complicated and requires the use of a suitable numerical optimization procedure.

We have so far discussed inverse problems in which the observations are assumed to be exact. In real world applications, the assumption of exact observations tends to be unreasonable. Observations are burdened by measurement errors which, even for well-posed problems, may entail that the correct solution can not be found. Measurement errors can be detrimental for so-called ill-conditioned problems, which are characterized by that small errors in the observations yield much greater errors in the solution. This type of inverse problems can be approached by regularization techniques. However, numerical approaches yield solutions that are point estimates; hence, the inherent uncertainties of the observations are not reflected. A probabilistic approach that incorporates the measurement errors therefore seems more reasonable, and it can provide a solution on the form of a probability statement. In a probabilistic setting, the forward model takes the form

$$\big[\mathbf{d}|\mathbf{r}\big] = f(\mathbf{r}) + \mathbf{e}_{d|r}, \tag{3}$$

where $\mathbf{e}_{d|r}$ is a stochastic variable representing the measurement errors. The probabilistic forward model has an associated probability density function (pdf) $p(\mathbf{d}|\mathbf{r})$, which is termed the likelihood model.

Solving the inverse problem in a statistical framework involves identifying an estimator $\hat{\mathbf{r}}(\mathbf{d}) : \mathcal{D} \to \mathcal{R}$, from which an estimate of the unknown causal variable $\mathbf{r}$ can be computed. Estimators are often evaluated according to a loss function $L(\mathbf{r}, \hat{\mathbf{r}}(\mathbf{d}))$, which is typically taken to be the sum of squared residuals, i.e., $L(\mathbf{r}, \hat{\mathbf{r}}(\mathbf{d})) = \big|\big|\mathbf{r} - \hat{\mathbf{r}}(\mathbf{d})\big|\big|^2$ .The expected loss of the estimator is captured by the risk function $S_{\hat{r}}(\mathbf{r})$, which is defined pointwise in $\mathcal{R}$ as

$$S_{\hat{r}}(\mathbf{r}) = \int_{\mathcal{D}} L(\mathbf{r}, \hat{\mathbf{r}}(\mathbf{d}))\ \mathrm{d}p(\mathbf{d}|\mathbf{r}). \tag{4}$$

In the minimax approach, the optimal estimator is the estimator that minimizes the maximum risk. That is, the minimax risk is $S_{mm} = \inf_{\hat{\mathbf{r}}} \sup_{\mathbf{r}\in\mathcal{R}} S_{\hat{r}}(\mathbf{r})$, and the optimal minimax estimator is the one that achieves this risk. The minimax approach is a general principle for estimator selection, but it tends to be infeasible due to computational complexity, and results regarding the assessment of estimator uncertainty are limited. However, Stark (1992) reports results on minimax confidence intervals for a linear forward function (Backus, 1989; Donoho, 1989). Alternatively, estimator selection can be based on minimization of the average risk. This is the framework of Bayesian statistics, in which optimal estimators minimize the average risk. In practice, estimation and uncertainty quantification are easily available in a Bayesian framework,

albeit substantial computational effort may be needed in some cases. We therefore further pursue inverse problems in a Bayesian framework.

Bayesian inversion

We now consider $\mathbf{r}$ and $\mathbf{d}$ to be real valued random vectors of dimension n_r and n_d, respectively. In a Bayesian framework, the solution to the inverse problem is the posterior model $p(\mathbf{r}|\mathbf{d})$, which can be used for estimation and uncertainty assessment. The posterior model is the probability density function given by

$$p(\mathbf{r}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{r})p(\mathbf{r})}{p(\mathbf{d})} \propto p(\mathbf{d}|\mathbf{r})p(\mathbf{r}). \tag{5}$$

Here, $p(\mathbf{r})$ is the prior model, which represents the prior knowledge and beliefs that the modeler has about $\mathbf{r}$. Furthermore, $p(\mathbf{d}|\mathbf{r})$ is the likelihood model describing the relationship between $\mathbf{r}$ and $\mathbf{d}$. The normalizing constant $p(\mathbf{d})$ tends to be challenging to compute and can make it difficult to obtain posterior models.

For certain combinations of prior models and likelihood models, the form of the posterior models are known. A class of prior models $\mathcal{C}$ is said to be conjugate with respect to a likelihood model, if the corresponding class of posterior models is also $\mathcal{C}$ (Casella and Berger, 2001). Gaussian prior models provide the perhaps most well known examples of conjugacy; Gaussian prior models are conjugate with respect to so-called Gauss-linear likelihood models (Tarantola, 2005), which are likelihood models that are Gaussian with expectation linear in $\mathbf{r}$. This relation provides the foundation of geostatistics and Kriging interpolation (Chilès and Delfiner, 1999), and because Gaussian prior models can adequately represent many phenomena, the relation is often used in Bayesian inversion frameworks due to analytically available posterior models. Moreover, predictive quantities can be computed directly from the parameters of the posterior model.

The n_r-dimensional random vector $\mathbf{r}$ is a Gaussian random field (GRF) with n_r-dimensional expectation vector $\boldsymbol{\mu}_r$ and $(n_r \times n_r)$ covariance matrix $\boldsymbol{\Sigma}_r$, if its pdf is of the form (Johnson and Wichern, 2007)

$$p(\mathbf{r}) = (2\pi)^{-n_r/2} \left|\boldsymbol{\Sigma}_r\right|^{-1/2} \exp\left\{-\frac{1}{2}\left(\mathbf{r}-\boldsymbol{\mu}_r\right)^T \boldsymbol{\Sigma}_r^{-1}\left(\mathbf{r}-\boldsymbol{\mu}_r\right)\right\}, \tag{6}$$

with the superscript T indicating the matrix transpose. We denote this Gaussian pdf by $\varphi_{n_r}\left(\mathbf{r}; \boldsymbol{\mu}_r, \boldsymbol{\Sigma}_r\right)$. A Gauss-linear forward model is of the form $[\mathbf{d}|\mathbf{r}] = \mathbf{F}\mathbf{r} + \mathbf{e}_{d|r}$, where $\mathbf{F}$ is an $(n_d \times n_r)$ matrix representing the

linear operation on $\mathbf{r}$ and the n_d-dimensional vector $\mathbf{e}_{d|r}$ is an error term with Gaussian distribution. We assume the error term to be zero in expectation; hence, the likelihood model can be expressed

$$p(\mathbf{d}|\mathbf{r}) = \varphi_{n_d}(\mathbf{d}; \mathbf{Fr}, \boldsymbol{\Sigma}_{d|r}), \tag{7}$$

where $\boldsymbol{\Sigma}_{d|r}$ is the associated $(n_d \times n_d)$ covariance matrix. The Gaussian posterior model $\varphi_{n_r}\left(\mathbf{r}; \boldsymbol{\mu}_{r|d}, \boldsymbol{\Sigma}_{r|d}\right)$ is easily available through computation of its conditional parameters (Johnson and Wichern, 2007),

$$\begin{aligned} \boldsymbol{\mu}_{r|d} &= \boldsymbol{\mu}_r + \boldsymbol{\Gamma}_{rd}\boldsymbol{\Sigma}_d^{-1}\left(\mathbf{d} - \boldsymbol{\mu}_d\right) \\ \boldsymbol{\Sigma}_{r|d} &= \boldsymbol{\Sigma}_r - \boldsymbol{\Gamma}_{rd}^T\boldsymbol{\Sigma}_d^{-1}\boldsymbol{\Gamma}_{rd}, \end{aligned} \tag{8}$$

where $\boldsymbol{\Gamma}_{rd} = \boldsymbol{\Sigma}_r\mathbf{F}^T$ contains the inter-variable covariances between $\mathbf{r}$ and $\mathbf{d}$. Moreover, $\boldsymbol{\mu}_d = \mathbf{F}\boldsymbol{\mu}_r$ and $\boldsymbol{\Sigma}_d = \mathbf{F}\boldsymbol{\Sigma}_r\mathbf{F}^T$.

If a non-conjugate prior model is used, assessment of the posterior model tends to be simulation based. Markov chain Monte Carlo (McMC) is a widely used simulation technique for posterior model assessment (Mosegaard and Tarantola, 1995; Sen and Stoffa, 1996; Eidsvik et al., 2004a).

Markov chain Monte Carlo simulation

McMC simulation schemes make use of Markov chains to simulate from the target distribution $\pi(\mathbf{r}) = p(\mathbf{r}|\mathbf{d})$. The chain requires a transition kernel $p_t\left(\mathbf{r}'|\mathbf{r}\right)$, by which the chain can enter new states $\mathbf{r}'$ given its current state $\mathbf{r}$. Such a transition kernel must ensure that the target distribution $\pi(\cdot)$ is the equilibrium distribution of the chain, i.e.,

$$\int_B \pi(\mathbf{r})\ \mathrm{d}\mathbf{r} = \int p_t(B|\mathbf{r})\pi(\mathbf{r})\ \mathrm{d}\mathbf{r}, \quad \forall B \in \mathcal{B}, \tag{9}$$

where

$$p_t(B|\mathbf{r}) = \int_B p_t(\mathbf{r}'|\mathbf{r})\ \mathrm{d}\mathbf{r}' \tag{10}$$

and $\mathcal{B}$ is the Borel σ-field on $\mathbb{R}^{n_r}$. This condition is ensured by using reversible chains where the transition kernel satisifies the detailed balance equation (Gamerman and Hedibert, 2006),

$$\pi(\mathbf{r})p_t(\mathbf{r}'|\mathbf{r}) = \pi(\mathbf{r}')p_t(\mathbf{r}|\mathbf{r}'), \quad \forall(\mathbf{r}, \mathbf{r}'). \tag{11}$$

The transition kernel $p_t(\mathbf{r}'|\mathbf{r})$ consists of a proposal kernel $q(\mathbf{r}'|\mathbf{r})$ and a probability $\alpha(\mathbf{r}'|\mathbf{r}) \in [0, 1]$, such that $p_t(\mathbf{r}'|\mathbf{r}) = \alpha(\mathbf{r}'|\mathbf{r})q(\mathbf{r}'|\mathbf{r})$ for $\mathbf{r}' \neq \mathbf{r}$

defines the probability to move to a new state. Hence, the complement of the integral of the transition kernel over all new states defines the probability to remain in the current state, and the transition kernel can be expressed

$$p_t(B|\mathbf{r}) = \int_B q(\mathbf{r}'|\mathbf{r})\alpha(\mathbf{r}'|\mathbf{r})\ \mathrm{d}\mathbf{r}' + I(\mathbf{r} \in B)\left[1 - \int q(\mathbf{r}'|\mathbf{r})\alpha(\mathbf{r}'|\mathbf{r})\ \mathrm{d}\mathbf{r}'\right], \quad \forall B \in \mathcal{B}, \tag{12}$$

where $I(\cdot)$ is the indicator function, being equal to 1 if its argument is true and 0 otherwise.

Identifying a transition kernel $p_t(\mathbf{r}'|\mathbf{r})$ that can produce a Markov chain with the target distribution $\pi(\cdot)$ as its equilibrium distribution may appear as a daunting task. Conveniently, the Metropolis-Hastings (M-H) algorithm (Hastings, 1970) provides the means by which to do so. The M-H algorithm designates an acceptance probability that ensures that the transition kernel $p_t(\cdot|\cdot)$ defines a reversible chain when combined with an arbitrary proposal kernel $q(\cdot|\cdot)$,

$$\alpha(\mathbf{r}'|\mathbf{r}) = \min\left\{1, \frac{\pi(\mathbf{r}')q(\mathbf{r}|\mathbf{r}')}{\pi(\mathbf{r})q(\mathbf{r}'|\mathbf{r})}\right\}. \tag{13}$$

The M-H algorithm defines an irreducible and aperiodic chain with transition kernel $p_t(\cdot|\cdot)$ with $\pi(\cdot)$ as its limiting distribution if the proposal kernel $q(\cdot|\cdot)$ is aperiodic and irreducible and $\alpha(\mathbf{r}'|\mathbf{r}) > 0$ for all possible values of $(\mathbf{r}', \mathbf{r})$ (Roberts and Smith, 1994).

The sequence of simulations $\{\mathbf{r}_k\}_{k=1}^n$ from the McMC algorithm converges in distribution to the target distribution $\pi(\mathbf{r})$ as $n \to \infty$. Hence, obtaining a result in finite time necessitates decisions about where to start and when to stop the algorithm. Usually, the initial state of a chain is drawn at random and will not be from a representative region of $\pi(\mathbf{r})$. A subsequent sequence of simulations are notably influenced by the initial state, before the chain is sufficiently close to $\pi(\mathbf{r})$. This notably affected sequence of simulations is termed the burn-in of the chain, and we discard it because these simulations are not representative of $\pi(\mathbf{r})$. For practical purposes, we draw inspiration from the Gelman-Rubin convergence diagnostic (Gelman and Rubin, 1992) to decide when to stop a chain, unless otherwise stated. We run several chains in parallel and compare the results to determine an appropriate stopping point. The time necessary to reach a stopping point depends on how efficiently the chain navigates the target distribution $\pi(\mathbf{r})$, which is captured by the concept of mixing. A chain is said to have good mixing if approximately

independent simulations are not far apart in the sequence of simulations, and conversely a chain has bad mixing if approximately independent simulations are far apart. Good mixing is preferable because it entails that the chain rapidly arrives at an acceptable degree of convergence. The mixing of a chain is influenced by the proposal kernel $q(\cdot|\cdot)$; hence, the proposal kernel should be selected with care if computational efficiency is a concern.

In a setting where the form of the posterior model is unknown, McMC simulation is traditionally performed by using the prior model $p(\mathbf{r})$ as proposal distribution (Mosegaard and Tarantola, 1995). The associated M-H acceptance probability is given by

$$\alpha(\mathbf{r}'|\mathbf{r}) = \min\left\{1, \frac{p(\mathbf{d}|\mathbf{r}')}{p(\mathbf{d}|\mathbf{r})}\right\}. \tag{14}$$

Although this approach is generally valid, it tends to suffer from low acceptance rates in multivariate settings. Moreover, in the absence of a very informative prior model there is an inverse relationship between acceptance rates and data quality. The acceptance rate increases as the uncertainty associated with the measured data increases, and decreases as the number of data points increases.

If a prior model that is conjugate with respect to a linear likelihood model is used with a non-linear likelihood model, the conjugate property can be exploited for more efficient simulation by use of an approximated linear likelihood model, $p_L(\mathbf{d}|\mathbf{r}) \approx p(\mathbf{d}|\mathbf{r})$. Proposals can then be made from an approximate posterior model $p_L(\mathbf{r}|\mathbf{d}) \propto p_L(\mathbf{d}|\mathbf{r})p(\mathbf{r})$. Simulating from an approximate posterior model is similar to simulating from the prior model due to the conjugate property, but the proposals are likely to be better; hence, this approach may increase the acceptance rate. The associated M-H acceptance probability is

$$\alpha(\mathbf{r}'|\mathbf{r}) = \min\left\{1, \frac{p(\mathbf{d}|\mathbf{r}')/p_L(\mathbf{d}|\mathbf{r}')}{p(\mathbf{d}|\mathbf{r})/p_L(\mathbf{d}|\mathbf{r})}\right\}. \tag{15}$$

Note that $\alpha(\mathbf{r}'|\mathbf{r}) \to 1$ if the likelihood model is close to linear $p(\mathbf{d}|\mathbf{r}) \to p_L(\mathbf{d}|\mathbf{r})$, as expected.

Reservoir data

Oil and gas reservoirs are formations of rock in the subsurface in which hydrocarbons have accumulated. Reservoirs are expensive to produce and necessitate the construction of production wells. Hence, attempts to produce reservoirs that are poorly suited for production entails big economical losses. However, producing well suited reservoirs can yield very large payoffs. The ability to identify a reservoirs suitability for production is therefore of utmost importance in the oil and gas industry. The process of doing so is called reservoir characterization, which we will return to later. Reservoir characterization relies on data from potential reservoirs, which usually consist of seismic data and well data. Seismic data has good spatial coverage and poor precision, whereas well data has poor spatial coverage and high precision.

Seismic AVO data

Seismic data are collected by emitting compressional waves into the subsurface and measuring and registering the amplitude of the waves that are reflected back. These data and can be collected from large subsurface volumes in search for potential hydrocarbon reservoirs at relatively low cost. The arrival times of the reflections relative to the time of emission are also registered. The arrival times make it possible to map the registered amplitudes to particular locations in the subsurface; hence, seismic data are a representation of the subsurface based on wave amplitudes and times.

The physics involved in the collection of seismic data is complex and relies on intimate details of the medium that the seismic waves propagate through. Seismic wave energy is reflected back at interfaces defined by rock or fluid inhomogeneities. Moreover, the proportion of reflected energy is dependent on the angle of incidence of the emitted seismic waves on the interfaces, which is described by the Zoeppritz equations (Zoeppritz, 1919). These equations relate the reflectivity coefficients at interfaces to changes in the elastic properties and angle of incidence. The elastic properties consist of P-wave velocity and S-wave velocity, which are jointly referred to as seismic velocities, and density.

The angle dependency of the reflected energy is very useful for detecting hydrocarbons, because the seismic responses of fluid transitions have a strong and characteristic angle dependency. This attribute forms the basis for seismic amplitude variation with offset (AVO) data and seismic AVO analysis. Seismic AVO data consist of seismic data associated with different angles of incidence of the emitted compressional waves on in-

terfaces in the subsurface. For a specific subsurface target, seismic AVO data are collected by varying the horizontal distance between the source that emits the compresssional waves and the receiver that measures the reflections, while keeping the subsurface target centered between them. In practice, receivers are usually placed at numerous distances from the source to obtain seismic data from different subsurface targets simultaneously. This principle is used in offshore seismic data collection, in which a large number of receivers are towed behind a moving ship that emits compressional waves into the subsurface.

As an emitted compressional wave travels through the subsurface, it becomes distorted and stretched due to dispersion, and the shape of the distorted pulse is termed a seismic wavelet. The measured reflection at a given point in time contains contributions from several reflectivity coefficients in the subsurface, weighted by the seismic wavelet. Hence, the received signal is modeled as a convolution of a sequence of reflectivity coefficients with a seismic wavelet in a time interval. In practice, the seismic wavelet is unknown and application dependent, because the dispersive process is sensitive to the medium that the seismic waves travel through. Therefore, in order to construct a sensible seismic forward model, seismic data should be supplemented by well data from which the seismic wavelet can be estimated.

Well data

In oil and gas exploration, it is usual to drill bore-holes at locations of particular interest in a seismic survey to obtain information that enables interpretation of the seismic data. Core samples, which are informative about the subsurface lithology, are collected during drilling. Once the wells are drilled, measuring instruments are placed in the bore-holes and hoisted up while producing well log data on a regular and relatively fine grid along the well profiles. Measurements of the seismic velocities and density, as well as measurements of petrophysical properties are recorded. The measuring instruments are highly accurate; hence, well log data are subject only to minor measurement errors and can be considered as good approximations to the truth.

The reflectivity coefficients that can be used to estimate the seismic wavelet along the wells can be computed from well data using the Zoeppritz equations. Estimating the seismic wavelet at a few locations in a seismic survey tends to be adequate because the wavelet shape varies slowly with lateral position (Walden and White, 1998), due to layering effects. Another important use of well data is in the construction of rock physics models, which are necessary to interpret seismic data in terms

of petrophysical properties. The seismic data are related to the elastic properties, and a rock physics model can in turn relate the elastic properties to petrophysical properties. A rock physics model can either be based on theoretical relations or be empirically approximated.

Bayesian reservoir characterization

The goal of reservoir characterization is to evaluate a reservoir's suitability for production. Reservoir characterization is a spatial problem in which the reservoir zone of interest is discretized into a reservoir grid $\mathcal{G}_r$, consisting of n_r grid points. A spatial reservoir variable $\mathbf{r}$ from which the production suitability of a reservoir can be inferred, is defined on the reservoir grid $\mathcal{G}_r$. That is, each random variable contained in $\mathbf{r}$ is associated with a location in $\mathcal{G}_r$, which enables spatial effects to influence the characterization. A reservoir variable consists of a few selected petrophysical properties related to the producibility of the reservoir under study and typically includes permeability/porosity and water saturation, which are volumetric fractions, i.e., quantities limited to $[0, 1]$. Porosity is informative about the presence of pores in the rock, which are pockets of empty space where hydrocarbons can settle. The degree to which the empty spaces in the rock are connected is described by permeability. The connectivity of the pores is important for fluid flow through the rock, which is crucial for the extraction of hydrocarbons. Permeability is usually strongly dependent on porosity, and since the former is complicated to measure, it is often derived from the latter. Water saturation is indicative of the location of hydrocarbons. Because water is denser than hydrocarbons, gravitational effects tend to separate the fluids. Hence, hydrocarbons can typically be found at locations where the water saturation is low. In reservoirs that are lithologically heterogeneous, porosity and water saturation may not be sufficiently informative and should preferably be accompanied by lithology variables, such as volume of clay, to explain certain phenomena. The reservoir variable $\mathbf{r}$ contains n_p selected petrophysical properties and can be expressed as $\mathbf{r} = [\mathbf{r}_1, ..., \mathbf{r}_{n_p}]$, with each $\mathbf{r}_k$ being defined on the reservoir grid $\mathcal{G}_r$, for $k = 1, ..., n_p$. Hence, $\mathbf{r}$ is an $n_p n_r$-dimensional vector. For ease of presentation we will consider $n_p = 1$ in the following.

Reservoir characterization of a subsurface reservoir volume is typically done by inversion of seismic data. This is an inverse problem in which seismic data $\mathbf{d}$ on a seismic grid $\mathcal{G}_d$, consisting of n_d grid points, is sought to be explained by the reservoir variables $\mathbf{r}$. We confine ourselves to seismic AVO data. The inversion is either cast into an optimization

setting or into a probabilistic setting. In an optimization setting, a highly accurate seismic forward model is used and the deviation between $\mathbf{d}$ and the seismics computed by the seismic forward model is minimized with respect to the elastic properties (Sen and Stoffa, 2013). Probabilistically, the inverse problem is usually approached in a Bayesian framework using approximate Zoeppritz equations (Buland and Omre, 2003; Gunning and Glinsky, 2004; Larsen et al., 2006). The Bayesian seismic reservoir characterization is represented by the posterior model $p(\mathbf{r}|\mathbf{d})$, which is proportional to the product of a likelihood model $p(\mathbf{d}|\mathbf{r})$ and a prior model $p(\mathbf{r})$, both of which has to be specified, see Equation 5.

Likelihood model

The likelihood model represents the chain of data acquisition, from the target variables of the inversion to the data, and is based on geophysics theory and well data. Well data from representative regions of the reservoir volume are particularly important for seismic wavelet estimation and may also be used to inform the rock physics model.

The Zoeppritz equations can in principle completely describe the relation between the PP reflectivity coefficients $c(t, \theta)$ and the elastic properties along a seismic trace. The elastic properties are canonical variables of the equation and consist of the seismic velocities, which are the P-wave velocities $V_p(t)$ and the S-wave velocities $V_s(t)$, and the densities $\rho(t)$. However, the equations are difficult to interpret and their solution is unstable due to nonlinearity. Therefore, the Zoeppritz equations tend to be linearly approximated, and several linear approximations exist, including Bortfeld's approximation (Bortfeld, 1961), Aki and Richards' approximation (Aki and Richards, 1980), and Shuey's approximation (Shuey, 1985). In Bayesian inversion frameworks, the Aki and Richards' approximation is a common choice (Buland and Omre, 2003; Larsen et al., 2006; Grana and Della Rossa, 2010; Rimstad et al., 2012). We use a time continuous reflectivity function for the PP reflection coefficients, based on the Aki and Richards' approximation (Buland and Omre, 2003),

$$\begin{aligned} c\left(t,\theta\right) = a_1\left(t,\theta\right)\frac{\delta}{\delta t}\ln\left(V_p\left(t\right)\right) + a_2\left(t,\theta\right)\frac{\delta}{\delta t}\ln\left(V_s\left(t\right)\right) \\ + a_3\left(t,\theta\right)\frac{\delta}{\delta t}\ln\left(\rho\left(t\right)\right), \end{aligned} \tag{16}$$

where

$$a_1(t,\theta) = \frac{1}{2}\left(1 + \tan^2(\theta)\right), \tag{17}$$

$$a_2(t,\theta) = -4\frac{\bar{V}_s^2(t)}{\bar{V}_p^2(t)}\sin^2(\theta), \tag{18}$$

$$a_3(t,\theta) = \frac{1}{2}\left(1 - 4\frac{\bar{V}_s^2(t)}{\bar{V}_p^2(t)}\sin^2(\theta)\right). \tag{19}$$

Moreover, $\bar{V}_p(t)$, $\bar{V}_s(t)$, and $\bar{\rho}(t)$ are time dependent averages which are assumed to be adequately represented by a constant or moving average in a time window. The above approximation can for n_θ offset angles be discretized and represented in matrix form as $\mathbf{c} = \mathbf{ADm}$. Here, we discretize according to $\mathcal{G}_r$. Hence, $\mathbf{c}$ is an $n_\theta n_r$-dimensional vector. Moreover, $\mathbf{A}$ is the sparse $(n_\theta n_r \times 3n_r)$ matrix

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1(\theta_1) & \mathbf{A}_2(\theta_1) & \mathbf{A}_3(\theta_1) \\ \vdots & \vdots & \vdots \\ \mathbf{A}_1(\theta_{n_\theta}) & \mathbf{A}_2(\theta_{n_\theta}) & \mathbf{A}_3(\theta_{n_\theta}), \end{bmatrix} \tag{20}$$

where $\mathbf{A}_1(\theta_i)$, $\mathbf{A}_2(\theta_i)$, and $\mathbf{A}_3(\theta_i)$ are $(n_r \times n_r)$ diagonal matrices containing discrete time samples of $a_1(t,\theta_i)$, $a_2(t,\theta_i)$, and $a_3(t,\theta_i)$, respectively, for $i = 1, ..., n_\theta$. Furthermore, the $(3n_r \times 3n_r)$ matrix $\mathbf{D}$ is a first order differential operator. Lastly, the $3n_r$-dimensional vector $\mathbf{m}$ contains the elastic properties, discretized on $\mathcal{G}_r$, in the logarithmic domain. Hence, a reflectivity likelihood model is

$$\big[\mathbf{c}|\mathbf{m}\big] = \mathbf{ADm} + \mathbf{e}_{c|m}, \tag{21}$$

where $\mathbf{e}_{c|m}$ is an $n_\theta n_r$-dimensional vector containing approximation errors.

The seismic AVO data $\mathbf{d}$ are the convolution of a seismic wavelet with the reflectivity coefficients $\mathbf{c}$. The wavelet is discretized to be consistent with the resolution of the reflectivity coefficients, and the convolutional likelihood model is

$$\big[\mathbf{d}|\mathbf{c}\big] = \mathbf{Wc} + \mathbf{e}_{d|c}. \tag{22}$$

Here, the $(n_\theta n_d \times n_\theta n_r)$ matrix $\mathbf{W}$ contains discretizations of the seismic wavelet $W(t,\theta)$ for all n_θ offset angles, and the $n_\theta n_d$-dimensional vector

$\mathbf{e}_{d|c}$ contains observation errors. The wavelet matrix is of the form

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_1 & 0\mathbf{i}_{n_r} & \dots & \dots & 0\mathbf{i}_{n_r} \\ 0\mathbf{i}_{n_r} & \mathbf{W}_2 & 0\mathbf{i}_{n_r} & \dots & \vdots \\ \vdots & 0\mathbf{i}_{n_r} & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & 0\mathbf{i}_{n_r} \\ 0\mathbf{i}_{n_r} & \dots & \dots & 0\mathbf{i}_{n_r} & \mathbf{W}_{n_\theta} \end{bmatrix}, \tag{23}$$

where the $(n_d \times n_r)$ matrices $\mathbf{W}_i$ contain n_d discretizations of the seismic wavelet, with each row corresponding to the seismic wavelet centered at a location in $\mathcal{G}_d$.

The seismic likelihood model is

$$p(\mathbf{d}|\mathbf{m}) = \int p(\mathbf{d}|\mathbf{c}, \mathbf{m}) p(\mathbf{c}|\mathbf{m}) \, \mathrm{d}\mathbf{c} = \int p(\mathbf{d}|\mathbf{c}) p(\mathbf{c}|\mathbf{m}) \, \mathrm{d}\mathbf{c}, \tag{24}$$

which can be expressed

$$[\mathbf{d}|\mathbf{m}] = \mathbf{WADm} + \mathbf{e}_{d|m}, \tag{25}$$

where $\mathbf{e}_{d|m} = \mathbf{W}\mathbf{e}_{c|m} + \mathbf{e}_{d|c}$. The error terms are typically assigned Gaussian distributions (Buland and Omre, 2003), which yields a Gauss-linear seismic likelihood model. This likelihood model can readily be used in a Bayesian seismic inversion framework for the elastic properties, given a prior model $p(\mathbf{m})$. A Gaussian prior model is advantageous for its conjugate property, but not required. However, if reservoir variables are the target of the inversion, a rock physics likelihood model $p(\mathbf{m}|\mathbf{r})$ is needed. The rock physics model can be integrated into the seismic inversion model or be used in rock physics inversion, i.e., after seismic inversion to elastic properties. However, for the uncertainties to propogate all the way from the seismic data to the reservoir variables, the data acquisition procedure should be described all the way from $\mathbf{r}$ to $\mathbf{d}$ by a likelihood model. The overall likelihood model is

$$p(\mathbf{d}|\mathbf{r}) = \int p(\mathbf{d}|\mathbf{m}, \mathbf{r}) p(\mathbf{m}|\mathbf{r}) \, \mathrm{d}\mathbf{m} = \int p(\mathbf{d}|\mathbf{m}) p(\mathbf{m}|\mathbf{r}) \, \mathrm{d}\mathbf{m}, \tag{26}$$

where the latter step follows because $\mathbf{m}$ are canonical variables of the Zoeppritz equations. The rock physics forward function may be non-linear and, if so, the forward function in the overall likelihood model $p(\mathbf{d}|\mathbf{r})$ is non-linear. Under these circumstances, simulating from the posterior model tends to be computationally inefficient in spatial settings. However, if a linear forward function seems feasible, either by empirical estimation

or theoretical approximation, computationally efficient inversion schemes can be based on a Gauss-linear rock physics model. The rock physics forward function is then represented by the $(3n_r \times n_r)$ matrix $\mathbf{B}$ and the overall likelihood model takes the form

$$[\mathbf{d}|\mathbf{r}] = \mathbf{G}\mathbf{r} + \mathbf{e}_{d|r} = \mathbf{WADB}\mathbf{r} + \mathbf{e}_{d|r}, \tag{27}$$

where $\mathbf{e}_{d|r} = \mathbf{WADe}_{m|r} + \mathbf{e}_{d|m}$ is assumed to be Gaussian, with $\mathbf{e}_{m|r}$ a $3n_r$-dimensional vector containing rock physics model errors. Hence, the overall likelihood model is defined to be Gauss-linear.

Prior & Posterior model

The prior model $p(\mathbf{r})$, see Equation 5, is assigned on a subjective basis, but should be representative of $\mathbf{r}$. A prior model is representative if it is accurately centered and realistically represents the uncertainty and spatial continuity in $\mathbf{r}$. Prior model assignment can be based on experience, expert knowledge, beliefs, data, or any combination thereof. If representative well data of the reservoir variables are available, it is natural to adopt an empirical Bayes approach to prior model assignment, and we do so. The empirical Bayes method entails estimation of the prior model from representative data, which can be done either non-parametrically or parametrically. We will be concerned with parametric empirical Bayes, to which an introduction can be found in Casella (1985). Reservoir variables tend to be multimodal due to underlying lithology/fluid (LF) classes (Grana and Della Rossa, 2010; Rimstad et al., 2012); hence, a Gaussian prior model may not be adequate. A prior model $p(\mathbf{r})$ should often have support for multimodality, and two Gaussian-based parametric model alternatives have emerged in the Bayesian seismic inversion literature; namely, Gaussian mixture models and selection Gaussian models.

Gaussian mixture models

Gaussian mixture models (GMMs) can represent multimodal distributions and have successfully been applied to reservoir characterization (Grana and Della Rossa, 2010; Rimstad et al., 2012; Fjeldstad et al., 2021). As the name implies, GMMs are generated by combining Gaussian models. Typically, the mixture is based on pre-defined LF classes that notably affect the reservoir variables; hence, the model relies on conditional Gaussian distributions for the reservoir variables, $p(\mathbf{r}|\boldsymbol{\kappa}) = \varphi_{n_r}(\mathbf{r}; \boldsymbol{\mu}_{r|\kappa}, \boldsymbol{\Sigma}_{r|\kappa})$, where $\boldsymbol{\kappa} \in \Omega_\kappa^{n_r}$ is a spatial mode indicator variable representing the LF classes. Here, $\Omega_\kappa = \{\omega_1, ..., \omega_{n_\kappa}\}$ is the set of LF classes. A Gaussian

mixture prior model $p(\mathbf{r})$ can be expressed as

$$p(\mathbf{r}) = \sum_{\Omega_\kappa^{n_r}} p(\mathbf{r}|\boldsymbol{\kappa})p(\boldsymbol{\kappa}), \tag{28}$$

where $p(\boldsymbol{\kappa})$ is a prior model for $\boldsymbol{\kappa}$. Defined as such, the prior model for the reservoir variables is a probability weighted sum of Gaussian distributions, each of which has a parameterization that is characteristic for a particular LF class.

The prior model for $\boldsymbol{\kappa}$ should honor geophysical laws and tendencies with respect to LF ordering and transitions, and spatial continuity. Markov models can impose such constraints and have a long standing tradition in geophysical applications, first appearing in the form of Markov chains in 1D applications (Krumbein and Dacey, 1969). A first-order Markov chain prior model for $\boldsymbol{\kappa}$ can be expressed

$$p(\boldsymbol{\kappa}) = p(\kappa_1) \prod_{i=2}^{n_r} p(\kappa_i|\kappa_{i-1}), \tag{29}$$

where $p(\kappa_1)$ is the stationary probability and $p(\kappa_i|\kappa_{i-1}), i = 2, ..., n_r$ are transition probabilities. Higher order chains can be defined to add additional constraints such as minimum layer thickness. The interpretability and functionality of Markov models have made them widely used in geophysical applications, and in particular to exploring the LF properties of reservoirs (Eidsvik et al., 2004b; Larsen et al., 2006; Ulvmoen et al., 2010).

Generally, Markov models are specified in the form of Markov random fields (MRFs), which are typically defined according to the Gibbs formulation (Besag, 1974; Kindermann and Snell, 1980). This formulation generalizes the Markov chain definition to higher spatial dimensions and is based on so-called cliques. A clique is defined on an undirectional graph as a set of mutually adjacent vertices, i.e., every pair of vertices in a clique are adjacent. Moreover, a maximal clique c is a clique that is not a subset of a larger clique, and the set of maximal cliques on the graph is denoted by $\mathcal{C}$. The MRF model can be expressed in terms of maximal cliques as

$$p(\boldsymbol{\kappa}) \propto \exp\left\{ -\sum_{c \in \mathcal{C}} \psi_c(\boldsymbol{\kappa}_c) \right\}, \tag{30}$$

where $\psi_c(\boldsymbol{\kappa}_c)$ is the clique potential function for the maximal clique c and $\boldsymbol{\kappa}_c$ are the LF classes in the maximal clique c.

The neighbordhood structure of an MRF is related to its cliques, as described by the Hammersley-Clifford theorem (Hammersley and Clifford, 1971). Intuitively, the neighborhood structure of an MRF can be defined locationwise from its defining cliques as the set of locations in the union of all cliques that include the target location, except the target location itself, as illustrated in Figure 1. The MRF model supports arbitrary spatial dimensionality and higher order neighborhoods; hence, the Markov chain model specified in Equation 29 is a special case of an MRF in 1D with neighborhoods consisting of the two nearest locations, except for border effects.

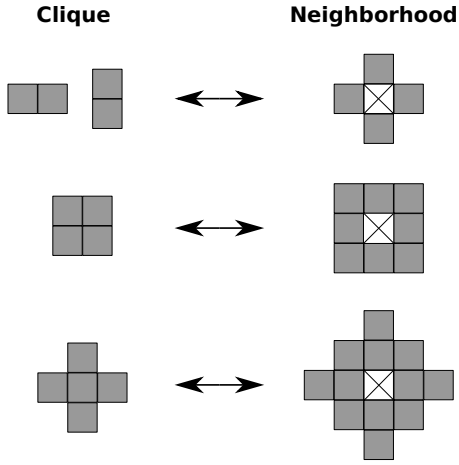

Figure 1: Cliques with corresponding neighborhoods. The neighborhoods are defined with respect to the cells marked by a cross.

The full conditional densities are $p(\kappa_i|\boldsymbol{\kappa}_{i^c})$, with the subscript i^c denoting the complement of location i on the graph, i.e., $i^c = \{1, ..., n_r\}\backslash i$, $i = 1, ..., n_r$. These densities are defined by the neighborhood structure of the MRF, which can be identified through application of the Hammersley-Clifford theorem. The typically large grids associated with reservoir characterization entails that the normalization constant of the MRF is computationally prohibitive. Therefore, assessment of $p(\boldsymbol{\kappa})$ tends to be simulation based and is often done by using the full conditional densities $p(\kappa_i|\boldsymbol{\kappa}_{i^c})$, that is, simulation is usually done by single-site Gibbs sampling.

GMMs are conjugate prior models subject to Gauss-linear likelihood models (Grana et al., 2017; Fjeldstad and Grana, 2018); hence, $p(\mathbf{r}|\mathbf{d})$ is

also a GMM and of the form

$$p(\mathbf{r}|\mathbf{d}) = \sum_{\Omega_\kappa^{n_r}} p(\mathbf{r}|\boldsymbol{\kappa}, \mathbf{d}) p(\boldsymbol{\kappa}|\mathbf{d}). \tag{31}$$

Note that the conditional pdf $p(\mathbf{r}|\boldsymbol{\kappa}, \mathbf{d})$ is Gaussian and therefore easy to evaluate. Evaluation of the mixing weights $p(\boldsymbol{\kappa}|\mathbf{d})$ is more involved. Assessment of the posterior model $p(\boldsymbol{\kappa}|\mathbf{d})$ is complicated due to the vertically convolutional seismic likelihood model, which results in $p(\boldsymbol{\kappa}|\mathbf{d})$ being a convolved hidden Markov model (Lindberg and Omre, 2014). Consequently, $p(\boldsymbol{\kappa}|\mathbf{d})$ will be a higher-order Markov model, irrespective of the order defined in $p(\boldsymbol{\kappa})$. This entails that assessment of the exact posterior Markov model becomes computationally infeasible. However, assessment can be based on approximations (Rimstad and Omre, 2013; Fjeldstad and Grana, 2018).

Selection Gaussian models

The selection Gaussian model is inspired by the concept of selection probability distributions (Azzalini, 1985; Arellano-Valle et al., 2006; Azzalini, 2013). The selection concept has been extended to spatial settings (Allard and Naveau, 2007; Omre and Rimstad, 2021) and has successfully been applied to seismic inversion (Karimi et al., 2010; Rimstad and Omre, 2014). The selection Gaussian model is a very flexible class of models and is a viable candidate for representing multimodal random variables. A selection Gaussian random field (S-GRF) is based on two interacting GRFs: the basis GRF $\tilde{\mathbf{r}}$ and the auxiliary GRF $\boldsymbol{\nu}$. The basis GRF is specified on the reservoir grid $\mathcal{G}_r$, whereas the auxiliary GRF is specified on a grid $\mathcal{G}_\nu$, which may differ. The conditional random field (RF) $[\boldsymbol{\nu}|\tilde{\mathbf{r}}]$ is Gauss-linear; hence, $(\tilde{\mathbf{r}}, \boldsymbol{\nu})$ are jointly Gaussian, with joint pdf

$$p\left(\begin{bmatrix}\tilde{\mathbf{r}}\\ \boldsymbol{\nu}\end{bmatrix}\right) = \varphi_{n_r+n_\nu}\left(\begin{bmatrix}\tilde{\mathbf{r}}\\ \boldsymbol{\nu}\end{bmatrix}; \begin{bmatrix}\boldsymbol{\mu}_{\tilde{r}}\\ \boldsymbol{\mu}_\nu\end{bmatrix}, \begin{bmatrix}\boldsymbol{\Sigma}_{\tilde{r}} & \boldsymbol{\Gamma}_{\tilde{r}\nu}\\ \boldsymbol{\Gamma}_{\tilde{r}\nu}^T & \boldsymbol{\Sigma}_\nu\end{bmatrix}\right). \tag{32}$$

Here, $\boldsymbol{\mu}_{\tilde{r}}$ is the n_r-dimensional expectation vector of $\tilde{\mathbf{r}}$ and $\boldsymbol{\Sigma}_{\tilde{r}}$ is its $(n_r \times n_r)$ covariance matrix. Similarly, $\boldsymbol{\mu}_\nu$ is the n_ν-dimensional expectation vector of $\boldsymbol{\nu}$ and $\boldsymbol{\Sigma}_\nu$ is its $(n_\nu \times n_\nu)$ covariance matrix. Lastly, $\boldsymbol{\Gamma}_{\tilde{r}\nu}$ is the $(n_r \times n_\nu)$ cross-covariance matrix between $\tilde{\mathbf{r}}$ and $\boldsymbol{\nu}$. A truncation of the auxiliary GRF $\boldsymbol{\nu}$ is defined according to the n_ν-dimensional selection set $\mathbf{A}$. The conditional RF $\mathbf{r} = [\tilde{\mathbf{r}}|\boldsymbol{\nu} \in \mathbf{A}]$ is selection Gaussian, with pdf

$$p(\mathbf{r}) = p(\tilde{\mathbf{r}}|\boldsymbol{\nu} \in \mathbf{A}) = \frac{\Phi_{n_\nu}(\mathbf{A}; \boldsymbol{\mu}_{\nu|\tilde{r}}, \boldsymbol{\Sigma}_{\nu|\tilde{r}})}{\Phi_{n_\nu}(\mathbf{A}; \boldsymbol{\mu}_\nu, \boldsymbol{\Sigma}_\nu)} \varphi_{n_r}(\tilde{\mathbf{r}}; \boldsymbol{\mu}_{\tilde{r}}, \boldsymbol{\Sigma}_{\tilde{r}}), \tag{33}$$

where the nominator and denominator are the Gaussian probabilities of the selection set. That is, the denominator is

$$\Phi_{n_\nu}(\mathbf{A}; \boldsymbol{\mu}_\nu, \boldsymbol{\Sigma}_\nu) = \int_{\mathbf{A}} \varphi_{n_\nu}(\boldsymbol{\nu}; \boldsymbol{\mu}_\nu, \boldsymbol{\Sigma}_\nu)\ \mathrm{d}\boldsymbol{\nu}, \tag{34}$$

and similarly for the nominator with appropriate distributional parameters. The conditional parameters in Equation 33 are computed by standard Gaussian conditioning formulas (Johnson and Wichern, 2007),

$$\begin{aligned} \boldsymbol{\mu}_{\nu|\tilde{r}} &= \boldsymbol{\mu}_\nu + \boldsymbol{\Gamma}_{\tilde{r}\nu}^T \boldsymbol{\Sigma}_{\tilde{r}}^{-1} (\tilde{\mathbf{r}} - \boldsymbol{\mu}_{\tilde{r}}) \\ \boldsymbol{\Sigma}_{\nu|\tilde{r}} &= \boldsymbol{\Sigma}_\nu - \boldsymbol{\Gamma}_{\tilde{r}\nu}^T \boldsymbol{\Sigma}_{\tilde{r}}^{-1} \boldsymbol{\Gamma}_{\tilde{r}\nu}. \end{aligned} \tag{35}$$

The flexibility of the distribution enters through the cross-covariance $\boldsymbol{\Gamma}_{\tilde{r}\nu}$ between $\tilde{\mathbf{r}}$ and $\boldsymbol{\nu}$, and the shape of the selection set $\mathbf{A}$.

To robustly represent the variability in the reservoir variables, a spatially stationary prior model is sensible, because it captures the total uncertainty reflected in the well data at every location. An RF is said to be stationary if its associated pdf is spatially shift invariant for any subset of its random variables. That is, the pdf of the chosen subset of random variables must depend only on the distances between them and not on their specific location. We specify an S-GRF model with $n_\nu = n_r$, which can support stationarity, except for border effects. Moreover, it enables mode transitions at every location in the reservoir grid, which can be important for precisely locating regions of interest in the spatial domain. To restrict model complexity, we specify an intervariable spatial correlation structure, which is contained in the $(n_r \times n_r)$ matrix $\boldsymbol{\Omega}$ and defined through a translation invariant and positive definite correlation function $\rho(\cdot)$. We define a stationary Gaussian distribution for $\tilde{\mathbf{r}}$ with expectation vector $\mu_{\tilde{r}}\mathbf{i}_{n_r}$ and covariance matrix $\boldsymbol{\Sigma}_{\tilde{r}} = \sigma_{\tilde{r}}^2\boldsymbol{\Omega}$. Here, $\mu_{\tilde{r}}$ and $\sigma_{\tilde{r}}^2$ are the locationwise expectation and variance of $\tilde{\mathbf{r}}$, respectively, and $\mathbf{i}_{n_r}$ is the n_r-dimensional vector of ones. Furthermore, because the influence of $[\boldsymbol{\nu}|\boldsymbol{\nu} \in \mathbf{A}]$ on $\tilde{\mathbf{r}}$ depends on the location of $\mathbf{A}$ only through the relative location of $\mathbf{A}$ to the distribution of the auxiliary variable, we define a stationary and locationwise standard Gaussian distribution for $\boldsymbol{\nu}$, with expectation vector $\boldsymbol{\mu}_\nu = 0\mathbf{i}_{n_r}$ and covariance matrix $\boldsymbol{\Sigma}_\nu = \gamma^2\boldsymbol{\Omega} + (1-\gamma^2)\mathbf{I}_{n_r}$, where $\mathbf{I}_{n_r}$ is the $(n_r \times n_r)$ identity matrix. Here, γ is the locationwise correlation between $\tilde{r}$ and ν. In this framework, the cross covariance matrix is an $(n_r \times n_r)$ matrix of the form $\boldsymbol{\Gamma}_{\tilde{r}\nu} = \gamma\sigma_{\tilde{r}}\boldsymbol{\Omega}$. Lastly, we use a location invariant selection set, i.e., $\mathbf{A} = A^{n_r}$. The locationwise selection set A consists of n_A disjoint real intervals, $A = \bigcup_{i=1}^{n_A}[a_i, b_i]$, $a_i < b_i$. The S-GRF specified above is stationary and the conditional parameters

involved in its pdf in Equation 33 can be expressed

$$\begin{aligned} \boldsymbol{\mu}_{\nu|\tilde{r}} &= 0\mathbf{i}_{n_r} + \gamma\sigma_{\tilde{r}}^{-1}\left(\tilde{\mathbf{r}} - \mu_{\tilde{r}}\mathbf{i}_{n_r}\right) \\ \boldsymbol{\Sigma}_{\nu|\tilde{r}} &= (1-\gamma^2)\mathbf{I}_{n_r}. \end{aligned} \tag{36}$$

Hence, the model parameters are $\Theta^{SG} = [\mu_{\tilde{r}}, \sigma\tilde{r}, \gamma, A, \rho]$. The first four parameters are primarily related to the locationwise distributions of the S-GRF, whereas $\rho(\cdot)$ primarily relates to the spatial correlation structure.

We now briefly consider S-GRFs of low dimensionality to build some model intuition. Figure 2 presents three examples of univariate selection Gaussian pdfs and associated cumulative probability functions (cdfs), with the effect of the selection set highlighted. The geometry of the selection set A notably affects the selection Gaussian distribution and is the source of vast model flexibility, provided that the correlation γ between the basis variable r and the auxiliary variable ν is sufficiently strong. Figure 3 illustrates the selection Gaussian distribution in a multivariate setting based on the bimodal and trimodal univariate distributions shown in Figure 2.

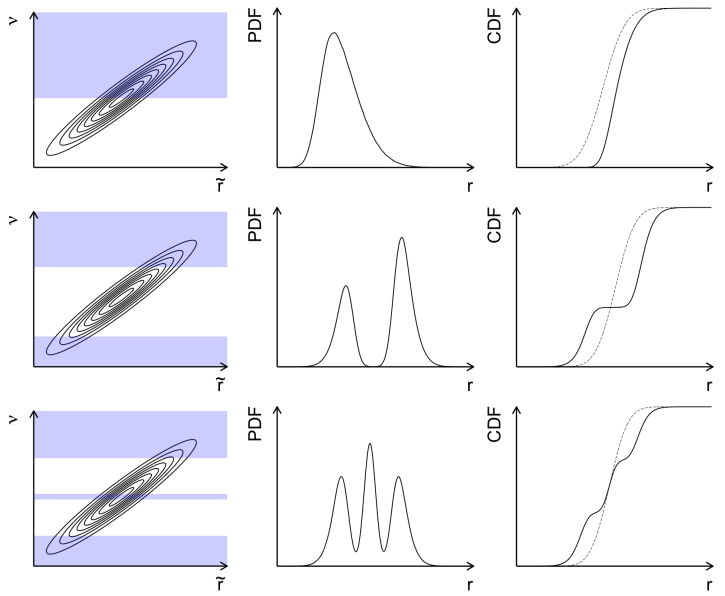

Figure 2: Univariate selection Gaussian distributions. Skewed (top row), bimodal (middle row), and trimodal (bottom row). The left column displays the shape of the selection set A superimposed on the joint distribution of $\tilde{r}$ and ν, whereas the middle and right column display the resulting pdf and cdf of r, respectively.

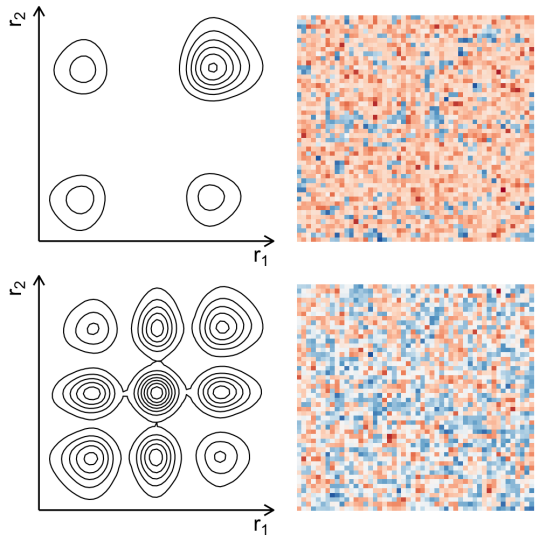

Figure 3: Multivariate selection Gaussian distributions with bimodal (top row) and trimodal (bottom row) marginals. Bivariate distributions (left column) and realizations from corresponding 2500-variate S-GRFs (right column). A moderate level of spatial correlation is used.

S-GRFs are conjugate prior models with respect to Gauss-linear likelihood models (Omre and Rimstad, 2021); hence $p(\mathbf{r}|\mathbf{d})$ is also an S-GRF and of the form

$$\begin{aligned} p(\mathbf{r}|\mathbf{d}) &= p(\tilde{\mathbf{r}}|\boldsymbol{\nu} \in \mathbf{A}, \mathbf{d}) \\ &= \frac{\Phi_{n_r}\left(\mathbf{A}; \boldsymbol{\mu}_{\nu|\tilde{r},d}, \boldsymbol{\Sigma}_{\nu|\tilde{r},d}\right)}{\Phi_{n_r}\left(\mathbf{A}; \boldsymbol{\mu}_{\nu|d}, \boldsymbol{\Sigma}_{\nu|d}\right)} \times \varphi_{n_r}\left(\tilde{\mathbf{r}}; \boldsymbol{\mu}_{\tilde{r}|d}, \boldsymbol{\Sigma}_{\tilde{r}|d}\right). \end{aligned} \tag{37}$$

The expressions for the parameters involved in the posterior model can be developed from classical Gaussian theory. The conditional expectation vectors are

$$\begin{bmatrix} \boldsymbol{\mu}_{\tilde{r}|d} \\ \boldsymbol{\mu}_{\nu|d} \end{bmatrix} = \begin{bmatrix} \mu_{\tilde{r}} \mathbf{i}_{n_r} \\ 0\mathbf{i}_{n_r} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\Sigma}_{\tilde{r}} \mathbf{G}^T \\ \boldsymbol{\Gamma}_{\nu\tilde{r}} \mathbf{G}^T \end{bmatrix} \boldsymbol{\Sigma}_d^{-1} \left(\mathbf{d} - \boldsymbol{\mu}_d\right), \tag{38}$$

and the conditional covariance matrices are

$$\begin{bmatrix} \boldsymbol{\Sigma}_{\tilde{r}|d} & \boldsymbol{\Gamma}_{\tilde{r}\nu|d} \\ \boldsymbol{\Gamma}_{\nu\tilde{r}|d} & \boldsymbol{\Sigma}_{\nu|d} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Sigma}_{\tilde{r}} & \boldsymbol{\Gamma}_{\tilde{r}\nu} \\ \boldsymbol{\Gamma}_{\nu\tilde{r}} & \boldsymbol{\Sigma}_{\nu} \end{bmatrix} - \begin{bmatrix} \boldsymbol{\Sigma}_{\tilde{r}} \mathbf{G}^T \\ \boldsymbol{\Gamma}_{\nu\tilde{r}} \mathbf{G}^T \end{bmatrix} \boldsymbol{\Sigma}_d^{-1} \begin{bmatrix} \mathbf{G}\boldsymbol{\Sigma}_{\tilde{r}} & \mathbf{G}\boldsymbol{\Gamma}_{\tilde{r}\nu} \end{bmatrix}, \tag{39}$$

hence

$$\begin{aligned}\boldsymbol{\mu}_{\nu|\tilde{r},d} &= \boldsymbol{\mu}_{\nu|d} + \boldsymbol{\Gamma}_{\nu\tilde{r}|d}\boldsymbol{\Sigma}_{\tilde{r}|d}^{-1}(\tilde{\mathbf{r}} - \boldsymbol{\mu}_{\tilde{r}|d}), \\ \boldsymbol{\Sigma}_{\nu|\tilde{r},d} &= \boldsymbol{\Sigma}_{\nu|d} + \boldsymbol{\Gamma}_{\nu\tilde{r}|d}\boldsymbol{\Sigma}_{\tilde{r}|d}^{-1}\boldsymbol{\Gamma}_{\tilde{r}\nu|d}.\end{aligned} \tag{40}$$

Recall that $\mathbf{G}$ is the forward function of the likelihood model. Furthermore, $\boldsymbol{\Sigma}_d = \mathbf{G}\boldsymbol{\Sigma}_r\mathbf{G}^T$.

The prior and posterior models are, therefore, both available once a suitable assessment strategy for S-GRFs is developed. The assessment strategy should be capable of dealing with non-stationary S-GRFs because the posterior model will, even if the prior model is stationary, tend to be non-stationary. In the following, we will discuss the assessment of a prior S-GRF model, but the same discussion applies to the assessment of the posterior model.

In the univariate case, the selection Gaussian distribution can easily be evaluated analytically by use of well known Gaussian cdfs,

$$\begin{aligned}p(r) &= p(\tilde{r}|\nu \in A) \\ &= \frac{\sum_{i=1}^{n_A}\left(\Phi_1(b_i, \mu_{\nu|\tilde{r}}, \sigma^2_{\nu|\tilde{r}}) - \Phi_1(a_i, \mu_{\nu|\tilde{r}}, \sigma^2_{\nu|\tilde{r}})\right)}{\sum_{i=1}^{n_A}\left(\Phi_1(b_i, \mu_\nu, \sigma^2_\nu) - \Phi_1(a_i, \mu_\nu, \sigma^2_\nu)\right)}\varphi_1(\tilde{r}; \mu_{\tilde{r}}, \sigma^2_{\tilde{r}}).\end{aligned} \tag{41}$$

Evaluation of the pdf becomes challenging for high dimensional S-GRFs and is usually simulation based. The simulation is performed in two steps; first the auxiliary variable $\boldsymbol{\nu}$ is simulated from the truncated Gaussian pdf $p\left(\boldsymbol{\nu}|\boldsymbol{\nu} \in \mathbf{A}\right) = I\left(\boldsymbol{\nu} \in \mathbf{A}\right)\varphi_{n_\nu}\left(\boldsymbol{\nu}; \boldsymbol{\mu}_\nu, \boldsymbol{\Sigma}_\nu\right)\left[\Phi_{n_\nu}\left(\mathbf{A}; \boldsymbol{\mu}_\nu, \boldsymbol{\Sigma}_\nu\right)\right]^{-1}$ and then the conditional basis variable $\left[\tilde{\mathbf{r}}|\boldsymbol{\nu}, \boldsymbol{\nu} \in \mathbf{A}\right]$ is simulated from the conditional Gaussian pdf $\varphi_{n_r}\left(\tilde{\mathbf{r}}; \boldsymbol{\mu}_{\tilde{r}|\nu}, \boldsymbol{\Sigma}_{\tilde{r}|\nu}\right)$. The latter step is straightforward because efficient algorithms for simulation from GRFs are available, whereas the former is more involved and typically relies on an McMC algorithm. A sensible simulation strategy is to simulate the truncated auxiliary GRF $[\boldsymbol{\nu}|\boldsymbol{\nu} \in \mathbf{A}]$ piece by piece; hence, the form of the conditional pdfs of the RF is important. For a block $\mathbf{b} \subseteq \{1, ..., n_r\}$ of size n_b, a block based decomposition of the pdf $p(\boldsymbol{\nu}|\boldsymbol{\nu} \in \mathbf{A})$ is

$$\begin{aligned}p(\boldsymbol{\nu}|\boldsymbol{\nu} \in \mathbf{A}) &= p(\boldsymbol{\nu}_b, \boldsymbol{\nu}_{b^c}|\boldsymbol{\nu}_b \in A^{n_b}, \boldsymbol{\nu}_{b^c} \in A^{n_r - n_b}) \\ &= p(\boldsymbol{\nu}_b|\boldsymbol{\nu}_{b^c}, \boldsymbol{\nu}_b \in A^{n_b})p(\boldsymbol{\nu}_{b^c}|\boldsymbol{\nu}_b \in A^{n_b}, \boldsymbol{\nu}_{b^c} \in A^{n_r - n_b}),\end{aligned} \tag{42}$$

where the subscript b^c denotes the complement of the block. In the following, we somewhere omit subscripting expectations and variances for ease of notation. The expectations and variances that appear without

a subscripted variable are to be understood as if subscripted by ν. We first consider the case of a single-site block, i.e., a block of size $n_b = 1$ that consists only of one location, $\mathbf{b} = i$, $i \in \{1, ..., n_r\}$. The associated block pdf is

$$p(\boldsymbol{\nu}_b|\boldsymbol{\nu}_{b^c}, \boldsymbol{\nu}_b \in A^{n_b}) = \frac{I(\boldsymbol{\nu}_b \in A^{n_b}) p(\boldsymbol{\nu}_b|\boldsymbol{\nu}_{b^c})}{P(\boldsymbol{\nu}_b \in A^{n_b}|\boldsymbol{\nu}_{b^c})} = \frac{I(\nu_i \in A)\varphi_1(\nu_i; \mu_{i|i^c}, \sigma^2_{i|i^c})}{\Phi_1(A; \mu_{i|i^c}, \sigma^2_{i|i^c})}. \tag{43}$$

This block pdf is a truncated Gaussian pdf and represents the locationwise full conditional pdfs of the truncated GRF. Hence, the locationwise full conditional pdfs of the truncated GRF are computationally easily available, which entails that single-site Gibbs sampling may be feasible. Single-site Gibbs sampling is a viable and efficient simulation strategy for RFs with weak spatial correlation structure that simulates sequentially from the locationwise full conditional distributions. The single-site Gibbs simulation algorithm is presented in pseudo code in Algorithm 1.

Algorithm 1: Do k simulation sweeps of $\left[\boldsymbol{\nu}|\boldsymbol{\nu} \in \mathbf{A}\right]$ by single-site Gibbs.

Precompute conditioning weights $\mathbf{k}_i$ and conditional marginal variances $\sigma^2_{i|i^c}$:

For i **from** 1 **to** n_r
  $\mathbf{k}_i = \boldsymbol{\Sigma}_{i,i^c}\boldsymbol{\Sigma}^{-1}_{i^c}$.
  $\sigma^2_{i|i^c} = \sigma^2_i - \mathbf{k}_i\boldsymbol{\Sigma}^T_{i,i^c}$.
End

Initialize $\boldsymbol{\nu}^0 \in \mathbf{A}$.

For j **from** 1 **to** k
  Set $\boldsymbol{\nu}^j = \boldsymbol{\nu}^{j-1}$
  For i **from** 1 **to** n_r
    Compute conditional mean and simulate:

$$\mu^j_{i|i^c} = \mu_i + \mathbf{k}_i\left(\boldsymbol{\nu}^j_{i^c} - \boldsymbol{\mu}_{i^c}\right).$$

$$\nu^j_i \sim \frac{I(\nu^j_i \in A)\varphi_1\left(\nu^j_i; \mu^j_{i|i^c}, \sigma^2_{i|i^c}\right)}{\Phi_1\left(A; \mu^j_{i|i^c}, \sigma^2_{i|i^c}\right)}.$$

  End
End

Note that the conditioning weights $\mathbf{k}_i$ are identical for locations that are not influenced by border effects in stationary S-GRFs, which may be exploited for more efficient precomputation. In non-stationary S-GRFs the conditioning weights tend to be unique for each location. Moreover, computation of the parameters involved in the full conditional marginals may be prohibitive if the grid under study is large and the range of the spatial correlation is long. In such situations it may be necessary to reduce computational time by approximating the full conditionals as $\pi(\nu_i|\boldsymbol{\nu}_{i^c}) \approx \pi(\nu_i|\boldsymbol{\nu}_{i^n})$, with $\boldsymbol{\nu}_{i^n}$ consisting of random variables that are notably correlated with ν_i.

The feasibility of the single-site Gibbs algorithm depends on the properties of the locationwise full conditional pdfs. These pdfs become increasingly constrained with increasing correlation in the RF, which is detrimental to the mixing of the algorithm in multimodal settings. Blockwise sampling, which is based on a partition of the grid into blocks consisting of collections of adjacent grid points, offers a possible solution to the limitations of single-site Gibbs sampling. The sampling is performed blockwise and sequentially, which enables reduction of the influence of the RF outside the block on the locationwise marginals within the block by strategic inter-block sampling. The pdf of a block $\mathbf{b} \subseteq \{1, ..., n_r\}$ of size $n_b > 1$ is

$$\begin{aligned} p(\boldsymbol{\nu}_b|\boldsymbol{\nu}_{b^c}, \boldsymbol{\nu}_b \in \mathbf{A}^{n_b}) &= \frac{I(\boldsymbol{\nu}_b \in \mathbf{A}^{n_b}) p(\boldsymbol{\nu}_b|\boldsymbol{\nu}_{b^c})}{P(\boldsymbol{\nu}_b \in \mathbf{A}^{n_b}|\boldsymbol{\nu}_{b^c})} \\ &= \frac{\prod_{i\in\mathbf{b}} I(\nu_i \in A)\varphi_1(\nu_i; \mu_{i|b^{c,v}}, \sigma^2_{i|b^{c,v}})}{\Phi_{n_b}(A^{n_b}; \boldsymbol{\mu}_{b|b^c}, \boldsymbol{\Sigma}_{b|b^c})}. \end{aligned} \tag{44}$$

Here, the subscript $b^{c,v}$ denotes the union of the complement of the block and the already visited locations within the block. Note that the normalization constant $\Phi_{n_b}(A^{n_b}; \boldsymbol{\mu}_{b|b^c}, \boldsymbol{\Sigma}_{b|b^c})$ can not be expressed in a sequentially conditional form; hence, assessment of the block pdf requires evaluation of a high dimensional Gaussian orthant probability, which is challenging. Therefore, we use the Metropolis Hastings approach presented in Rimstad and Omre (2014) with proposal

$$q(\boldsymbol{\nu}_b|\boldsymbol{\nu}_{b^c}, \boldsymbol{\nu}_b \in \mathbf{A}^{n_b}) = \prod_{i\in b} I(\nu_i \in A) \frac{\varphi_1\left(\nu_i; \mu_{i|b^{c,v}}, \sigma^2_{i|b^{c,v}}\right)}{\Phi_1\left(A; \mu_{i|b^{c,v}}, \sigma^2_{i|b^{c,v}}\right)}, \tag{45}$$

and associated Metropolis-Hastings acceptance probability

$$\alpha(\boldsymbol{\nu}_b'|\boldsymbol{\nu}_b) = \min\left\{1, \frac{p(\boldsymbol{\nu}_b'|\boldsymbol{\nu}_{b^c}, \boldsymbol{\nu}_b \in \mathbf{A}^{n_b})}{p(\boldsymbol{\nu}_b|\boldsymbol{\nu}_{b^c}, \boldsymbol{\nu}_b \in \mathbf{A}^{n_b})} \cdot \frac{q(\boldsymbol{\nu}_b|\boldsymbol{\nu}_{b^c}, \boldsymbol{\nu}_b \in \mathbf{A}^{n_b})}{q(\boldsymbol{\nu}_b'|\boldsymbol{\nu}_{b^c}, \boldsymbol{\nu}_b \in \mathbf{A}^{n_b})}\right\}$$
$$= \min\left\{1, \prod_{i \in b} \frac{\Phi_1\left(A; \mu'_{i|b^{c,v}}, \sigma^2_{i|b^{c,v}}\right)}{\Phi_1\left(A; \mu_{i|b^{c,v}}, \sigma^2_{i|b^{c,v}}\right)}\right\}.$$

Note that the expectation in the nominator of the product is taken with respect to $\boldsymbol{\nu}_b'$, whereas the expectation in the denominator is taken with respect to $\boldsymbol{\nu}_b$. The initialization of the blockwise M-H McMC algorithm is presented in Algorithm 2 and the simulation algorithm is presented in Algorithm 3.

Algorithm 2: Initialize blockwise McMC simulation algorithm.

Partition the grid into blocks $\mathbf{b}_i \subseteq \{1, ..., n_r\}$ of sizes n_{b_i}, with associated block complements $\mathbf{b}_i^c = \{1, ..., n_r\} / \mathbf{b}_i$, $i = 1, ..., N_b$.

Construct inter-block sampling orders $\mathbf{O}_i, i = 1, ..., N_b$.

Precompute inter-block conditioning weights $\mathbf{k}_i$ and conditional marginal variances $\sigma^2_{i|b^{c,v}}$:

For i from 1 **to** N_b
  For j from 1 **to** n_{b_i}
    Set m to be the j-th element of $\mathbf{O}_i$.
    $\mathbf{k}_m = \boldsymbol{\Sigma}_{m,b_i^{c,v}} \boldsymbol{\Sigma}^{-1}_{b_i^{c,v}}$.
    $\sigma^2_{m|,b_i^{c,v}} = \sigma^2_m - \mathbf{k}_m \boldsymbol{\Sigma}^T_{m,b_i^{c,v}}$.
  End
End

Initialize $\boldsymbol{\nu}^0 \in \mathbf{A}$.

Algorithm 3: Do k simulation sweeps of $\left[\boldsymbol{\nu}|\boldsymbol{\nu} \in \mathbf{A}\right]$ by blockwise M-H McMC simulation.

Initialize according to Algorithm 2.

For j **from** 1 **to** k
  Set $\boldsymbol{\nu}^j = \boldsymbol{\nu}^{j-1}$
  For i **from** 1 **to** N_b
    For l **from** 1 **to** n_{b_i}
      Set m to be the l-th element of $\mathbf{O}_i$.
      Compute conditional mean and simulate:

$$\mu^j_{m|b_i^{c,v}} = \mu_m + \mathbf{k}_m \left(\boldsymbol{\nu}^j_{b_i^{c,v}} - \boldsymbol{\mu}_{b_i^{c,v}} \right).$$

      Simulate proposal $\nu'_{m|b_i^{c,v}}$ from $q(\nu'_{m|b_i^{c,v}})$.
    End
    Accept $\boldsymbol{\nu} = [\boldsymbol{\nu}'_{b_i|b_i^c}, \boldsymbol{\nu}_{b_i^c}]^T$ with probability $\alpha(\boldsymbol{\nu}'_b|\boldsymbol{\nu}_b)$.
  End
End

Note that both the presented simulation approaches rely on simulation from conditional locationwise pdfs, and that simulation within the selection set A is necessary at each grid location. Rejection sampling from $\varphi_1(\nu_i; \mu_{i|b^c}, \sigma^2_{i|b^c})$ is a natural approach, but its efficiency depends on the degree of support $\varphi_1(\nu_i; \mu_{i|b^c}, \sigma^2_{i|b^c})$ has for A. There is an inverse relationship between the computational efficiency and the distance between $\mu_{i|b^c}$ and the selection set A. However, this should not be an issue in sensibly defined S-GRFs, because the modeler has full control over the position of A. Nevertheless, unsensible models that impede efficient simulation from the truncated Gaussian distributions can arise in conditional S-GRFs.

The cdf associated with $[\nu_i|\nu_i \in A]$ is a piecewise rescaled Gaussian cdf and can be related to the Gaussian cdf $\Phi_1\left(\nu_i; \mu_i, \sigma_i^2\right)$, which can easily be used to sample $[\nu_i|\nu_i \in A]$ by inverse transform sampling. This approach is not sensitive to the location of A in terms of computational efficiency and can produce samples in fixed computation time. The method of inverse transform sampling is illustrated in Figure 4 for the truncated Gaussian distributions associated with the univariate bimodal and trimodal selection Gaussian distributions displayed in Figure 2. A value u of the truncated Gaussian cdf is simulated from $\mathcal{U}[0, 1]$ and the associated value of the Gaussian cdf, u_G is computed. Lastly, the simulated variable ν_u is obtained from the inverse Gaussian cdf applied to u_G. Inverse

transform sampling from a univariate truncated normal distribution is outlined in Algorithm 4.

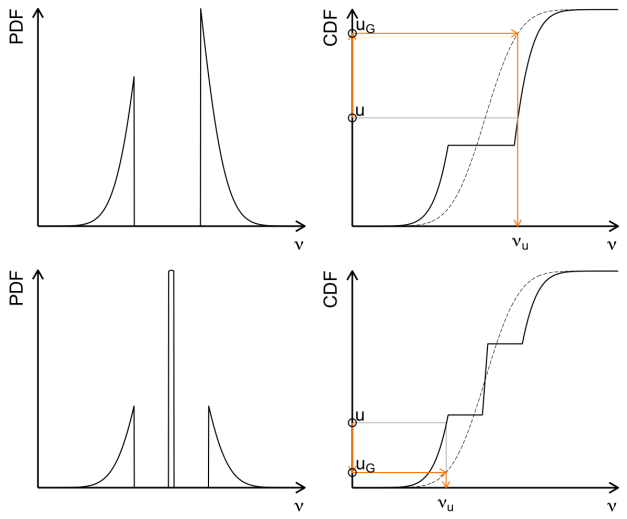

Figure 4: Inverse transform sampling from a bimodal truncated Gaussian distribution (top row) and a trimodal truncated Gaussian distribution (bottom row). In the right column, the solid black lines represent the selection Gaussian cdfs, and the dashed black lines represent the Gaussian cdfs. Moreover, the orange lines indicate the flow of the inverse transform algorithm and the grey lines mark the naive inverse transform.

Algorithm 4: Draw a sample $\left[\nu|\nu \in A\right]$ by inverse transform sampling.

Compute $p(A_i) = \Phi(b_i; \mu_\nu, \sigma_\nu^2) - \Phi(a_i; \mu_\nu, \sigma_\nu^2), \qquad i = 1, 2, ..., n_A$
Construct $\mathbf{s}$ with entries $s_0 = 0$ and $s_j = \sum_{i=1}^{j} p(A_i)$,
$j = 1, 2, ..., n_A$
Define normalizing constant $C = s_{n_A}$
Generate $u \sim \mathcal{U}[0, 1]$.
Find the mode m satisfying $s_{m-1} \leq Cu \leq s_m$,
$m = 1, 2, ..., n_A$
Compute u relative to the Gaussian cdf,
$u_G = \Phi(a_m, \mu_\nu, \sigma_\nu^2) + C(u - \frac{s_{m-1}}{C})$
Compute $\nu \in A$ from $\Phi^{-1}(u_G; \mu_\nu, \sigma_\nu^2)$.

References

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

Allard, D. and Naveau, P. (2007). A New Spatial Skew-Normal Random Field Model. *Communications in Statistics - Theory and Methods*, 36:1821–1834.

Arellano-Valle, R. B., Branco, M. D., and Genton, M. G. (2006). A Unified View on Skewed Distributions Arising from Selections. *The Canadian Journal of Statistics*, 34:581–601.

Azzalini, A. (1985). A Class of Distributions Which Includes the Normal Ones. *Scandinavian Journal of Statistics*, 12(2):171–178.

Azzalini, A. (2013). *The Skew-Normal and Related Families*. Cambridge University Press.

Backus, G. E. (1989). Confidence Set Inference with a Prior Quadratic Bound. *Geophysical Journal International*, 97:119–150.

Besag, J. (1974). Spatial Interaction and the Statistical Analysis of Lattice Systems. *Journal of the Royal Statistical Society: Series B (Methodological)*, 36:192–225.

Bortfeld, R. (1961). Approximations to the Reflection and Transmission Coefficients of Plane Longitudinal and Transverse Waves. *Geophysical Prospecting*, 9(4):485–502.

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

Casella, G. (1985). An Introduction to Empirical Bayes Data Analysis. *The American Statistician*, 39(2):83–87.

Casella, G. and Berger, R. L. (2001). *Statistical Inference*. Cengage Learning.

Chilès, J.-P. and Delfiner, P. (1999). *Geostatistics: Modeling Spatial Uncertainty*. Wiley.

Donoho, D. L. (1989). Statistical estimation and optimal recovery, Technical report No. 214. *Department of Statistics, University of California, Berkeley*.

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

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

Fjeldstad, T., Avseth, P., and Omre, H. (2021). A one-step Bayesian inversion framework for 3D reservoir characterization based on a Gaussian mixture model - A Norwegian Sea demonstration. *Geophysics*, 86(2):R221–R236.

Fjeldstad, T. M. and Grana, D. (2018). Joint probabilistic petrophysics-seismic inversion based on Gaussian mixture and Markov chain prior models. *Geophysics*, 83:1JF–Z3.

Gamerman, D. and Hedibert, F. L. (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, volume 2. Chapman & Hall/CRC Taylor & Francis Gruop.

Gelman, A. and Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7(4):457–472.

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. M., and Omre, H. (2017). Bayesian Gaussian Mixture Linear Inversion for Geophysical Inverse Problems. *Mathematical Geosciences*, 49:493–515.

Gunning, J. and Glinsky, M. E. (2004). Delivery: an open-source model-based Bayesian seismic inversion program. *Computers & Geosciences*, 30(6):619–636.

Hadamard, J. (1902). Sur les problèmes aux dériveés partielles et leur signification physique. *Princetion Univ. Bull.*, 13:49–52.

Hadamard, J. (1923). Lectures on Cauchy's Problems in Linear Partial Differential Equations. *Yale University Press*.

Hammersley, J. M. and Clifford, P. (1971). Markov fields on finite graphs and lattices. *Not submitted for publication*.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109.

Johnson, R. A. and Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis*. Pearson Prentice Hall, Upper Saddle River, 6 edition.

Karimi, O., Omre, H., and Mohammadzadeh, M. (2010). Bayesian closed-skew Gaussian inversion of seismic AVO data for elastic material properties. *Geophysics*, 75:R1–R11.

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

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

Larsen, A. L., Ulvmoen, M., Omre, H., and Buland, A. (2006). Bayesian lithology/fluid prediction and simulation on the basis of a Markov-chain prior model. *Geophysics*, 71:R69–R78.

Lindberg, D. V. and Omre, H. (2014). Blind categorical deconvolution in two-level hidden Markov models. *IEEE Transactions on Geoscience and Remote Sensing*, 52(11):7435–7447.

Mosegaard, K. and Tarantola, A. (1995). Monte Carlo sampling of solutions to inverse problems. *Journal of Geophysical Research*, 100:12431–12447.

Omre, H. and Rimstad, K. (2021). Bayesian Spatial Inversion and Conjugate Selection Gaussian Prior Models. *Journal of Uncertainty Quantification*, To appear.

Rimstad, K., Avseth, P., and Omre, H. (2012). Hierarchical Bayesian lithology/fluid prediction: A North Sea case study. *Geophysics*, 77:B69–B85.

Rimstad, K. and Omre, H. (2013). Approximate posterior distributions for convolutional two-level hidden Markov models. *Computational Statistics & Data Analysis*, 58(1):187–200.

Rimstad, K. and Omre, H. (2014). Skew-Gaussian Random Fields. *Spatial Statistics*, 10:43–62.

Roberts, G. O. and Smith, A. F. M. (1994). Simple conditions for the convergence of the Gibbs sampler and Metropolis-Hastings algorithms. *Stochastic Processes and their Applications*, 49(2):207–216.

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

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

Shuey, R. T. (1985). A simplification of the Zoeppritz equations. *Geophysics*, 50(4):609–614.

Stark, P. B. (1992). Minimax confidence intervals in geomagnetism. *Geophysical Journal International*, 108:329–338.

Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. SIAM.

Tikhonov, A. N. (1963). Solution to incorrectly formulated problems and the regularization method. *Soviet Mathematics Doklady*, 4:1035–1038.

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

Walden, A. T. and White, R. E. (1998). Seismic Wavelet Estimation: A Frequency Domain Solution to a Geophysical Noisy Input-Output Problem. *IEEE Transactions on Geoscience and Remote Sensing*, 36(1):287–297.

Zoeppritz, K. B. (1919). Vllb. On reflection and transmission of seismic waves by surfaces of discontinuity. *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-physikalische Klasse*, pages 66–84.

Summary of papers

Paper l: Bayesian seismic amplitude variation with offset inversion for reservoir variables with bimodal spatial histograms

Ole Bernhard Forberg, Øyvind Kjøsnes, and Henning Omre

Published in *Geophysics*, 2021, vol. 86, no. 3.

Abstract: An oil reservoir in the Kneler field in the Alvheim oil and gas field in the North sea is studied. The reservoir is studied along a vertical 1D profile through which the reservoir is primarily lithologically homogeneous. Therefore, the reservoir profile can be adequately characterized by the petrophysical properties porosity and water saturation. The reservoir is considered at the initial state; hence, gravitational effects dominate and keep hydrocarbons from mixing with water. Histograms based on measured water saturations along the reservoir profile are consequently clearly bimodal, which is challenging to model. Reservoir characterization along the reservoir profile by seismic amplitude variation with offset (AVO) inversion is carried out in a Bayesian framework. The prior spatial model for porosity and water saturation is specified to be a selection Gaussian random field (S-GRF), which is capable of representing multimodal spatial variables. We use linear models for the seismic and rock physics likelihood models, which entail that the posterior model is also an S-GRF and that the bimodal characteristics of water saturation can be reproduced. The methodology is defined and demonstrated on two synthetic cases inspired by the real data, and is thereafter applied to the real case. The S-GRF inversion results are compared to results from a standard Gaussian model, and are considered to be substantial improvements.

Main contribution: Formulation of a Bayesian S-GRF seismic inversion framework for reservoir variables and comparison with an alternative Bayesian Gaussian random field seismic inversion framework. The methodology is demonstrated on real seismic data.

Paper ll: Bayesian inversion of Time-Lapse Seismic AVO Data for Multimodal Reservoir Properties

Ole Bernhard Forberg, Dario Grana, and Henning Omre

Published in *IEEE: Transactions on Geoscience and Remote Sensing*, 2021, vol. 59, no. 11.

Abstract: Reservoir fluid monitoring by inversion of time-lapse seismic amplitude variation with offset (AVO) data for porosity and water saturation is studied. The inverse problem is cast into a Bayesian framework and a selection Gaussian random field (S-GRF) prior spatial model is assigned for porosity and water saturation. Water saturation is of particular monitoring interest since production entails fluid replacement; hence, the time-lapse model incorporates water saturation at different time points. The likelihood model is Gauss-linear and is based on linearized seismic and rock physics models, which entails that the associated posterior model is also an S-GRF, with analytically assessable parameters. Hence, the posterior model is capable of representing multimodality and abrupt spatial mode transitions. Two realistic case studies are considered. The first case study is on production of an oil reservoir in the North Sea, where the movement of oil-water-contacts along a vertical profile is of interest. The second case study is on injection of CO_2 into a subsurface potential CO_2 storage unit, where the expansion of the CO_2 region in a cross section is of interest. In both case studies, the main property of interest can be inferred from the changes in the predicted water saturations. Based on the inversion results, which are considered to be very encouraging, the proposed statistical formulation appears to be particularly well suited for fluid monitoring problems of the described type.

Main contribution: Extension of the Bayesian S-GRF seismic inversion framework to time-lapse reservoir characterization. The methodology is applied to two synthetic cases and is demonstrated to be well suited for fluid monitoring problems.

Paper lll: Bayesian seismic AVO inversion using a laterally coupled multimodal prior model

Ole Bernhard Forberg, Øyvind Kjøsnes, and Henning Omre

To appear in *IEEE: Transactions on Geoscience and Remote Sensing.*

Abstract: A 2D section and a 3D volume of an oil reservoir in the Kneler field in the Alvheim oil and gas field in the North Sea is studied. The chosen reservoir zones are lithologically heterogeneous with considerable variability in the clay content; hence, we propose a reservoir characterization scheme including volume of clay in addition to porosity and water saturation. The characterization is based on Bayesian inversion of seismic amplitude variation with offset (AVO) data. We use a laterally coupled prior model for porosity, water saturation, and volume of clay, which may reduce the impact of trace unique signal errors in the seismic AVO data. The reservoir is considered at the initial state, which entails gravitationally induced bimodal characteristics in water saturation due to poor fluid mixing. Therefore, the prior model is specified to be a selection Gaussian random field (S-GRF), which can support multimodality. The likelihood model is linear and Gaussian, and consequently the posterior model is also an S-GRF; hence, the form of the posterior model is known and its parameters can be analytically computed. The inversion results of real seismic AVO data from the two reservoir zones appear to be reliable along validation wells. Moreover, a notable variance reduction is demonstrated in the posterior model based on the prior model with lateral coupling, compared to a posterior model based on an alternative prior model without.

Main contribution: Formulation of a three-dimensional laterally coupled Bayesian S-GRF seismic inversion framework and comparison with a trace-by-trace approach. Moreover, the prior model is specified to be lithologically robust and to support different anisotropy for the reservoir variables. The methodology is demonstrated on real seisimc data.

Synthesis

Gaussian prior models are often adopted in the framework of Bayesian inversion due to their parametric simplicity and conjugate properties. A Gaussian prior model is conjugate with respect to linear and Gaussian likelihood models, i.e., the posterior model is also Gaussian. This conjugate property entails that predictive quantities can be analytically computed by well known Gaussian conditioning formulas, which is very computationally efficient. Although Gaussian prior models are suitable for a vast range of phenomena, they can be inadequate for some, including phenomena that exhibit multimodality and abrupt spatial mode transitions. For these types of phenomena, a selection Gaussian prior model, which has support for multimodality, can be more representative. A selection Gaussian prior model is, like a Gaussian prior model, conjugate with respect to linear and Gaussian likelihood models. However, the posterior model assessment is simulation based and computationally more expensive. In Paper l, we characterize a reservoir along a vertical 1D profile in terms of porosity and water saturation. The reservoir is studied at the initial state, which entails gravitationally induced bimodality in water saturation, due to poor fluid mixing. We formulate a Bayesian selection Gaussian seismic inversion framework through definition of a linear and Gaussian likelihood model and a selection Gaussian random field prior model and demonstrate that the bimodal characteristics of water saturation can be reliably reproduced in predictions. Moreover, the bimodal characteristics are shown to be less accurately reproduced in predictions based on an alternative Gaussian model. The methodology is tested on real seismic data. In Paper lll, we formulate a laterally coupled three-dimensional selection Gaussian prior model and apply it to a 2D case and a 3D case. Due to lithological heterogeneity, the resevoir zones are characterized in terms of porosity, water saturation, and volume of clay. The main objective of the paper is to investigate whether lateral coupling in the prior model is advantageous compared with a trace-by-trace approach to seismic inversion, in light of seismic data having strong lateral continuity. The methodology is tested on real seisimc data and we conclude that lateral coupling in the prior model has a positive effect. In Paper ll, we apply the Bayesian selection Gaussian seismic inversion framework to time-lapse characterization of a CO_2 reservoir, in terms of porosity and water saturation. The focus of the paper is on fluid monitoring through prediction of water saturation at two time points. Because the seismic signal from CO_2 reservoirs are relatively strong, we perform the seismic inversion trace-by-trace. The obtained predictions are highly accurate and have clearly defined regions of gas and water. This suggests

that trace-by-trace inversion of seismic data from CO_2 reservoirs is highly suitable; hence, large 3D volumes can be characterized at relatively low computational cost.

Bayesian seismic amplitude variation with offset inversion for reservoir variables with bimodal spatial histograms

Ole Bernhard Forberg, Øyvind Kjøsnes, and Henning Omre

Published in *Geophysics*, 2021, vol. 86, no. 3.

DOI: 10.1190/geo2020-0086.1

URL: https://library.seg.org/doi/abs/10.1190/geo2020-0086.1

GEOPHYSICS, VOL. 86, NO. 3 (MAY-JUNE 2021); P. R331–R350, 27 FIGS., 6 TABLES.
10.1190/GEO2020-0086.1

Bayesian seismic amplitude variation with offset inversion for reservoir variables with bimodal spatial histograms

Ole Bernhard Forberg[1], Øyvind Kjøsnes[2], and Henning Omre[1]

ABSTRACT

We consider seismic amplitude variation with offset (AVO) inversion for prediction of the reservoir properties porosity and water saturation. An oil reservoir at the initial state is studied; hence, gravitational effects dominate and keep hydrocarbons from mixing with water. Histograms of observations of water saturation along wells are consequently clearly bimodal, which is challenging to model. The seismic AVO inversion is cast into a Bayesian framework. The prior spatial model for porosity and water saturation is specified to be a selection Gaussian random field (S-GRF), which is capable of representing spatial variables with multimodal histograms. By using linear models for the seismic and rock-physics likelihoods, the posterior model is also an S-GRF. Hence, the Bayesian seismic inversion can be solved analytically, and the bimodal characteristics of water saturation can be reproduced. The methodology is defined and demonstrated on two synthetic cases inspired by real data from an oil reservoir, and thereafter applied to the real case. The well observations are fairly accurately reproduced and the inversion results are considered to be substantial improvements compared to standard spatial Gaussian models.

INTRODUCTION

Reliable reservoir characterization is important for reservoir management and drilling decisions. Geoscientific experience and seismic data with good spatial coverage along with well-log data provide the basis for this characterization. Seismic data are a wavelet convolved source of reservoir information, subject to multiple sources of uncertainty. Extracting the reservoir information contained in these data constitutes the challenging seismic inverse problem.

Seismic amplitude variation with offset (AVO) inversion is a widely used inversion scheme in which the ultimate goal is the prediction of lithology and fluid (LF) properties (Larsen et al., 2006; Buland et al., 2008); hence, it is highly applicable for reservoir characterization. The technique uses the relationship that the strength of the seismic reflections is dependent on the angle of incidence of the seismic waves on interfaces in the subsurface, and the understanding of this dates back to the conception of the Zoeppritz (1919) equations, which establish a dependence of the seismic reflections on the elastic rock properties (compressional wave [P-wave] and shear wave [S-wave] velocity and mass density). Characterizing a reservoir in terms of LF properties therefore also requires a rock-physics model that links these elastic rock properties to the LF properties of interest, called reservoir properties. The rock-physics model can, for example, be based on an empirical relation obtained by fitting well-log observations (Mavko et al., 2009).

Bayesian inversion is a common framework for many geophysical problems (Tarantola, 2005). This approach provides reliable reservoir characterization because the solution consists not only of predictions, but it also provides information about their associated uncertainty. The Bayesian approach requires the definition of a likelihood model and a prior model. The likelihood model is a probabilistic geophysical model relating the reservoir properties to the seismic AVO data, and the prior model reflects our a priori beliefs about the reservoir properties and acts as an inversion constraint. The ultimate solution in the framework of Bayesian inversion is the posterior model, which is uniquely defined by the likelihood and prior models.

The traditional assumptions in Bayesian seismic AVO inversion are that the likelihood model is linear and Gaussian (Gauss-linear) and that the prior model is a Gaussian random field (GRF) (Buland and Omre, 2003). Assuming a Gauss-linear likelihood model ensures conjugacy for the Gaussian prior distribution, that is, that the form of the posterior is the same as the form of the prior. Consequently, the solution can be analytically evaluated. The spatial

Manuscript received by the Editor 12 February 2020; revised manuscript received 24 November 2020; published ahead of production 19 January 2021; published online 08 April 2021.
[1]Norwegian University of Science and Technology, Department of Mathematical Sciences, Trondheim 7034, Norway. E-mail: ole.b.forberg@ntnu.no (corresponding author); omre@math.ntnu.no.
[2]Aker BP ASA, Trondheim, Norway. E-mail: oyvind.kjosnes@akerbp.com.

reservoir variables are usually represented on a very large grid; hence, the computational demand of the inversion is important and the analytical tractability is favorable (Buland et al., 2003). A frequently used approach to evaluate the posterior model in the absence of these assumptions is Markov chain Monte Carlo (MCMC) methods (Mosegaard and Tarantola, 1995). These algorithms are iterative and usually propose samples from the prior model, each of which are accepted or rejected according to the Metropolis rule applied to the likelihood of the current and proposed samples. The convergence rates of these algorithms are usually low, which may cause problems in real high-dimensional studies. Alternatively, nonparametric kernel estimation of the posterior model may be an option (Ayani and Grana, 2020), but for seismic data these methods tend to become computationally unfeasible.

The nonlinear Zoeppritz equations are a natural point of departure for defining the likelihood model, and many seismic AVO inversion schemes use the Aki and Richards (1980) approximation of these equations. In the traditional Bayesian framework for seismic AVO inversion introduced in Buland and Omre (2003), the likelihood model is Gauss-linear and is based on a linearization of the Aki and Richards (1980) approximation. If the rock-physics relationship is nonlinear, but only weakly so, the Gauss-linearity of the likelihood model can be preserved by using a linear approximation (Landrø, 2001).

Many reservoir properties, such as porosity and water saturation, are naturally multimodal because of varying underlying LF classes (Grana and Della Rossa, 2010). Classifying reservoir zones in terms of such classes are prevailing and important problems (Larsen et al., 2006; Ulvmoen et al., 2010). These are categorical inverse problems in a spatial setting, usually modeled by Markov random fields (Besag, 1974; Eidsvik et al., 2004). Another common approach is truncated Gaussian simulation (Matheron et al., 1987). Spatial variables that are continuous, with multimodal histograms over value per location (spatial histograms), due to different underlying LF classes, are challenging to model. The Gaussian mixture model offers an alternative for modeling multimodality (Hasselblad, 1966); however, in spatial inversion it is necessary to model a spatially defined mode indicator, and Markov random fields are frequently used for this purpose (Rimstad et al., 2012; Grana et al., 2017). Assessment of the resulting Gaussian mixture random fields relies on MCMC algorithms that are computationally demanding, particularly when conditioned on convolved seismic data.

In this study, we use a Gauss-linear likelihood model based on the linear Aki and Richards (1980) approximation of the Zoeppritz equations and an approximated linear rock-physics model. We explore an alternative prior model for continuous spatial variables with multimodal spatial histograms; the so-called selection Gaussian random field (S-GRF), developed and coined as such in Omre and Rimstad (2018) based on the selection concept introduced in Arellano-Valle et al. (2006). This random field (RF) may represent continuous spatial variables with spatial histograms being multimodal, skewed, or peaked. Moreover, the S-GRF prior model is demonstrated to be conjugate subject to a Gauss-linear likelihood model (Omre and Rimstad, 2018), which makes it suitable as a prior model in the Bayesian seismic AVO inversion framework. We demonstrate the model on two synthetic cases inspired by real reservoir data, and we test its practical capacity on a real data set from the Alvheim oil and gas field in the North Sea.

NOTATION

The n-dimensional vector of all ones is denoted $\mathbf{i}_n$, and $\mathbf{I}_n$ is the $n \times n$ identity matrix. Furthermore, let $p(\cdot)$ be a probability density function (pdf) of its argument, and the corresponding cumulative distribution function is $P(\cdot)$. The Boolean function $I(\cdot)$ is the indicator function, and it is for logical argument A defined as $I(A) = 1$ if A is true, and $I(A) = 0$ otherwise. Finally, the superscript T is used to indicate the matrix transpose.

PROBLEM DEFINITION

We consider a 1D profile along a vertical well in an oil reservoir whose dominant lithology is sandstone. The goal of our study is to assess the storage capacity and fluid content in the reservoir along the vertical profile, which we will refer to as the reservoir profile. For this purpose, we examine the distribution of porosity ϕ and water saturation s_w hereafter referred to as saturation, along the profile.

The reservoir profile is discretized into a regular grid $\mathcal{L}_r$, consisting of n_r grid points. Let $r(z) = [\phi(z), s_w(z)]^T$ be the reservoir variables at depth $z \in \mathcal{L}_r$. Furthermore, $\boldsymbol{\phi} = [\phi(z_1), \phi(z_2), \ldots, \phi(z_{n_r})]^T$ and $\mathbf{s}_w = [s_w(z_1), s_w(z_2), \ldots s_w(z_{n_r})]^T$ are the porosity and saturation on all of $\mathcal{L}_r$, respectively. The $2n_r$-dimensional vector $\mathbf{r} = [\boldsymbol{\phi}, \mathbf{s}_w]^T$ is the reservoir variables on all of $\mathcal{L}_r$.

Seismic data are our gateway to predict $\mathbf{r}$. We have seismic AVO data from the reservoir profile, that is, seismic data gathered with a common midpoint, but at different offsets. Seismic data are collected along the profile at n_θ offset angles, and they are processed to a regular grid of times $\mathcal{L}_d$, consisting of n_d grid points. This seismic grid is much coarser than the reservoir variable grid $\mathcal{L}_r$. The seismic AVO data on all of $\mathcal{L}_d$ are represented by the $n_\theta n_d$-dimensional vector $\mathbf{d}$.

Data from the Kneler discovery in the Alvheim oil and gas field in the North Sea are used in our study. Seismic data were acquired in 1996 before the initiation of hydrocarbon production. In Figure 1, well porosities and saturations from well logging the reservoir profile are displayed. Although sandstone dominates the lithology of the profile, it contains four thin layers of shale, the locations of which are indicated by the transparent gray zones. The saturation increases consistently in all these shale layers, but the porosity seems to be affected more strongly in the second layer from the top than in the others. Moreover, it can be seen that the oil-water-contact (OWC) is located at approximately 2132 m in depth. The lower part of Figure 1 presents spatial histograms of the data. We observe that the distribution of porosity appears to be unimodal and symmetric, whereas the distribution of saturation, representing the fluid content, appears to be bimodal.

The seismic AVO data associated with the reservoir profile are presented in Figure 2. These reflections result from heterogeneities in the elastic rock properties along the well profile. The elastic rock properties are represented by the P- and S-wave velocities and the mass density, denoted by V_P, V_S, and ρ, respectively. For modeling purposes, we represent the natural logarithm of the elastic rock variables on all of $\mathcal{L}_r$ by the $3n_r$-dimensional vector $\mathbf{m} = [\log(\mathbf{V}_P), \log(\mathbf{V}_S), \log(\boldsymbol{\rho})]^T$, and we hereafter refer to $\mathbf{m}$ as the elastic rock variables. Observations of the elastic rock properties along the reservoir profile are available, but these observations are only used to assess the model parameters; hence, they are not displayed.

To demonstrate and evaluate the suggested Bayesian seismic AVO inversion technique, we introduce two synthetic case studies strongly inspired by the observations from Kneler. The synthetic

case studies are presented in parallel with the methodology, and after demonstrating the inversion method on synthetic data, we also invert the real seismic observations from Kneler.

Synthetic case studies

The well observations used in our synthetic case studies are presented in Figures 3 and 4, corresponding to cases A and B, respectively. The two case studies share a lot of similarities because they are both constructed from the observations in Figure 1. In both cases, an extension of the reservoir profile is introduced to avoid the border problems associated with generating synthetic convolved seismic data. The top and bottom of the reservoir profile are indicated by dashed horizontal lines. Because the well observations are highly spatially correlated, we thin the reservoir profile grid $\mathcal{L}_r$ to

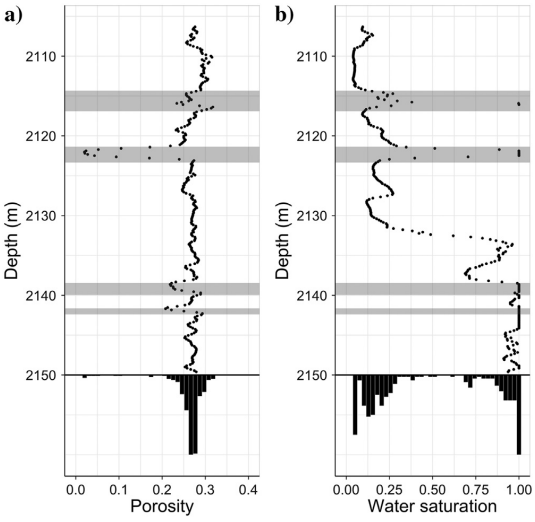

Figure 1. Well logs from the reservoir profile in Kneler in the Alvheim oil and gas field, collected in 1996: (a) porosity and (b) saturation profiles.

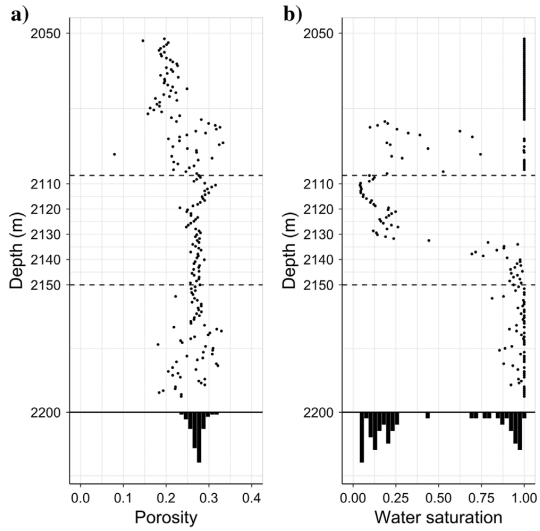

Figure 3. Synthetic well in case A: (a) porosity and (b) saturation profiles.

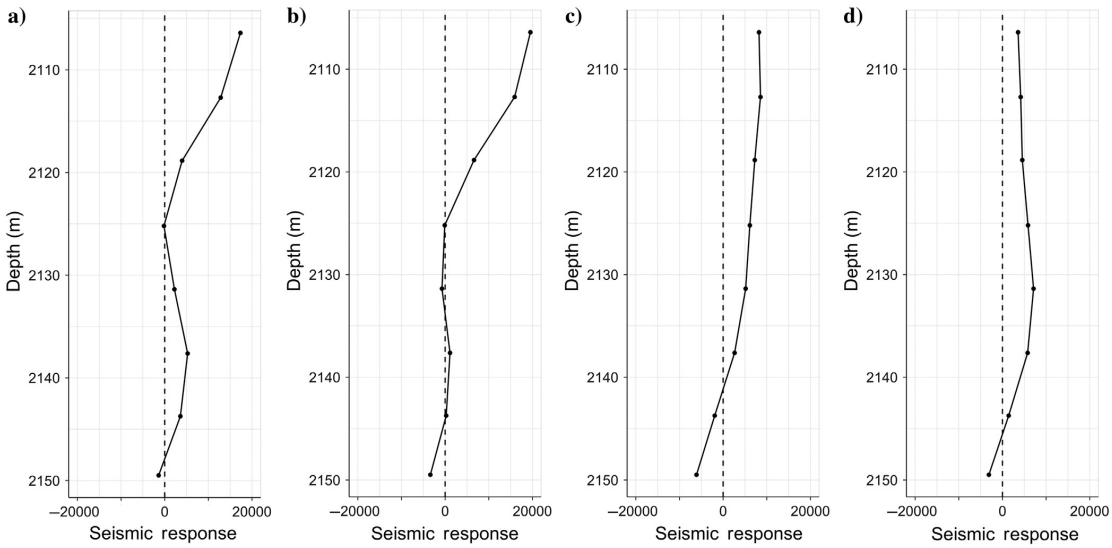

Figure 2. Seismic AVO data, in the form of picked amplitudes along the reservoir profile, collected in 1996 from Kneler in the Alvheim oil and gas field: (a) 10°, (b) 18°, (c) 27°, and (d) 35°.

$n_r = 58$ grid points, which is coarser than in Figure 1. Lastly, spatial histograms of the synthetic well observations located within the reservoir profile are shown at the bottom of the displays. Observe that the porosity histograms are unimodal and fairly symmetric, whereas both saturation histograms are strongly bimodal with slightly less water in case A than in case B.

Case A

Case study A corresponds to an all-sand reservoir profile. It is constructed by replacing the shale layers in the observations by consistently simulated sand sequences. This case represents the simplest reservoir design with one clearly defined OWC.

Case B

Case study B corresponds to a reservoir zone consisting of two separately producible compartments. It is constructed by keeping the dominant shale layer between 2122 and 2123 m depth, which acts as an impermeable barrier, and replacing the other layers by consistently simulated sand sequences. Furthermore, we introduce a thin water zone underlying the oil zone on top of the major shale layer. This case represents a more complex reservoir design with two OWCs and mixed lithology.

MODEL DESCRIPTION

We phrase the assessment of $\mathbf{r}$ given $\mathbf{d}$, $[\mathbf{r}|\mathbf{d}]$, in a Bayesian spatial inversion setting, which entails defining a prior model for the reservoir variables $p(\mathbf{r})$ and a likelihood model $p(\mathbf{d}|\mathbf{r})$ for the seismic observations. The solution to the inversion is termed the posterior model, and it is represented by the pdf $p(\mathbf{r}|\mathbf{d})$ obtained by an application of Bayes' rule,

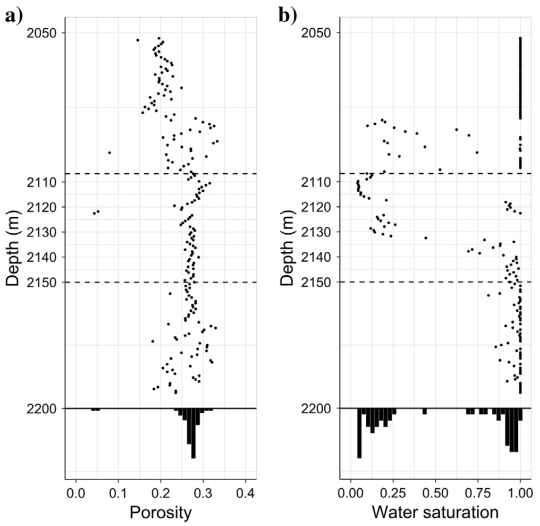

Figure 4. Synthetic well in case B: (a) porosity and (b) saturation profiles.

$$p(\mathbf{r}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{r})p(\mathbf{r})}{p(\mathbf{d})} = \text{const} \times p(\mathbf{d}|\mathbf{r})p(\mathbf{r}), \quad (1)$$

with const being a normalizing constant which is usually very computationally demanding to calculate.

Probability distributions

The class of Gaussian distributions is a natural candidate to represent variables that appear as unimodal and symmetric. An n-dimensional real-valued random vector $\mathbf{x}$ with expected value $\boldsymbol{\mu}_x$ and covariance matrix $\boldsymbol{\Sigma}_x$ is Gaussian if its pdf, $p(\mathbf{x})$, is of the form

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}_x|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_x)^T\boldsymbol{\Sigma}_x^{-1}(\mathbf{x}-\boldsymbol{\mu}_x)\right\}. \quad (2)$$

We denote this n-dimensional Gaussian pdf by $\varphi_n(\mathbf{x};\boldsymbol{\mu}_x,\boldsymbol{\Sigma}_x)$.

The probability that the n-dimensional Gaussian vector $\mathbf{x}$ belongs to the n-dimensional set $\mathbf{U}$ is given by

$$P(\mathbf{x} \in \mathbf{U};\boldsymbol{\mu}_x,\boldsymbol{\Sigma}_x) = \int_{\mathbb{R}^n} I(\mathbf{x} \in \mathbf{U})\varphi_n(\mathbf{x};\boldsymbol{\mu}_x,\boldsymbol{\Sigma}_x)\mathrm{d}\mathbf{x}. \quad (3)$$

We denote this n-dimensional Gaussian subset probability by $\boldsymbol{\Phi}_n(\mathbf{U};\boldsymbol{\mu}_x,\boldsymbol{\Sigma}_x)$.

In this study, we introduce the class of selection Gaussian distributions, which is very flexible, supporting multimodal, skewed, and truncated distribution types. A selection Gaussian random variable $\mathbf{x}$ is defined by two Gaussian random variables: the basis variable $\tilde{\mathbf{x}}$ and the auxiliary variable $\boldsymbol{\nu}$. These variables are jointly Gaussian, and the selection Gaussian random variable $\mathbf{x}$ is obtained by conditioning $\mathbf{x} = [\tilde{\mathbf{x}}|\boldsymbol{\nu} \in \mathbf{A}]$ with $\mathbf{A}$ a nonempty subset with the same dimension as $\boldsymbol{\nu}$. The distribution of the basis variable primarily controls the location and variability of the associated selection Gaussian distribution, whereas the correlation between $\tilde{\mathbf{x}}$ and $\boldsymbol{\nu}$ and the geometry of the selection set $\mathbf{A}$ primarily controls its shape. Constructing an n-dimensional selection Gaussian random variable starts with defining the n-dimensional basis variable $\tilde{\mathbf{x}}$. This basis variable is Gauss-linearly extended by an m-dimensional auxiliary variable $\boldsymbol{\nu}$; that is, $\boldsymbol{\nu}$ conditioned on $\tilde{\mathbf{x}}$ is Gaussian with expectation linear in $\tilde{\mathbf{x}}$. Consequently, the joint distribution of $\tilde{\mathbf{x}}$ and $\boldsymbol{\nu}$ is Gaussian,

$$p\left(\begin{bmatrix}\tilde{\mathbf{x}}\\ \boldsymbol{\nu}\end{bmatrix}\right) = \varphi_{n+m}\left(\begin{bmatrix}\tilde{\mathbf{x}}\\ \boldsymbol{\nu}\end{bmatrix}; \begin{bmatrix}\boldsymbol{\mu}_{\tilde{x}}\\ \boldsymbol{\mu}_{\nu}\end{bmatrix}, \begin{bmatrix}\boldsymbol{\Sigma}_{\tilde{x}} & \boldsymbol{\Gamma}_{\tilde{x}\nu}\\ \boldsymbol{\Gamma}_{\nu\tilde{x}} & \boldsymbol{\Sigma}_{\nu}\end{bmatrix}\right), \quad (4)$$

with $\boldsymbol{\mu}_{\tilde{x}}$, $\boldsymbol{\mu}_{\nu}$ and $\boldsymbol{\Sigma}_{\tilde{x}}$, $\boldsymbol{\Sigma}_{\nu}$ being the mean vectors and covariance matrices of $\tilde{\mathbf{x}}$ and $\boldsymbol{\nu}$, respectively. The covariance between $\tilde{\mathbf{x}}$ and $\boldsymbol{\nu}$ is defined by $\boldsymbol{\Gamma}_{\tilde{x}\nu} = \boldsymbol{\Gamma}_{\nu\tilde{x}}^T$. Note that the auxiliary variable $\boldsymbol{\nu}$ is introduced only for modeling and has no physical interpretation. Furthermore, define the nonempty selection set $\mathbf{A}$ with the same dimension as $\boldsymbol{\nu}$. Then, the conditional variable $\mathbf{x} = [\tilde{\mathbf{x}}|\boldsymbol{\nu} \in \mathbf{A}]$ belongs to the class of selection Gaussian distributions (Arellano-Valle et al., 2006):

$$p(\mathbf{x}) = p(\tilde{\mathbf{x}}|\boldsymbol{\nu} \in \mathbf{A}) = \frac{P(\boldsymbol{\nu} \in \mathbf{A}|\tilde{\mathbf{x}})p(\tilde{\mathbf{x}})}{P(\boldsymbol{\nu} \in \mathbf{A})} = \frac{\Phi_m(\mathbf{A}; \boldsymbol{\mu}_{\nu|\tilde{x}}, \boldsymbol{\Sigma}_{\nu|\tilde{x}})\varphi_n(\tilde{\mathbf{x}}; \boldsymbol{\mu}_{\tilde{x}}, \boldsymbol{\Sigma}_{\tilde{x}})}{\Phi_m(\mathbf{A}; \boldsymbol{\mu}_{\nu}, \boldsymbol{\Sigma}_{\nu})}, \quad (5)$$

where

$$\boldsymbol{\mu}_{\nu|\tilde{x}} = \boldsymbol{\mu}_{\nu} + \boldsymbol{\Gamma}_{\nu\tilde{x}}\boldsymbol{\Sigma}_{\tilde{x}}^{-1}(\tilde{\mathbf{x}} - \boldsymbol{\mu}_{\tilde{x}}),$$
$$\boldsymbol{\Sigma}_{\nu|\tilde{x}} = \boldsymbol{\Sigma}_{\nu} - \boldsymbol{\Gamma}_{\nu\tilde{x}}\boldsymbol{\Sigma}_{\tilde{x}}^{-1}\boldsymbol{\Gamma}_{\tilde{x}\nu}. \quad (6)$$

The flexibility of the selection Gaussian distributions enters through the correlation between $\tilde{\mathbf{x}}$ and $\boldsymbol{\nu}$, specified in $\boldsymbol{\Gamma}_{\tilde{x}\nu} = \boldsymbol{\Gamma}_{\nu\tilde{x}}^T$, and in the geometry of the selection set $\mathbf{A}$. Note that the class of Gaussian distributions is contained in the class of selection Gaussian distributions because the class of selection Gaussian distributions where $\tilde{\mathbf{x}}$ and $\boldsymbol{\nu}$ are uncorrelated or the selection set constitutes the entire m-dimensional space, coincides with the class of Gaussian distributions.

In Figure 5, the concept of the selection Gaussian distribution is illustrated for a 1D example. Figure 5a depicts the joint distribution of $\tilde{x}$ and ν together with a selection set $\mathbf{A}$, whereas Figure 5b shows the selection set with its corresponding selection Gaussian distribution embedded.

Later, we will extend the Gaussian and selection Gaussian distribution classes to represent spatial variables along the reservoir profile, the so-called GRF and S-GRF.

In a Bayesian framework (see equation 1), the class of Gaussian distributions provides conjugate prior models with respect to Gauss-linear likelihood models; that is, the resulting posterior model also belongs to the class of Gaussian distributions. It can be demonstrated (Omre and Rimstad, 2018) that the class of selection Gaussian distributions is also conjugate with respect to Gauss-linear likelihood models.

The class of high-dimensional selection Gaussian distributions is challenging to assess because of the high-dimensional Gaussian slice $\Phi_m(\mathbf{A}; \boldsymbol{\mu}_{\nu|\tilde{x}}, \boldsymbol{\Sigma}_{\nu|\tilde{x}})$ that must be computed. We approach this challenge by simulation, we use a block-independent MCMC algorithm to first simulate $\boldsymbol{\nu} \in \mathbf{A}$ (Omre and Rimstad, 2018), and then we simulate $\mathbf{x} = [\tilde{\mathbf{x}}|\boldsymbol{\nu}]$, which is Gaussian and therefore easy to simulate from to obtain selection Gaussian realizations. The simulation algorithm is presented in pseudocode in Appendix A.

Likelihood model

In our study, the likelihood model $p(\mathbf{d}|\mathbf{r})$, representing the seismic AVO acquisition procedure given the reservoir variables, is decomposed into

$$p(\mathbf{d}|\mathbf{r}) = \int p(\mathbf{d}|\mathbf{m}, \mathbf{r})p(\mathbf{m}|\mathbf{r})\mathrm{d}\mathbf{m} = \int p(\mathbf{d}|\mathbf{m})p(\mathbf{m}|\mathbf{r})\mathrm{d}\mathbf{m}. \quad (7)$$

The last equality holds because $\mathbf{m}$ are canonical variables of the wave equation, which is used to model the seismic wave propagation.

In the likelihood decomposition, $p(\mathbf{m}|\mathbf{r})$ is termed the rock-physics likelihood model and $p(\mathbf{d}|\mathbf{m})$ is termed the seismic likelihood model. Inspired by Landrø (2001), we use a linear rock-physics model, which, together with a linearized seismic likelihood model (Buland and Omre, 2003), form the likelihood model.

The rock-physics model is Gauss-linear and relates the reservoir variables to the elastic rock variables $[\mathbf{m}|\mathbf{r}] = \mathbf{B}\mathbf{r} + \mathbf{e}_{m|r}$. Here, $\mathbf{B}$ is a $(3n_r \times 2n_r)$-matrix containing the regression coefficients from a multiple linear regression model for the elastic rock variables with the reservoir variables as explanatory variables. The model error induced by assuming this linear relation is represented by the centered Gaussian error term $\mathbf{e}_{m|r}$, a $3n_r$-dimensional vector with the covariance matrix

$$\boldsymbol{\Sigma}_{m|r} = \begin{bmatrix} \sigma^2_{V_P|r}\mathbf{I}_{n_r} & \xi_{V_PV_S}\mathbf{I}_{n_r} & \xi_{V_P\rho}\mathbf{I}_{n_r} \\ \xi_{V_PV_S}\mathbf{I}_{n_r} & \sigma^2_{V_S|r}\mathbf{I}_{n_r} & \xi_{V_S\rho}\mathbf{I}_{n_r} \\ \xi_{V_P\rho}\mathbf{I}_{n_r} & \xi_{V_S\rho}\mathbf{I}_{n_r} & \sigma^2_{\rho|r}\mathbf{I}_{n_r} \end{bmatrix}, \quad (8)$$

where $\xi_{V_PV_S}\mathbf{I}_{n_r}$, $\xi_{V_P\rho}\mathbf{I}_{n_r}$, and $\xi_{V_S\rho}\mathbf{I}_{n_r}$ are the covariances between V_P and V_S, V_P and ρ, and V_S and ρ, respectively. Hence, the rock-physics likelihood model is

$$p(\mathbf{m}|\mathbf{r}) = \varphi_{3n_r}(\mathbf{m}; \mathbf{B}\mathbf{r}, \boldsymbol{\Sigma}_{m|r}). \quad (9)$$

Figure 6 displays the estimated regression models based on the observations in Figure 1 after outlier removal. Note that the bimodal characteristics of the saturation profile can be observed in the plots in the right column.

In Table 1, the numerical values of the estimated coefficients and of the standard deviations associated with the Gauss-linear rock-physics likelihood model are displayed. Each regression model is defined by the intercept α and the slopes β_ϕ and β_{s_w} of porosity and saturation, respectively. Furthermore, σ is the associated standard deviation. The covariance between V_P and V_S is estimated to be $\xi_{V_PV_S} = 6.16 \cdot 10^{-4}$, which corresponds to a correlation of 0.94. The estimated correlations between ρ and the velocities are very low; hence, we assume independence. We see that porosity is assigned more explanatory power than saturation, which is also reflected in Figure 6.

The seismic likelihood model is Gauss-linear and of the form $[\mathbf{d}|\mathbf{m}] = \mathbf{W}\mathbf{A}\mathbf{D}\mathbf{m} + \mathbf{e}_{d|m}$ (Buland and Omre, 2003). The matrices $\mathbf{A}$ and $\mathbf{D}$ emerge from the linearization of the Aki and Richards approximation of the Zoeppritz equations, with $\mathbf{D}$ a $(3n_r \times 3n_r)$-matrix functioning as a contrast operator with respect to the profile

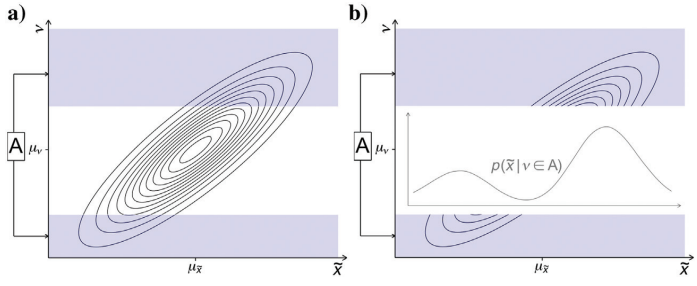

Figure 5. Selection Gaussian distribution.

direction and $\mathbf{A}$ being a $(n_\theta n_r \times 3n_r)$-matrix containing the Aki and Richards angle-dependent coefficients. The $(n_\theta n_d \times n_\theta n_r)$-matrix $\mathbf{W}$ contains discretizations of a wavelet and functions as a convolution operator. Lastly, $\mathbf{e}_{d|m}$ is an $n_\theta n_d$-dimensional vector containing centered Gaussian error terms encompassing model error and observation error, with the covariance matrix $\mathbf{\Sigma}_{d|m} = \mathbf{W}\sigma^2_{c|m}\mathbf{I}_{n_\theta n_r}\mathbf{W}^T + \sigma^2_{d|c}\mathbf{I}_{n_\theta n_d}$. Note that the first term indicates that the error is colored, due to convolution. Furthermore, $\sigma^2_{c|m}$ is the variance associated with the model error resulting from the linearization of the Zoeppritz equations and $\sigma^2_{d|c}$ is the variance associated with observation error. Hence, the seismic likelihood model is

$$p(\mathbf{d}|\mathbf{m}) = \varphi_{n_\theta n_d}(\mathbf{d}; \mathbf{WADm}, \mathbf{\Sigma}_{d|m}). \tag{10}$$

Figure 7 shows the estimated wavelet that defines the convolution matrix $\mathbf{W}$. The estimation is performed in a wide depth interval of 300 m of the well containing the reservoir profile. Several estimation methods are tested: parametric and nonparametric estimation methods, including the conventional Roy White method (White and Simm, 2003), with and without angular wavelet dependency. The best seismic fit and inversion results are obtained by using a cubic spline smoothed version of the wavelet estimated by regressing the seismic data on the reflection coefficients obtained by applying parts of the seismic likelihood model on the elastic rock variables along the well profile. The estimation is performed simultaneously for all angles; that is, the resulting wavelet is the one wavelet that jointly fits all angle stacks best. Note that the estimated wavelet has two sidelobes and appears slightly skewed. Moreover, because the estimation interval is more than six times wider than the study interval, we expect no overfitting in the estimation.

Combining the rock physics and the seismic likelihood models, the likelihood model takes the form $[\mathbf{d}|\mathbf{r}] = \mathbf{Gr} + \mathbf{e}_{d|r}$, with $\mathbf{G} = \mathbf{WADB}$ the forward operator, an $(n_\theta n_d \times 2n_r)$-matrix, and $\mathbf{e}_{d|r} = \mathbf{WADe}_{m|r} + \mathbf{e}_{d|m}$ an $n_\theta n_d$-dimensional vector, containing the total error, which is centered and Gaussian. The corresponding covariance matrix is $\mathbf{\Sigma}_{d|r} = \mathbf{WAD\Sigma}_{m|r}(\mathbf{WAD})^T + \mathbf{\Sigma}_{d|m}$. Hence, the likelihood model is Gauss-linear,

$$p(\mathbf{d}|\mathbf{r}) = \varphi_{n_\theta n_d}(\mathbf{d}; \mathbf{Gr}, \mathbf{\Sigma}_{d|r}). \tag{11}$$

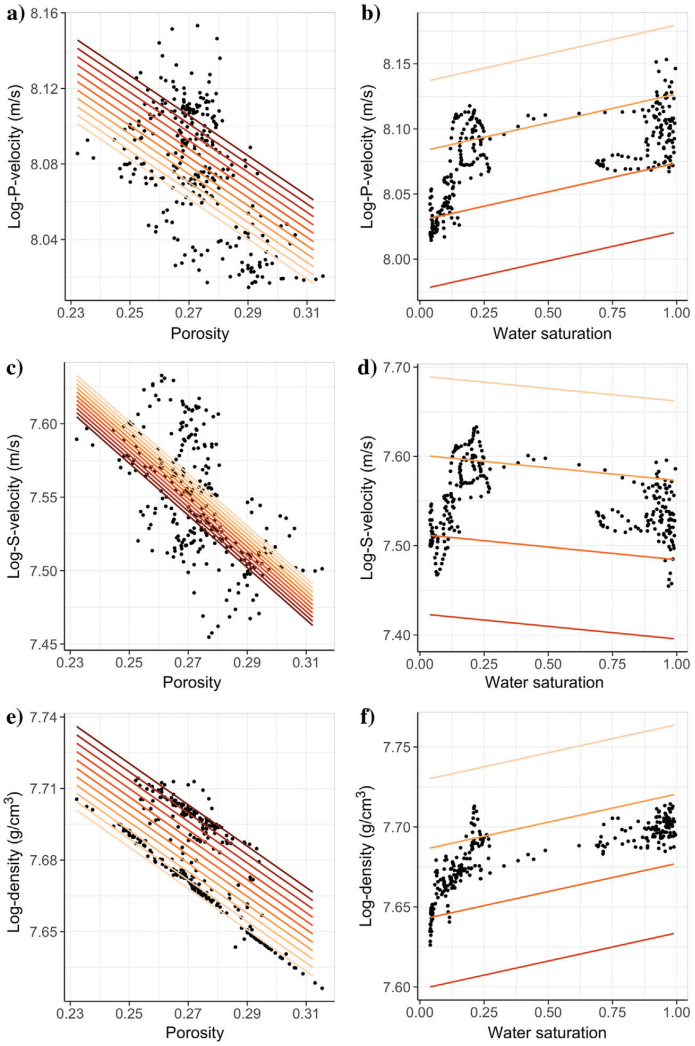

Figure 6. Rock-physics model. Estimated relationship between porosity and (a) P-wave velocity, (c) S-wave velocity, and (e) density, with water saturation fixed to a value in $[0, 1]$ with step size 0.1. Estimated relationship between water saturation and (b) P-wave velocity, (d) S-wave velocity, and (f) density, with porosity fixed to a value in $[0.2, 0.35]$ with step size 0.05. The color of the lines gradually transitions from yellow to red, corresponding to the minimum and maximum values, respectively. The observed elastic rock properties are indicated by points.

Synthetic case studies

Figures 8 and 9 display synthetic seismic AVO data generated from the reservoir variables in cases A and B, respectively, using the likelihood model $p(\mathbf{d}|\mathbf{r})$; see equation 11.

The synthetic seismic data are generated with a signal-to-noise ratio (S/N) of 5. The S/N governs the magnitude of the total error, and we assume that it consists of 1% white noise. The remaining 99% of the total error is wavelet-colored errors, resulting from propagation and convolution of model errors from the rock physics and reflection models.

The appearance of the synthetic seismic AVO data in Figures 8 and 9 should be compared with the corresponding observed data in Figure 2. Because the synthetic seismic data are based on the reservoir variables associated with cases A and B, the profiles should be similar but not identical

to the observed data in Figure 2. The data profiles have many similarities, although some deviations occur at the boundary depths.

Prior model

The prior model for the reservoir variables $p(\mathbf{r})$ should capture the characteristics of the observed porosity and saturation profiles displayed in Figure 1. The following decomposition of the prior distribution is useful for modeling:

$$p(\mathbf{r}) = p(\boldsymbol{\phi}, \mathbf{s}_w) = p(\boldsymbol{\phi}|\mathbf{s}_w)p(\mathbf{s}_w), \tag{12}$$

because it enables step-wise model construction. We start by defining a prior model for saturation. The saturation profile in sand in Figure 1b appears as spatially smooth with the exception of an abrupt fluid transition at the OWC, and the spatial histogram is bimodal. A traditional GRF model cannot capture bimodal marginal characteristics; hence, this class of models is unsuitable as a prior model for saturation. However, a generalization of the GRF, termed S-GRF, may capture bimodality (Omre and Rimstad, 2018). This spatial RF model is based on the selection Gaussian model presented in the "Probability distributions" subsection. It is defined by a basis GRF model, an auxiliary GRF, and a selection set. The basis GRF model for saturation is

$$p(\tilde{\mathbf{s}}_w) = \varphi_{n_r}(\tilde{\mathbf{s}}_w; \mu_{\tilde{s}_w}\mathbf{i}_{n_r}, \sigma^2_{\tilde{s}_w}\boldsymbol{\Omega}), \tag{13}$$

with expected level $\mu_{\tilde{s}_w}$, variance level $\sigma^2_{\tilde{s}_w}$, and $(n_r \times n_r)$ correlation matrix $\boldsymbol{\Omega}$. The correlation matrix is defined by the powered exponential spatial correlation function $\rho(\tau;\alpha_r,\beta_r) = \exp\{-(\tau/\alpha_r)^{\beta_r}\}$ with $\tau > 0$ being the grid point interdistances, whereas $\alpha_r > 0$ and $0 \le \beta_r \le 2$ are the scale and shape parameters, respectively.

The n_r-dimensional auxiliary variable $\boldsymbol{\nu}$ is Gauss-linearly and conditionally independently related to the basis GRF,

$$\begin{aligned} p(\boldsymbol{\nu}|\tilde{\mathbf{s}}_w) &= \varphi_{n_r}(\boldsymbol{\nu}; \boldsymbol{\mu}_{\nu|\tilde{s}_w}, (1-\gamma^2_{\nu\tilde{s}_w})\mathbf{I}_{n_r}) \\ &= \prod_{i=1}^{n_r} \varphi_1(\nu_i; \mu_{\nu_i|\tilde{s}_w}, (1-\gamma^2_{\nu\tilde{s}_w})), \end{aligned} \tag{14}$$

where

$$\boldsymbol{\mu}_{\nu|\tilde{s}_w} = 0\mathbf{i}_{n_r} + \boldsymbol{\Gamma}_{\nu\tilde{s}_w}[\sigma^2_{\tilde{s}_w}\boldsymbol{\Omega}]^{-1}(\tilde{\mathbf{s}}_w - \mu_{\tilde{s}_w}\mathbf{i}_{n_r}), \tag{15}$$

with $\boldsymbol{\Gamma}_{\nu\tilde{s}_w} = \gamma_{\nu\tilde{s}_w}\sigma_{\tilde{s}_w}\boldsymbol{\Omega}$ being the covariance between $\boldsymbol{\nu}$ and $\tilde{\mathbf{s}}_w$, where $\gamma_{\nu\tilde{s}_w}$ is the correlation between ν and $\tilde{s}_w$. The n_r-dimensional selection set $\mathbf{A}$ associated with the auxiliary variable is specified such that $\mathbf{A} = A^{n_r}$, with $A = (-\infty, a] \cup [b, \infty)$ being a subset of the real numbers. The prior model for saturation is the discretized S-GRF with pdf

$$\begin{aligned} p(\mathbf{s}_w) &= p(\tilde{\mathbf{s}}_w|\boldsymbol{\nu} \in \mathbf{A}) \\ &= \frac{\Phi_{n_r}(\mathbf{A}; \boldsymbol{\mu}_{\nu|\tilde{s}_w}, (1-\gamma^2_{\nu\tilde{s}_w})\mathbf{I}_{n_r})}{\Phi_{n_r}(\mathbf{A}; 0\mathbf{i}_{n_r}, \gamma^2_{\nu\tilde{s}_w}\boldsymbol{\Omega} + (1-\gamma^2_{\nu\tilde{s}_w})\mathbf{I}_{n_r})}\varphi_{n_r}(\tilde{\mathbf{s}}_w; \mu_{\tilde{s}_w}\mathbf{i}_{n_r}, \sigma^2_{\tilde{s}_w}\Omega) \\ &= \frac{\prod_{i=1}^{n_r}\Phi_1(A; \mu_{\nu_i|\tilde{s}_w}, (1-\gamma^2_{\nu\tilde{s}_w}))}{\Phi_{n_r}(\mathbf{A}; 0\mathbf{i}_{n_r}, \gamma^2_{\nu\tilde{s}_w}\boldsymbol{\Omega} + (1-\gamma^2_{\nu\tilde{s}_w})\mathbf{I}_{n_r})}\varphi_{n_r}(\tilde{\mathbf{s}}_w; \mu_{\tilde{s}_w}\mathbf{i}_{n_r}, \sigma^2_{\tilde{s}_w}\boldsymbol{\Omega}). \end{aligned} \tag{16}$$

This prior model will be spatially stationary in the sense that all location-wise distributions are identical.

The porosity profile in sand in Figure 1a appears as spatially smoothly varying, and the spatial histogram appears unimodal and symmetric. A stationary GRF model therefore seems suitable,

$$p(\boldsymbol{\phi}) = \varphi_{n_r}(\boldsymbol{\phi}; \mu_\phi\mathbf{i}_{n_r}, \sigma^2_\phi\boldsymbol{\Omega}), \tag{17}$$

with expected level μ_ϕ, variance level σ^2_ϕ, and $(n_r \times n_r)$ correlation matrix $\boldsymbol{\Omega}$. The joint pdf $p(\boldsymbol{\phi}, \tilde{\mathbf{s}}_w, \boldsymbol{\nu})$ can be decomposed into

Table 1. Estimated parameters in the rock-physics likelihood model. The estimated intercept $\hat{\alpha}$, slope of porosity $\hat{\beta}_\phi$, and slope of water saturation $\hat{\beta}_{s_w}$ are displayed together with the estimated standard deviation $\hat{\sigma}$ for each regression model in the top table. Each row of the table contains the parameters associated with the model for the elastic rock property specified in the leftmost column. The estimated cross-covariances between the elastic rock properties, $\hat{\xi}_{V_P V_S}$, $\hat{\xi}_{V_P\rho}$, and $\hat{\xi}_{V_S\rho}$, are shown in the lower table.

	$\hat{\alpha}$	$\hat{\beta}_\phi$	$\hat{\beta}_{s_w}$	$\hat{\sigma}$
V_P	8.348	−1.060	0.044	0.022
V_S	8.046	−1.777	−0.029	0.030
ρ	7.902	−0.868	0.035	0.006

$\hat{\xi}_{V_P V_S}$	$\hat{\xi}_{V_P\rho}$	$\hat{\xi}_{V_S\rho}$
$6.16 \cdot 10^{-4}$	0	0

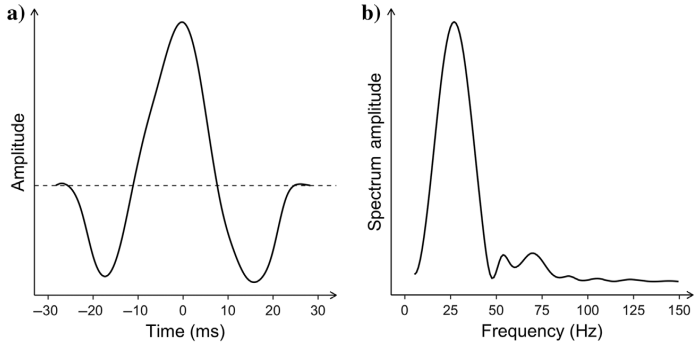

Figure 7. Estimated wavelet.

$$p(\boldsymbol{\phi}, \tilde{\mathbf{s}}_w, \boldsymbol{\nu}) = p(\boldsymbol{\nu}|\boldsymbol{\phi}, \tilde{\mathbf{s}}_w)p(\boldsymbol{\phi}|\tilde{\mathbf{s}}_w)p(\tilde{\mathbf{s}}_w) = p(\boldsymbol{\phi}|\tilde{\mathbf{s}}_w)p(\boldsymbol{\nu}|\tilde{\mathbf{s}}_w)p(\tilde{\mathbf{s}}_w), \quad (18)$$

where the latter equality follows because $[\boldsymbol{\nu}|\tilde{\mathbf{s}}_w]$ is conditionally independent of porosity. The porosity is defined to be Gauss-linearly dependent on saturation,

$$p(\boldsymbol{\phi}|\tilde{\mathbf{s}}_w) = \varphi_{n_r}(\boldsymbol{\phi}; \boldsymbol{\mu}_{\phi|\tilde{s}_w}, \boldsymbol{\Sigma}_{\phi|\tilde{s}_w}). \quad (19)$$

Here,

$$\boldsymbol{\mu}_{\phi|\tilde{\mathbf{s}}_w} = \mu_\phi \mathbf{i}_{n_r} + \boldsymbol{\Gamma}_{\phi\tilde{s}_w}\boldsymbol{\Sigma}_{\tilde{s}_w}^{-1}(\tilde{\mathbf{s}}_w - \mu_{\tilde{s}_w}\mathbf{i}_{n_r}),$$
$$\boldsymbol{\Sigma}_{\phi|\tilde{s}_w} = \sigma_\phi^2\boldsymbol{\Omega} - \boldsymbol{\Gamma}_{\phi\tilde{s}_w}\boldsymbol{\Sigma}_{\tilde{s}_w}^{-1}\boldsymbol{\Gamma}_{\tilde{s}_w\phi}, \quad (20)$$

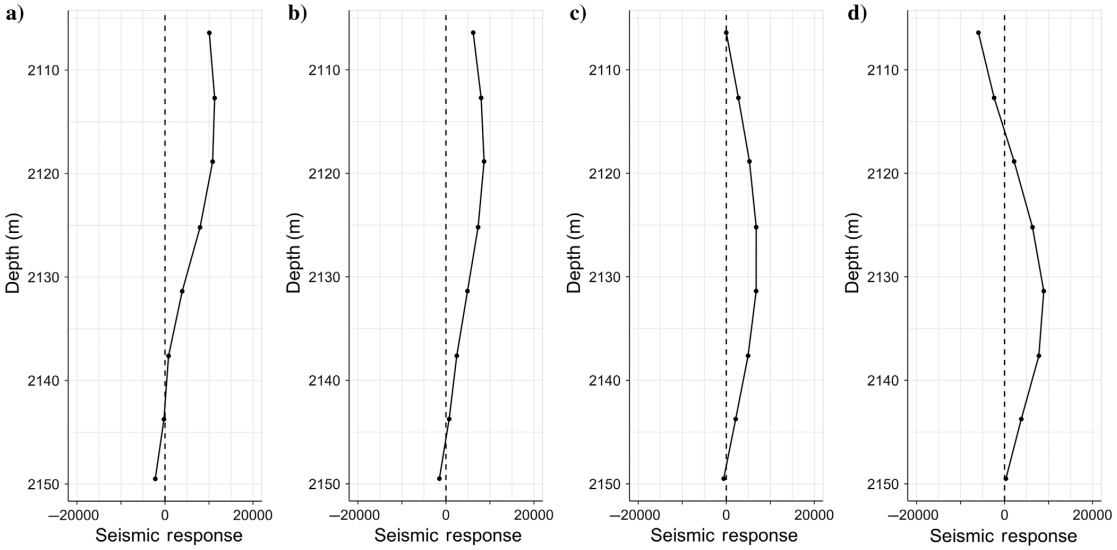

Figure 8. Synthetic seismic AVO data associated with case study A, in the form of picked amplitudes along the reservoir profile: (a) 10°, (b) 18°, (c) 27°, and (d) 35°.

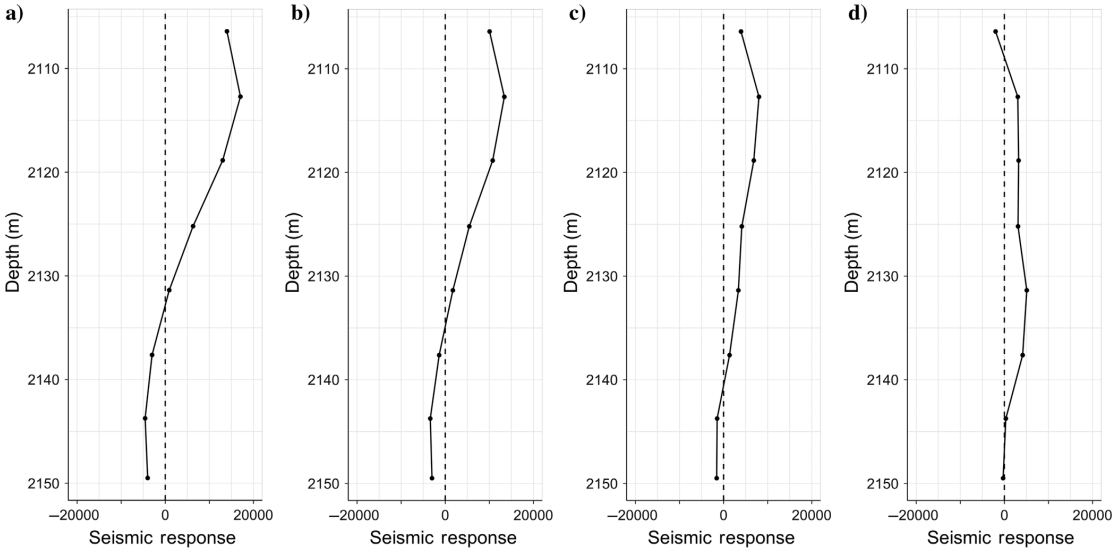

Figure 9. Synthetic seismic AVO data associated with case study B, in the form of picked amplitudes along the reservoir profile: (a) 10°, (b) 18°, (c) 27°, and (d) 35°.

with $\mathbf{\Gamma}_{\phi\tilde{s}_w} = \mathbf{\Gamma}_{\tilde{s}_w\phi} = \rho_{\phi\tilde{s}_w}\sigma_\phi\sigma_{\tilde{s}_w}\mathbf{\Omega}$ being the covariance between $\boldsymbol{\phi}$ and $\tilde{\mathbf{s}}_w$, where $\rho_{\phi\tilde{s}_w}$ is the correlation between ϕ and $\tilde{s}_w$.

The stationary GRF $\tilde{\mathbf{r}} = [\boldsymbol{\phi}, \tilde{\mathbf{s}}_w]^T$ and the stationary auxiliary GRF $\boldsymbol{\nu}$ are correlated only through the correlation between $\tilde{s}_w$ and ν, represented by $\gamma_{\nu\tilde{s}_w}$. However, note that the auxiliary variable $\boldsymbol{\nu}$ affects porosity through $\tilde{\mathbf{s}}_w$ if $|\rho_{\phi\tilde{s}_w}| > 0$. The joint pdf of the GRFs is

$$p\left(\begin{bmatrix}\boldsymbol{\phi}\\ \tilde{\mathbf{s}}_w\\ \boldsymbol{\nu}\end{bmatrix}\right) = \varphi_{3n_r}\left(\begin{bmatrix}\boldsymbol{\phi}\\ \tilde{\mathbf{s}}_w\\ \boldsymbol{\nu}\end{bmatrix}; \begin{bmatrix}\mu_\phi \mathbf{i}_{n_r}\\ \mu_{\tilde{s}_w}\mathbf{i}_{n_r}\\ 0\mathbf{i}_{n_r}\end{bmatrix}, \begin{bmatrix}\sigma_\phi^2\mathbf{\Omega} & \mathbf{\Gamma}_{\phi\tilde{s}_w} & \mathbf{\Gamma}_{\phi\nu}\\ \mathbf{\Gamma}_{\tilde{s}_w\phi} & \sigma_{\tilde{s}_w}^2\mathbf{\Omega} & \mathbf{\Gamma}_{\nu\tilde{s}_w}\\ \mathbf{\Gamma}_{\nu\phi} & \mathbf{\Gamma}_{\nu\tilde{s}_w} & \mathbf{\Sigma}_\nu\end{bmatrix}\right), \tag{21}$$

where

$$\begin{aligned}\mathbf{\Gamma}_{\phi\nu} &= \mathbf{\Gamma}_{\nu\phi} = \rho_{\phi\tilde{s}_w}\gamma_{\nu\tilde{s}_w}\sigma_\phi\mathbf{\Omega},\\ \mathbf{\Sigma}_\nu &= \gamma_{\nu\tilde{s}_w}^2\mathbf{\Omega} + (1-\gamma_{\nu\tilde{s}_w}^2)\mathbf{I}_{n_r}.\end{aligned} \tag{22}$$

It can be shown that the covariance $\mathbf{\Gamma}_{\phi\nu|\tilde{s}_w} = 0\mathbf{I}_{n_r}$, that is, $p(\boldsymbol{\phi}|\mathbf{s}_w) = p(\boldsymbol{\phi}|\tilde{\mathbf{s}}_w)$. Consequently, the conditional pdf defining the joint prior model is

$$p(\boldsymbol{\phi}|\mathbf{s}_w) = p(\boldsymbol{\phi}|\tilde{\mathbf{s}}_w) = \varphi_{n_r}(\boldsymbol{\phi}; \boldsymbol{\mu}_{\phi|\tilde{s}_w}, \mathbf{\Sigma}_{\phi|\tilde{s}_w}), \tag{23}$$

and the joint prior model is a stationary bivariate discretized S-GRF,

$$\begin{aligned}p(\mathbf{r}) &= p(\boldsymbol{\phi}|\mathbf{s}_w)p(\mathbf{s}_w) = p(\boldsymbol{\phi}|\tilde{\mathbf{s}}_w)p(\tilde{\mathbf{s}}_w|\boldsymbol{\nu}\in\mathbf{A})\\ &= \varphi_{n_r}(\boldsymbol{\phi}; \boldsymbol{\mu}_{\phi|\tilde{s}_w}, \mathbf{\Sigma}_{\phi|\tilde{s}_w})\frac{\prod_{i=1}^{n_r}\Phi_1(A; \mu_{\nu_i|\tilde{s}_{w,i}}, (1-\gamma_{\nu\tilde{s}_w}^2))}{\Phi_{n_r}(\mathbf{A}; 0\mathbf{i}_{n_r}, \gamma_{\nu\tilde{s}_w}^2\mathbf{\Omega} + (1-\gamma_{\nu\tilde{s}_w}^2)\mathbf{I}_{n_r})}\\ &\times\varphi_{n_r}(\tilde{\mathbf{s}}_w; \mu_{\tilde{s}_w}\mathbf{i}_{n_r}, \sigma_{\tilde{s}_w}^2\mathbf{\Omega}),\end{aligned} \tag{24}$$

parameterized by $\Theta_p^{S-G} = [\mu_\phi, \mu_{\tilde{s}_w}, \sigma_\phi^2, \sigma_{\tilde{s}_w}^2, \rho_{\phi\tilde{s}_w}, \gamma_{\nu\tilde{s}_w}, A, \rho(\tau;\alpha_r,\beta_r)]$. The seven first parameters are primarily related to the bivariate location-wise selection Gaussian pdf for $r(z) = [\phi(z), s_w(z)]^T$. The last parameter is primarily related to the spatial coupling along the depth profile.

In traditional inversion of seismic AVO data for reservoir variables (Buland and Omre, 2003), a GRF prior model is used; hence, we define a GRF model as a benchmark for the S-GRF model suggested in the current study. We assign a discretized stationary GRF prior model,

$$p(\mathbf{r}) = \varphi_{2n_r}\left(\begin{bmatrix}\boldsymbol{\phi}\\ \mathbf{s}_w\end{bmatrix}; \begin{bmatrix}\mu_\phi\mathbf{i}_{n_r}\\ \mu_{s_w}\mathbf{i}_{n_r}\end{bmatrix}, \begin{bmatrix}\sigma_\phi^2\mathbf{\Omega} & \mathbf{\Gamma}_{\phi s_w}\\ \mathbf{\Gamma}_{s_w\phi} & \sigma_{s_w}^2\mathbf{\Omega}\end{bmatrix}\right), \tag{25}$$

with $\mathbf{\Gamma}_{\phi s_w} = \mathbf{\Gamma}_{s_w\phi} = \rho_{\phi s_w}\sigma_\phi\sigma_{s_w}\mathbf{\Omega}$, and parameterized by $\Theta_p^G = [\mu_\phi, \mu_{s_w}, \sigma_\phi^2, \sigma_{\tilde{s}_w}^2, \rho_{\phi s_w}, \rho(\tau;\alpha_r,\beta_r)]$.

Note that porosity and saturation are defined in the range $[0, 1]$, whereas the location-wise prior models are defined in $[-\infty, \infty]$. The possible inconsistencies due to this mismatch will depend on the actual values of the model parameters, and we will discuss this subject subsequently.

All model parameters of the prior stationary S-GRF Θ_p^{S-G} can formally be estimated by maximum likelihood (Omre and Rimstad, 2018). Given well observations $\mathbf{r}_{\text{obs}}$, the maximum likelihood estimate is $\hat{\Theta}_p^{S-G} = \underset{\Theta_p^{S-G}}{\text{argmax}}\{p(\mathbf{r}_{\text{obs}}; \Theta_p^{S-G})\}$, with $p(\mathbf{r}_{\text{obs}}; \Theta_p^{S-G})$ an n_r-dimensional selection Gaussian distribution. Alternatively, maximum likelihood estimation can be based on the seismic data; because the likelihood model is Gauss-linear, the seismic data have a selection Gaussian distribution with parameters linearly related to the parameters of the prior model. However, the maximum likelihood estimation is challenging due to the high degree of multimodality in the likelihood function.

We circumvent the difficulties posed by the maximum likelihood estimation by using a step-wise heuristic approach. We initiate the spatial correlation function to be $\rho(\tau;\alpha_r,\beta_r) = I(\tau = 0)$; hence, there is no spatial dependence. Because the prior model is stationary, the histograms and scatterplot of $[\phi, s_w]$ provided by the well logs are used to infer all parameters except the spatial correlation — this is the first step. The expectation and variance of the location-wise Gaussian prior model for porosity are estimated by maximum likelihood from the well logs. For saturation, we perform parametric optimization by gridding the parameters and minimizing the mean squared error (MSE) between the selection Gaussian distribution and the spatial histogram of saturation. The resulting approximate optimal parameters give a very good starting point, but some small parameter adjustments may be necessary as spatial correlation is introduced. In the second step, we find a suitable spatial correlation by performing and evaluating the smoothness of synthetic inversions, with the parameters estimated from the first step fixed and varying spatial correlation parameters. Lastly, we reinspect the synthetic inversion subject to the parameters estimated from the first and second steps and determine whether to adjust the parameter estimates from the first step or not. This last step requires some experience and understanding of the parameters.

The model parameters Θ_p^{S-G} of the prior stationary S-GRF, are estimated from the outlier corrected observations along the reservoir profile in Figure 1, and they are listed in Table 2. In the top row, the model parameters associated with the bivariate location-wise GRF basis model are displayed. The parameters associated with the auxiliary GRF are shown in the middle row, and the parameters associated with the spatial correlation structure are listed in the bottom row. Because the prior model is relatively low-parameterized, overfitting is unlikely. Note that the estimated intervariable correlation is low; hence, the auxiliary variable only weakly affects porosity and

Table 2. Parameter values used in the prior S-GRF. The parameters associated with the basis GRF are shown at the top: the expected levels ($\mu_\phi, \mu_{\tilde{s}_w}$), variance levels ($\sigma_\phi^2, \sigma_{\tilde{s}_w}^2$), and intervariable correlation $\rho_{\phi\tilde{s}_w}$. The parameters associated with the auxiliary GRF are shown in the middle: the correlation with the basis GRF $\gamma_{\nu\tilde{s}_w}$ and the marginal selection set A. The parameters associated with the spatial correlation structure are listed at the bottom: the scale parameter α_r and shape parameter β_r.

μ_ϕ	$\mu_{\tilde{s}_w}$	σ_ϕ^2	$\sigma_{\tilde{s}_w}^2$	$\rho_{\phi\tilde{s}_w}$
0.27	0.5	0.00019	0.04375	−0.164
$\gamma_{\nu\tilde{s}_w}$	A			
0.9	$[-\infty, -1.1] \cup [1.1, \infty]$			
α_r	β_r			
1.1	1.8			

the univariate location-wise prior models are unimodal for $\phi(z)$ and bimodal for $s_w(z)$.

Figure 10 contains four realizations from the stationary S-GRF prior model, and these realizations should be representative of the data shown in Figure 1, except for in the shale layers. Spatial histograms of the four realizations are displayed at the bottom of the displays. We observe that the porosity varies smoothly spatially and has a unimodal spatial histogram, whereas the saturation appears with small spatial variability except for some identifiable steps and has a bimodal spatial histogram. Note that the abrupt mode transitions occur randomly because the prior realizations are not conditioned on seismic AVO data. Moreover, observe that the saturation values sometimes are slightly outside the physical range $[0, 1]$.

The model parameters of the benchmarking stationary GRF prior model are also estimated from the outlier corrected observations along the reservoir profile in Figure 1, and they are presented in Table 3. Parameter estimation of the location-wise expectation and variance is done by maximum likelihood, whereas the correlations are assigned the same values as in the S-GRF prior model. The parameters associated with the location-wise bivariate GRF model are displayed in the top row, and the parameters associated with the spatial correlation structure are listed in the bottom row.

Figure 11 contains four realizations from the stationary GRF prior model, which should be representative of the data shown in Figure 1, except for in the shale layers. Note that the porosity profiles appear as smoothly spatially varying, whereas the saturation profiles appear with large variability and without identifiable steps. The spatial histograms in the lower part of the displays are unimodal, and very wide for the saturation, which sometimes falls outside the physical range $[0, 1]$ by a substantial amount.

For the GRF and S-GRF prior models, the actual parameters $(\mu_\phi, \sigma^2_\phi)$ inferred from the observations are such that the probability for ϕ to be within $[0, 1]$ is close to 1. The impact of the prior model being defined in $[-\infty, \infty]$ is therefore not reflected in the inversion results of porosity. However, the probability for s_w to be outside $[0, 1]$ is not negligible, particularly in the GRF prior model. This may impact the inversion results for saturation, and we describe our approach for correction next.

Posterior model

In Bayesian inversion, the posterior model is defined by the likelihood and prior models; see equation 1. Because the likelihood model $p(\mathbf{d}|\mathbf{r})$ is Gauss-linear and the prior model $p(\mathbf{r})$ is an S-GRF, the posterior model $p(\mathbf{r}|\mathbf{d})$ is also an S-GRF (Omre and Rimstad, 2018),

$$\begin{aligned} p(\mathbf{r}|\mathbf{d}) &= \text{const} \times p(\mathbf{d}|\mathbf{r})p(\mathbf{r}) \\ &= \text{const} \times p(\mathbf{d}|\boldsymbol{\phi}, \mathbf{s}_w)p(\boldsymbol{\phi}, \mathbf{s}_w) \\ &= \text{const} \times p(\tilde{\mathbf{d}}|\boldsymbol{\phi}, \tilde{\mathbf{s}}_w)p(\boldsymbol{\phi}, \tilde{\mathbf{s}}_w|\boldsymbol{\nu} \in \mathbf{A}). \end{aligned} \quad (26)$$

The latter equality follows from the likelihood function $p(\mathbf{d}|\mathbf{r})$ being a function of $\mathbf{r}$ with $\mathbf{d}$ fixed to the actual observations. More-

Table 3. Parameter values used in the prior GRF. The expected levels (μ_ϕ, μ_{s_w}), variance levels $(\sigma^2_\phi, \sigma^2_{s_w})$, and intervariable correlation $\rho_{\phi s_w}$ are shown at the top. The parameters associated with the spatial correlation structure are listed at the bottom: the scale parameter α_r and shape parameter β_r.

μ_ϕ	μ_{s_w}	σ^2_ϕ	$\sigma^2_{s_w}$	$\rho_{\phi s_w}$
0.27	0.45	0.00019	0.14583	−0.164
α_r	β_r			
1.1	1.8			

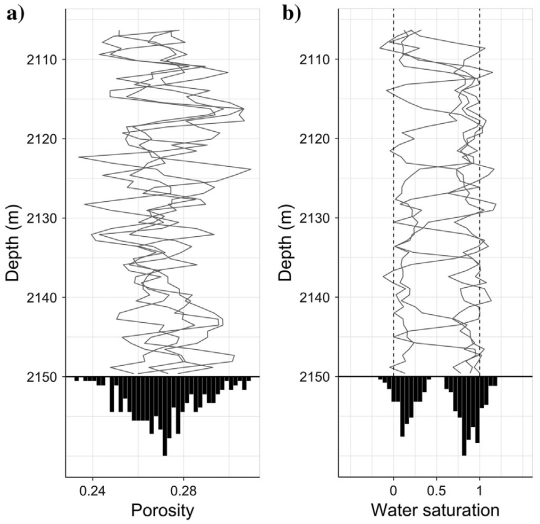

Figure 10. Four realizations from the S-GRF prior model: (a) porosity and (b) saturation profiles.

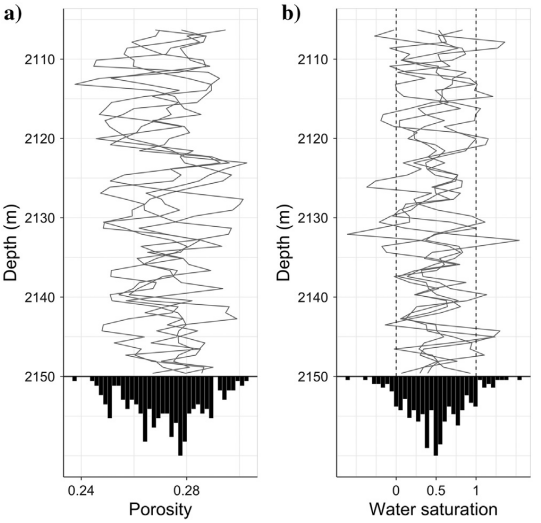

Figure 11. Four realizations from the GRF prior model: (a) porosity and (b) saturation profiles.

over, the likelihood function is invariant to the specification of prior model $p(\mathbf{r})$. From the relation above, one may demonstrate that $p(\tilde{\mathbf{d}}, \boldsymbol{\phi}, \tilde{\mathbf{s}}_w, \boldsymbol{\nu})$ is Gaussian and, hence, also that $p(\boldsymbol{\phi}, \tilde{\mathbf{s}}_w, \boldsymbol{\nu}|\mathbf{d})$ is Gaussian with $\tilde{\mathbf{d}} = \mathbf{d}$ being the actual observations. This entails that the posterior model is an S-GRF with exact pdf

$$p(\mathbf{r}|\mathbf{d}) = p(\boldsymbol{\phi}, \mathbf{s}_w|\mathbf{d}) = p(\boldsymbol{\phi}|\tilde{\mathbf{s}}_w, \mathbf{d}) p(\tilde{\mathbf{s}}_w|\mathbf{d}, \boldsymbol{\nu} \in \mathbf{A})$$
$$= \varphi_{n_r}(\boldsymbol{\phi}; \boldsymbol{\mu}_{\phi|\tilde{s}_w,d}, \boldsymbol{\Sigma}_{\phi|\tilde{s}_w,d}) \frac{\Phi_{n_r}(\mathbf{A}; \boldsymbol{\mu}_{\nu|\tilde{s}_w,d}, \boldsymbol{\Sigma}_{\nu|\tilde{s}_w,d})}{\Phi_{n_r}(\mathbf{A}; \boldsymbol{\mu}_{\nu|d}, \boldsymbol{\Sigma}_{\nu|d})}$$
$$\times \varphi_{n_r}(\tilde{\mathbf{s}}_w; \boldsymbol{\mu}_{\tilde{s}_w|d}, \boldsymbol{\Sigma}_{\tilde{s}_w|d}). \tag{27}$$

Expressions for these parameters are developed in Appendix B. The posterior distribution is assessed by simulation using the MCMC algorithm described in Appendix A. Based on the posterior distribution, we use the marginal (location-wise) maximum posterior (MMAP) as the predictor and the highest posterior density (HPD) interval (HPDI) as the prediction interval (Casella and Berger, 2001).

The MMAP can be expressed as

$$\mathbf{r}^* = \{\mathrm{MAP}(r_j^i|\mathbf{d}) = \underset{r_j^i}{\mathrm{argmax}}\{p(r_j^i|\mathbf{d})\},\ i = 1,2;\ j = 1,\ldots,n_r\}, \tag{28}$$

with $(\mathbf{r}^1, \mathbf{r}^2) = (\boldsymbol{\phi}, \mathbf{s}_w)$. The $(1-\alpha)\times 100\%$ HPDI provides the shortest marginal intervals with $(1-\alpha)\times 100\%$ coverage. The MMAP prediction and HPD prediction intervals for unimodal and symmetric and for bimodal posterior distributions are illustrated in Figure 12. Note that for unimodal and symmetric pdfs, such as a Gaussian pdf, the MMAP prediction and HPD prediction interval coincide with the traditional prediction and prediction interval. For bimodal pdfs, the HPDI may appear as two intervals with the MMAP located within one of them.

The posterior distribution may support nonphysical values outside the interval $[0, 1]$. Therefore, we correct the predictor and prediction intervals by truncation to $[0, 1]$. Later, we observe that, for our case studies, this truncation does not appear as a major problem.

We use the mean absolute error (MAE) as the criterion when comparing MMAP predictions with observations. This criterion is suitable for multimodal variables because it does not favor central expectation predictions as strongly as the alternative MSE. It is favorable to have MAE close to zero. To evaluate the $(1-\alpha)\times 100\%$ HPD prediction intervals, we use coverage, defined as the proportion of observations inside the intervals. An observed coverage of $(1-\alpha)\times 100\%$ is of course optimal.

Synthetic case studies

The S-GRF model is used to predict porosity and saturation from the synthetic seismic AVO data shown in Figures 8 and 9. These data originate from synthetic case studies A and B, respectively. The results are compared to results based on the GRF model.

Case A

We investigate the convergence of the MCMC algorithm used to assess the posterior distribution before running the inversion method. The posterior pdf of saturation is sampled 100,000 times, and Figure 13 shows the location-wise convergence. The location-wise posteriors of saturation are unimodal and bimodal along the reservoir profile, and the convergence at the locations where the posterior is strongly bimodal is slow (e.g., at depth 2132 m). The MCMC algorithm appears to be stable after 50,000 iterations, so we set the burn-in to 50,000 samples and base the results on the last 50,000 samples. In all subsequent inversions, the convergence of the MCMC algorithm is checked before obtaining the results by inspecting similar convergence plots. The plots are used to set the burn-in, and the results are based on 50,000 samples following the burn-in. The posterior pdf of porosity is unimodal at all grid locations; hence, convergence is relatively fast and does not influence the burn-in.

Figure 14 shows the inversion results of the synthetic seismic AVO data associated with case A; see Figure 8. The model parameters of the S-GRF prior model are listed in Table 2.

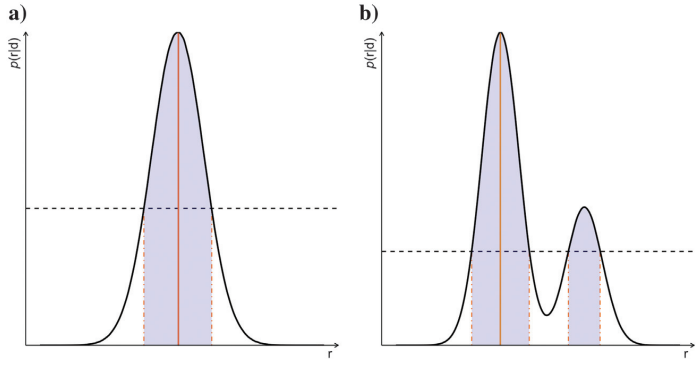

Figure 12. HPDI illustrated for (a) unimodal symmetric and (b) bimodal posterior. The limits of the prediction intervals (the dotted-dashed orange lines) and the posterior within (the blue fill), the MMAPs (the orange lines), and the minimum probability values associated with prediction interval inclusion (the dashed horizontal lines) are shown.

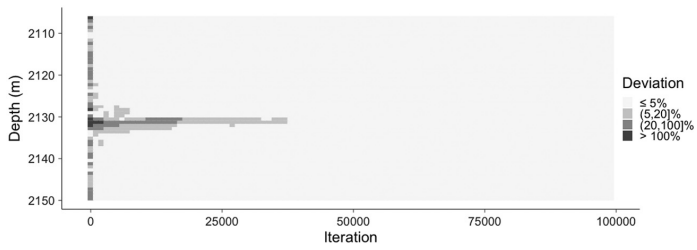

Figure 13. Convergence of the posterior distribution of saturation displayed as percentage deviation between the running sample average and the total sample average of 100,000 samples for every 1000 samples.

The porosity predictions (see Figure 14a) appear as well centered with respect to the observations, except for some border effects at the top. Furthermore, the prediction intervals look reasonable. The saturation predictions (see Figure 14b) are also relatively well centered in the observations, and are not truncated. They have a clear OWC transition, which occurs approximately 1 m higher than indicated by the well observations. The prediction intervals appear to have too high coverage. Note that the posterior pdf is clearly bimodal in the approximate depth range of 2125–2135 m. Furthermore, the pdf of the OWC location (see Figure 14c) indicates that the OWC is most likely to be in this depth range (2128–2135 m), and it is centered at approximately 2131 m depth. The pdf of the OWC location is obtained by reading the OWC location out of every realization according to a reasonable rule. In Figure 14e, we observe clear asymmetry in the marginal posterior pdf of saturation at depth 2132 m, and that the observed saturation is not very well predicted at this particular depth.

To evaluate the sensitivity of the inversions to the prior model assumptions, we use an alternative S-GRF prior model with asymmetric and bimodal location-wise pdf for saturation. Figure 15 shows the results from the inversion in a format identical to Figure 14. Note in particular the location-wise prior pdf of saturation in Figure 15d. The results appear as very similar to the ones obtained with the previous model, and the pdf of the OWC location (see Figure 15c) is now centered at approximately 2132 m depth. The asymmetric prior model favors oil and causes a lowering of the predicted OWC to approximately 0.5 m above the well observation indications. This result is as expected because oil tends to be assigned to the locations where the uncertainty is high. Reproduction of the observed OWC can probably be obtained by assigning an even more asymmetric prior saturation model, but in practice the true location of the OWC is unknown and the relative mode sizes should reflect the expected fluid proportions.

Figure 16 shows the inversion results based on the GRF prior model, in the same format as Figure 14. The model parameters of the GRF prior model are listed in Table 3. The porosity inversion results (see Figure 16a) are very similar to the corresponding results based on the S-GRF prior model in Figure 14a. This is as expected because porosity has a symmetric and unimodal prior model in both models. The saturation predictions (see Figure 16b) appear as very smooth and are frequently truncated. Moreover, the prediction intervals are severely overestimated, that is, far too wide. The pdf of the OWC location (see Figure 16c) indicates that the OWC is very likely to be in the depth range 2125–2138 m, and it is centered at approximately 2132 m depth. Note that this is a bigger depth range than for the S-GRF results. The pdf of the OWC location is obtained by reading the OWC location out of every realization, after smoothing, according to the same rule applied to the S-GRF realizations. The marginal prior and posterior saturation pdfs at depth 2132 m in Figure 16d and 16e are consistent with these results, and they expose a predictor that is clearly regressed toward the prior mean.

Figures 17 and 18 display four realizations of porosity and saturation profiles from the S-GRF and GRF posterior models, respectively. Spatial histograms of the realizations are presented at the bottom of each display. The simulated porosity profiles appear very similar regardless of model, as expected, being relatively smooth with unimodal spatial histograms. The simulated saturation profiles are, however, very different for the two models. The realizations from the S-GRF model (see Figure 17b) appear with a clear transition between fluids, indicating an OWC. Note that the depth of this OWC is fairly uncertain, and it appears in the depth range of 2129–2134 m. The OWC depth indicated by the well observations, 2132 m, is within this range. Moreover, the spatial histogram of the realizations is clearly bimodal and fairly symmetric. The saturation profile realizations from the GRF model (see Figure 18b) are fairly smooth without a clear fluid transition and the realizations fre-

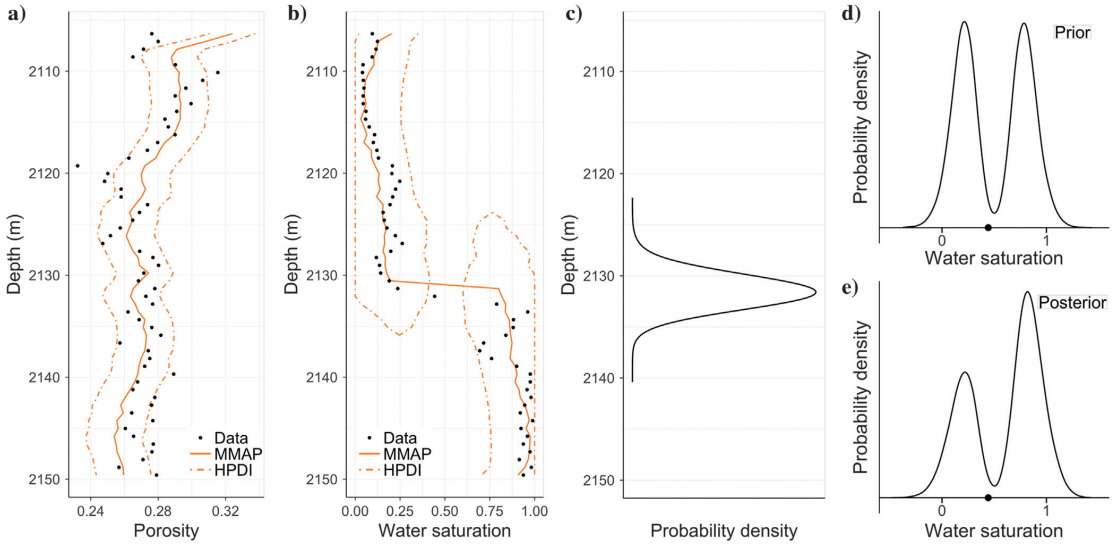

Figure 14. S-GRF inversion results in case A: (a) porosity and (b) saturation profiles with 80% HPDI, (c) pdf of the OWC location, and (d) marginal prior and (e) marginal posterior pdfs of saturation at depth 2132 m.

quently fall outside the [0,1] range. The corresponding spatial histogram is unimodal and very wide.

The inversion results based on the two models — the S-GRF and the traditional GRF models, displayed in Figures 14 and 16, respectively — appear as dramatically different for saturation. The former inversion results are more concordant to the well observations and the physical understanding that an abrupt OWC transition exists.

Case B

The challenge in this case, relative to case A, is to identify the two OWCs with the oil zone between them. Figure 19 shows the inversion results of the synthetic seismic AVO data associated with case B (see Figure 9) based on an S-GRF prior model. The prior model parameters are identical to the ones in case A (see Table 2), and the format of the figure is identical to Figure 14.

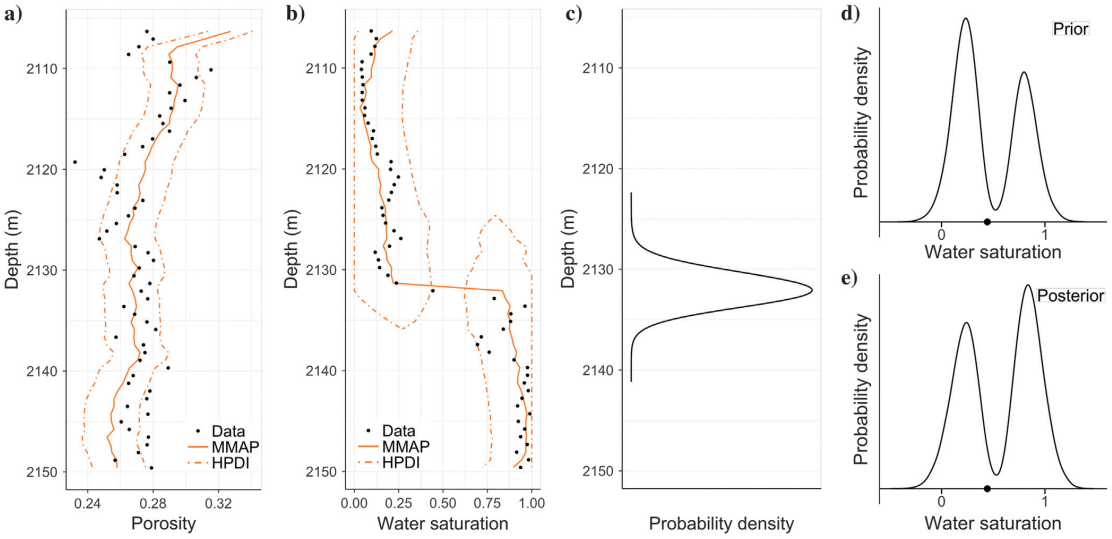

Figure 15. S-GRF inversion results in case A with alternative prior: (a) porosity and (b) saturation profiles with 80% HPDI, (c) pdf of the OWC location, and (d) marginal prior and (e) marginal posterior pdfs of saturation at depth 2132 m.

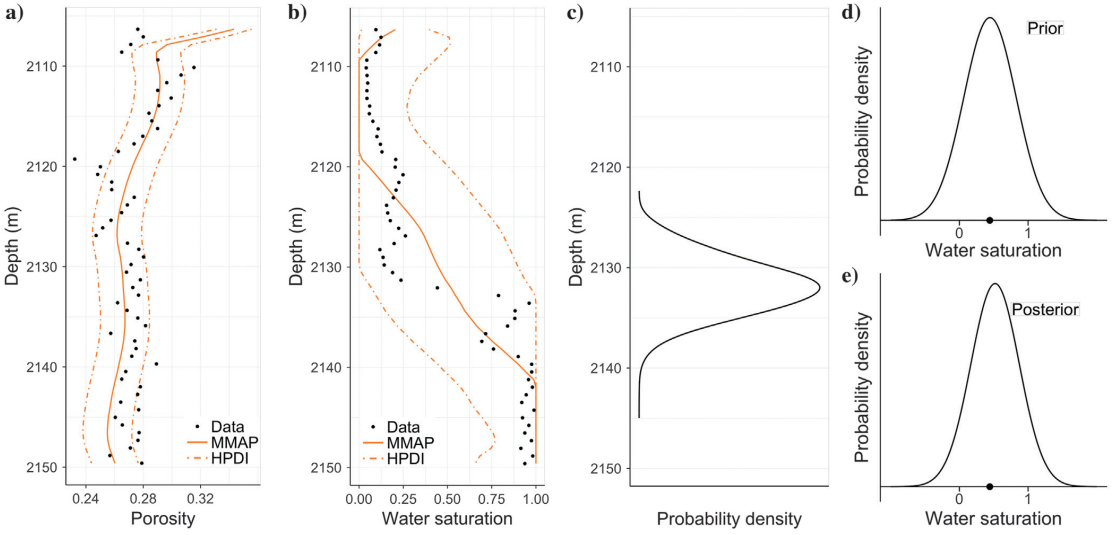

Figure 16. GRF inversion results in case A: (a) porosity and (b) saturation profiles with 80% HPDI, (c) pdf of the OWC location, and (d) marginal prior and (e) marginal posterior pdfs of saturation at depth 2132 m.

The porosity predictions and prediction intervals (see Figure 19a) are very similar to the ones in case A. The saturation predictions (see Figure 19b) reproduce the well observations reliably and are not truncated. They appear with a few fluid transitions and with two distinct OWCs at approximately 2119 and 2135 m depth. It is very encouraging that multiple OWCs can be identified, although both contact levels are approximately 2–3 m too deep. The thickness of the oil zone between the OWCs is underpredicted; the intermediate oil zone is roughly 9 m, whereas the predictions indicate that it is approximately 7 m. However, the prediction intervals indicate that the posterior has support for up to 10 m thickness. The pdf of the main OWC location (see Figure 19c) indicates that the main OWC is very likely to be in the depth range of 2131–2139 m, and it is centered at approximately 2135 m depth. The marginal prior and posterior pdfs of saturation at depth 2132 m, shown in Figure 19d and 19e, respectively, appear as expected based on the prediction intervals.

Figure 20 shows the inversion results based on the GRF model, in a format identical to Figure 14. The model parameters are identical to the ones in case A (see Table 3). The porosity results are very similar to the porosity results obtained with the S-GRF model, whereas the saturation results (see Figure 20b) are very different. The predictions are overly smooth and regressed toward the prior mean; hence, no abrupt OWC transitions can be identified. Moreover, the predictions do not reflect the well observations and they are frequently truncated. The marginal prior and posterior pdfs of saturation at depth 2132 m (see Figure 20d and 20e, respectively), reflect these features. The pdf of the main OWC location (see Figure 20c) indicates that the main OWC is very likely to be in the depth range of 2132–2141 m, and it is centered at approximately 2136 m depth. Note that this is a larger depth range than for the S-GRF results.

Figures 21 and 22 display realizations of porosity and saturation profiles from the S-GRF and GRF posterior models, respectively. The spatial histograms are also presented. The simulated porosity profiles appear very similar for the two models, as expected. The simulated saturation profiles from the S-GRF model (see Figure 21b) contain abrupt transitions, but the simulated OWC locations are fairly uncertain. The spatial histogram is clearly bimodal. The saturation profile realizations from the GRF model (see Figure 22b) appear as smooth without clear OWC locations, and with a unimodal and wide spatial histogram. The realizations frequently fall outside the valid $[0, 1]$ range.

The inversion results based on the two models, the S-GRF and GRF models, displayed in Figures 19 and 20, respectively, appear as dramatically different for saturation. The results based on the former model are more concordant to the well observations and to our experience.

DISCUSSION

Table 4 contains the prediction MAE values and prediction interval coverages for the two models, for cases A and B. For porosity, the model performances are very similar. The MAE values for the S-GRF model are favorable for saturation: the MAE value for the S-GRF model is approximately 45% lower than the MAE value for the GRF model in case A, and in case B the MAE value for the S-GRF model is slightly lower than that of the GRF model. Overall, the S-GRF model therefore appears to be favorable. The picture is more mixed when it comes to prediction interval coverages. In theory, the coverage should be 0.80, but because the intervals are truncated, the interpretation is more complex. The coverages are mostly close to 0.80, but none of the models appear as favorable.

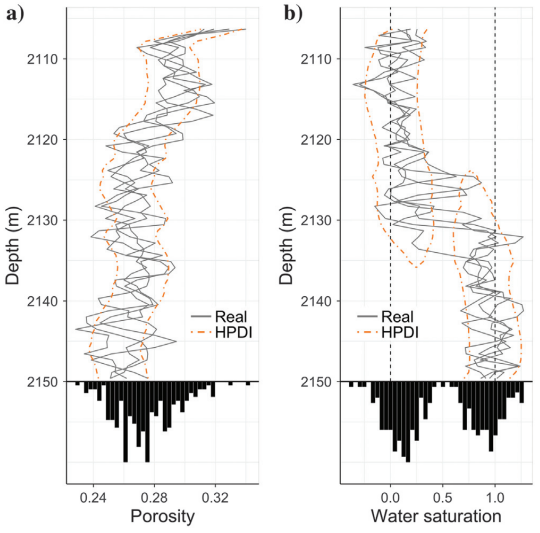

Figure 17. Four realizations (real) from the S-GRF posterior model in case A: (a) porosity and (b) saturation profiles with 80% HPDI. Spatial histograms of the realizations (the black bars) are shown in the lower part of the displays.

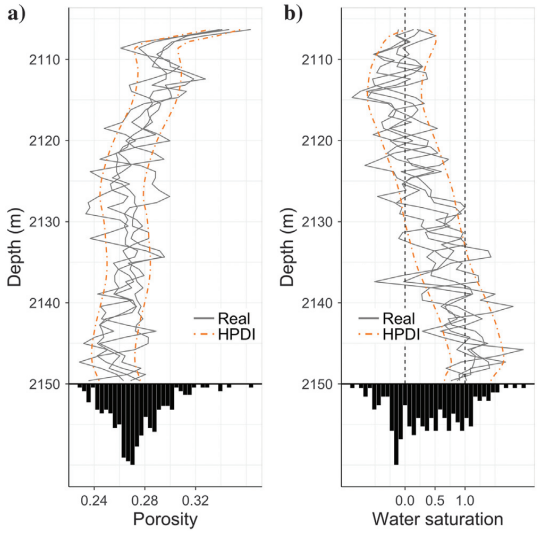

Figure 18. Four realizations (real) from the GRF posterior model in case A: (a) porosity and (b) saturation profiles with 80% HPDI. Spatial histograms of the realizations (the black bars) are shown in the lower part of the displays.

Both models have difficulties accurately predicting the saturations in case B. This is most evident for the S-GRF model because the fluid transitions are abrupt, and the mislocation of the intermediate oil zone and the underprediction of its thickness stands out as an issue. These issues are most likely not a limitation of the methodology; rather, a reasonable explanation lies in the coarseness of the seismic grid $\mathcal{L}_d$ relative to the reservoir grid $\mathcal{L}_r$, which reduces the accuracy of the predictions.

Computationally, the S-GRF inversions are more demanding than the corresponding GRF inversions. The S-GRF inversions in this study require approximately 6 min on a standard laptop computer, whereas the GRF inversions are made within one second.

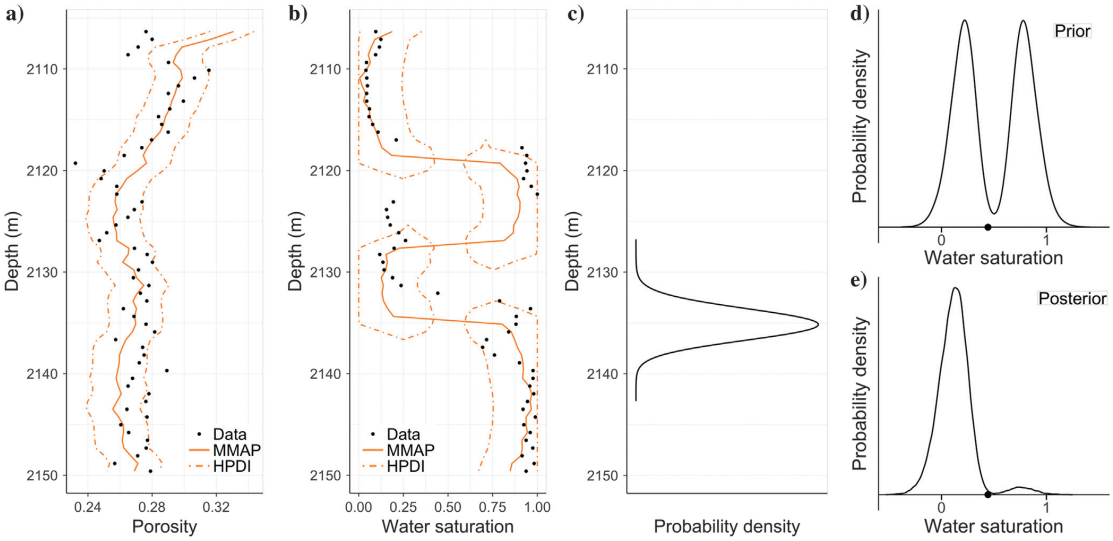

Figure 19. S-GRF inversion results in case B: (a) porosity and (b) saturation profiles with 80% HPDI, (c) pdf of the OWC location, and (d) marginal prior and (e) marginal posterior pdfs of saturation at depth 2132 m.

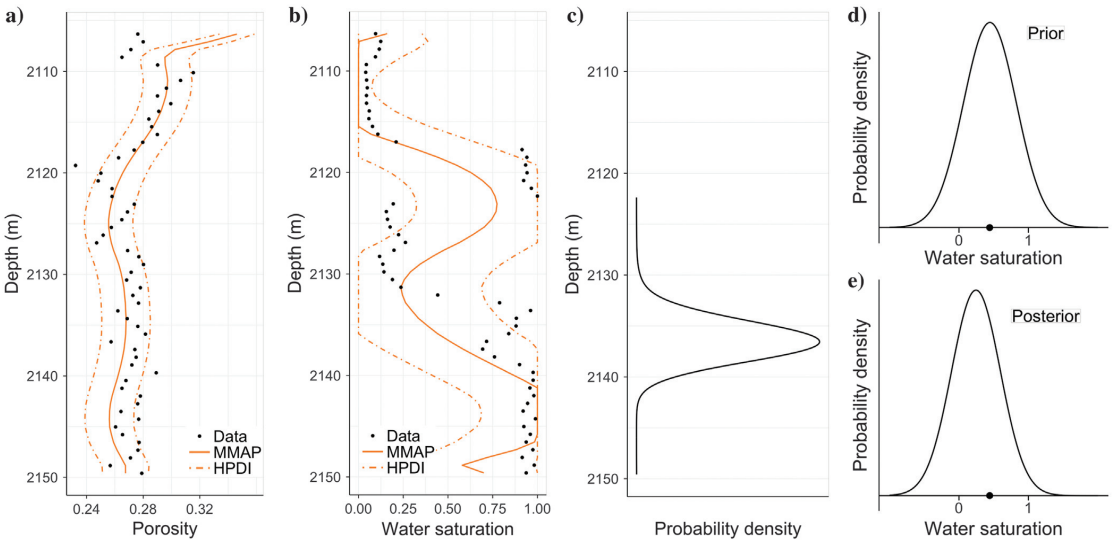

Figure 20. GRF inversion results in case B: (a) porosity and (b) saturation profiles with 80% HPDI, (c) pdf of the OWC location, and (d) marginal prior and (e) marginal posterior pdfs of saturation at depth 2132 m.

RESULTS — REAL SEISMIC DATA

Inversion of the real seismic AVO data displayed in Figure 2 is more challenging than inversion of synthetic cases. The corresponding real well observations (see Figure 1) contain several thin shale layers not included in our prior models. Moreover, the fairly simple Gauss-linear likelihood model may not be adequately representative of the real physics of the seismic data collection. Lastly, the magnitude and composition of the noise in the data are not known.

Figure 23 shows the real seismic data together with synthetically generated seismic data. The expected synthetic seismic data from the extended reservoir profile are used; that is, the seismic data are generated without noise. The seismic fit is good; the correlation between corresponding angle stacks is 0.96, 0.95, 0.56, and 0.50 for angles 10°, 18°, 27°, and 35°, respectively. Furthermore, the S/N is estimated to be 10.

Figure 24 shows the inversion results of the real seismic AVO data displayed in Figure 2 based on the S-GRF prior model with the parameters listed in Table 2. The format of the figure is identical to Figure 14. The predicted porosity (see Figure 24a) is in good agreement with the observed porosity, except for in the shale layers and at the top and bottom of the depth interval, where we observe mild border effects. The predicted saturation (see Figure 24b) is for the most part in good agreement with the observed saturation. The OWC is predicted to be at approximately 2135 m depth, roughly 3 m below what the well observations indicate. We also see that a water zone is predicted between 2120 and 2123 m depth, coinciding with the location of the major shale layer observed in the well. The volume of the oil column in the reservoir is accurately predicted. Moreover, the saturation predictions are centered in the corresponding well observations. The prediction intervals indicate a posterior model with clear bimodality in the depth interval of the major shale layer, which opens for an alternative saturation interpretation. The pdf of the OWC location (see Figure 24c) indicates that the OWC is very likely to be in the depth range of 2132–2139 m, and it is centered at approximately 2136 m depth. Lastly, the marginal pdfs at depth 2132 m (see Figure 24d and 24e) appear as expected based on the prediction intervals.

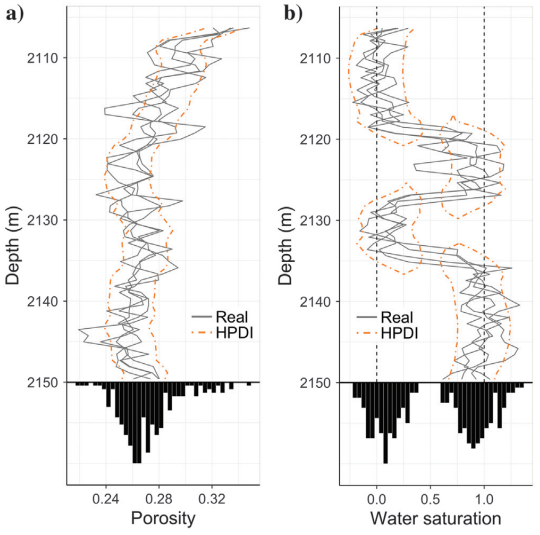

Figure 21. Four realizations (real) from the S-GRF posterior model in case B: (a) porosity and (b) saturation profiles with 80% HPDI. Spatial histograms of the realizations (the black bars) are shown in the lower part of the displays.

Figure 25 shows the inversion results of the real seismic AVO data displayed in Figure 2 based on the GRF model, in a format identical to Figure 14. The model parameters are listed in Table 3. The predicted porosity (see Figure 25a) is in good agreement with the observed porosity, except for in the shale layers and at the top of the depth interval, where we observe mild border effects. The predicted saturation (see Figure 25b) is much smoother than the well observations, and it is difficult to identify the location of the OWC or of other changes in the fluid filling. This effect is caused by regression toward the prior mean. Furthermore, the prediction intervals are very wide. The pdf of the OWC location (see Figure 25c) indicates that the OWC is very likely to be in the depth range of 2132–2141 m, and it is centered at approximately 2136 m depth.

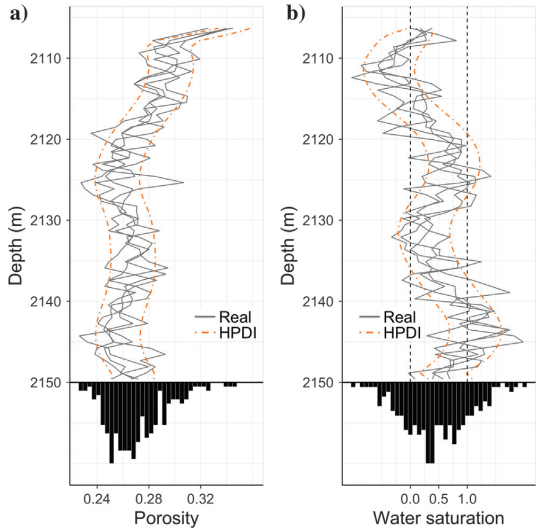

Figure 22. Four realizations (real) from the GRF posterior model in case B: (a) porosity and (b) saturation profiles with 80% HPDI. Spatial histograms of the realizations (the black bars) are shown in the lower part of the displays.

Table 4. Quantitative comparison: prediction MAE values and prediction interval coverages for the S-GRF and GRF model for cases A and B.

		MAE		Coverage	
Case	Model	ϕ	s_w	ϕ	s_w
A	GRF	0.0126	0.1063	0.7414	1.0000
A	S-GRF	0.0113	0.0600	0.7414	0.9828
B	GRF	0.0120	0.2118	0.7931	0.8276
B	S-GRF	0.0113	0.1796	0.7931	0.8276

The marginal prior and posterior pdfs at depth 2132 m (see Figure 25d and 25e) appear as expected based on the prediction intervals.

Figures 26 and 27 display realizations of porosity and saturation profiles from the S-GRF and GRF posterior models, respectively. The spatial histograms are also presented. As expected, we observe that the simulated porosity profiles from the two posterior distributions are very similar, with unimodal spatial histograms. The simulated saturation profiles, however, are very different for the S-GRF and GRF posterior models. The realizations from the former (see Figure 26b) appear with abrupt mode transitions, indicating changes in the fluid filling. There are local mode transitions above the OWC, between 2119 and 2124 m depth, which probably are related to the major shale layer. Note that this local apparent water zone may not be significant because the marginal posterior pdf is strongly bimodal in this depth interval, which is reflected by some realizations not changing mode. Furthermore, the spatial histogram is

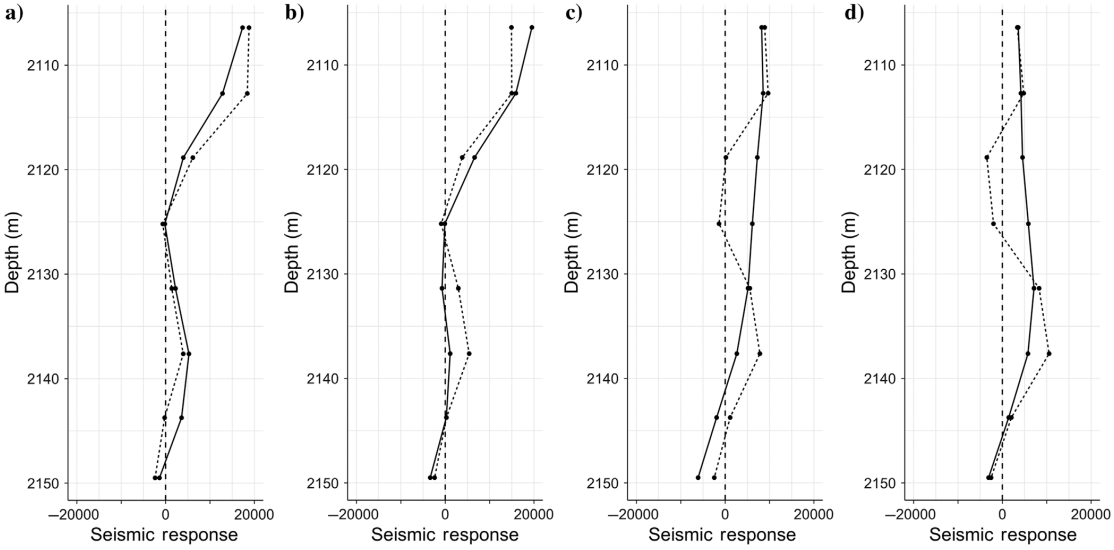

Figure 23. Seismic AVO fit: (a) 10°, (b) 18°, (c) 27°, and (d) 35°; real seismic data (the solid lines) and synthetic seismic data (the dashed lines).

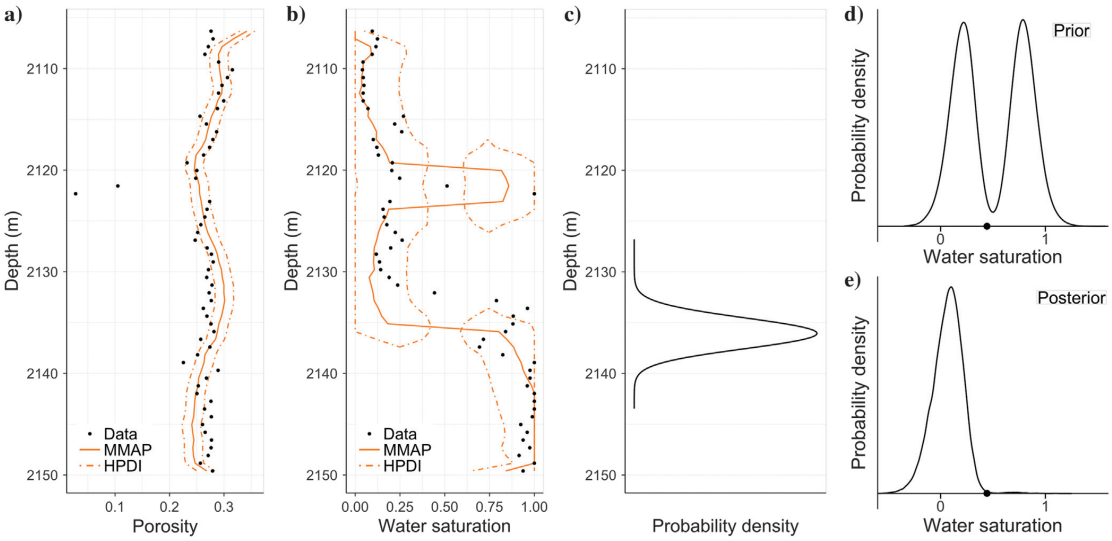

Figure 24. S-GRF inversion results for real seismic data: (a) porosity and (b) saturation profiles with 80% HPDI, (c) pdf of the OWC location, and (d) marginal prior and (e) marginal posterior pdfs of saturation at depth 2132 m.

clearly bimodal as seen in the well observations. The realizations from the GRF posterior model (see Figure 27b) appear with large variance, but no abrupt mode transitions that may clearly be interpreted as changes in the fluid filling. The spatial histograms are unimodal, and the realizations frequently appear notably outside the $[0, 1]$ range.

The inversion results based on the S-GRF and GRF models, displayed in Figures 24 and 25, respectively, appear as very similar for porosity, whereas the saturation results are dramatically different. The former model produces inversion results that reproduce the well observations better than the latter model.

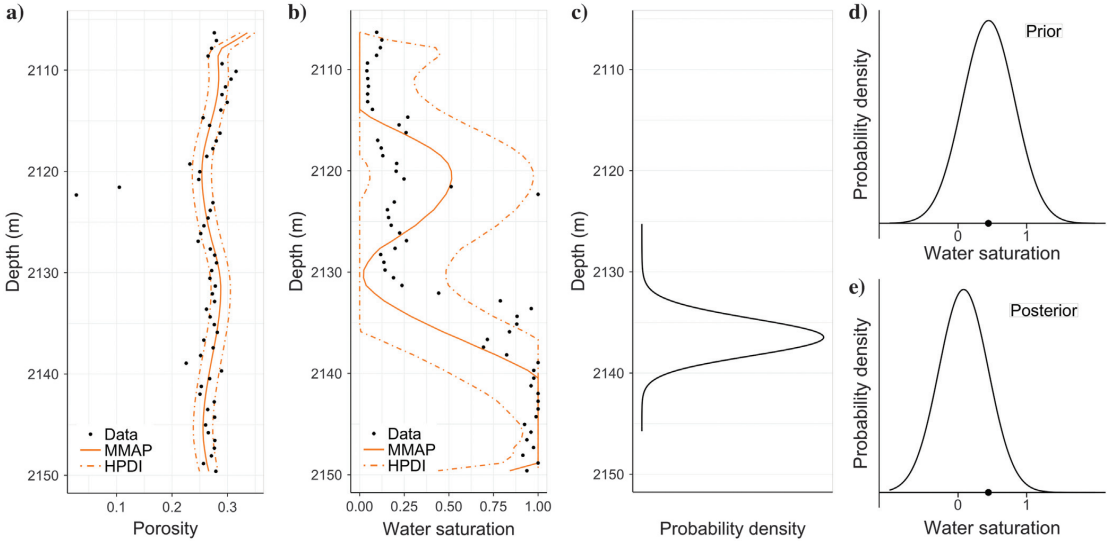

Figure 25. GRF inversion results for real seismic data: (a) porosity and (b) saturation profiles with 80% HPDI, (c) pdf of the OWC location, and (d) marginal prior and (e) marginal posterior pdfs of saturation at depth 2132 m.

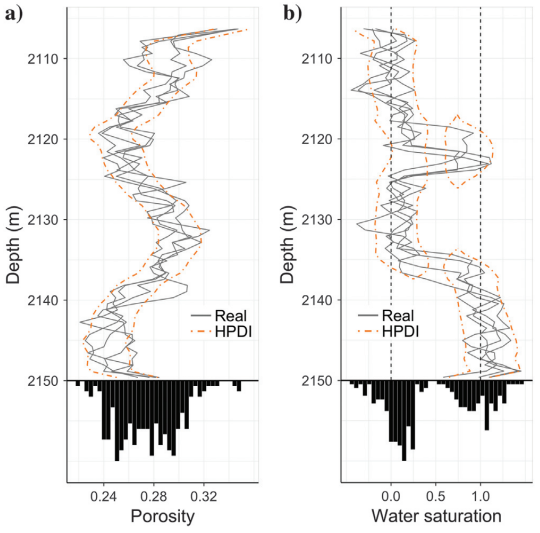

Figure 26. Four realizations (real) from the S-GRF posterior model for real seismic data: (a) porosity and (b) saturation profiles with 80% HPDI. Spatial histograms of the realizations (the black bars) are shown in the lower part of the displays.

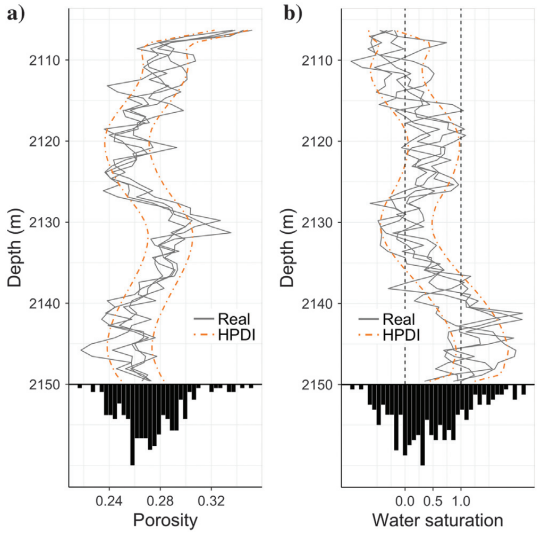

Figure 27. Four realizations (real) from the GRF posterior model for real seismic data: (a) porosity and (b) saturation profiles with 80% HPDI. Spatial histograms of the realizations (the black bars) are shown in the lower part of the displays.

Table 5. Quantitative comparison: prediction MAE values and prediction interval coverages for the S-GRF and GRF model for real seismic data.

	MAE		Coverage	
Model	ϕ	s_w	ϕ	s_w
GRF	0.0206	0.1563	0.6207	0.8966
S-GRF	0.0202	0.1485	0.5172	0.8966

Table 5 contains the prediction MAE values and the prediction interval coverages for the two models for the real data case. The MAE values for porosity are very similar for the two models, whereas the S-GRF MAE value for saturation is better than that of the GRF. The prediction interval coverages for porosity differ in favor of the GRF model, but are severely below the ideal coverage of 0.80 for both models. This underestimation is a shale layer effect. The prediction interval coverages for saturation are identical and appear reasonable for both models, but truncation makes a clear interpretation difficult.

CONCLUSION

The seismic AVO inversion problem is cast in a Bayesian framework. The likelihood model is Gauss-linear, and the prior model is an S-GRF, which is chosen due to its capability of representing the observed multimodality in the reservoir variables. Seismic AVO inversion of two synthetic data sets and of real seismic AVO data is performed. The synthetic studies are used to demonstrate the methodology and to explore its ability to identify fairly thin sections of locally extreme saturation values along the profile. The inversion results based on the S-GRF prior model are compared to corresponding results based on a traditional GRF prior model. The former inversion results are clearly favorable because they reproduce the abrupt changes in fluid filling for saturation and have lower MAE values than the latter. The inversion results based on the real seismic AVO data are very encouraging. The real porosity and saturation observations along the well are reliably reproduced. The saturation profile is predicted with abrupt transitions as observed in the well. Inversion based on a traditional GRF prior model reproduces these saturation transitions very poorly.

The methodology can be extended to 2D and 3D inversion problems at the expense of increased computational demands; however, the computation time need not increase severely because the current implementation can be improved and parallelized. The extension can be done on a trace-by-trace basis, or it can include a lateral correlation structure. This is currently a work in progress.

ACKNOWLEDGMENTS

This research is part of the Uncertainty in Reservoir Evaluation activity at the Norwegian University of Science and Technology. The authors would like to thank Aker BP for data access and the right to publish. The authors acknowledge the use of the R-package MASS in the computational algorithms used.

DATA AND MATERIALS AVAILABILITY

Data associated with this research are confidential and cannot be released.

APPENDIX A

SAMPLING FROM AN S-GRF

To draw samples from an S-GRF $\mathbf{x}$, we make use of the decomposition

$$p(\mathbf{x}) = p(\tilde{\mathbf{x}}|\boldsymbol{\nu}\in\mathbf{A}) = p(\tilde{\mathbf{x}}|\boldsymbol{\nu})p(\boldsymbol{\nu}|\boldsymbol{\nu}\in\mathbf{A}) = \varphi_n(\tilde{\mathbf{x}};\boldsymbol{\mu}_{\tilde{x}|\nu},\boldsymbol{\Sigma}_{\tilde{x}|\nu})\cdot\frac{I(\boldsymbol{\nu}\in\mathbf{A})\varphi_n(\boldsymbol{\nu};\boldsymbol{\mu}_\nu,\boldsymbol{\Sigma}_\nu)}{\int_{\mathbf{R}^n} I(\boldsymbol{\nu}\in\mathbf{A})\varphi_n(\boldsymbol{\nu};\boldsymbol{\mu}_\nu,\boldsymbol{\Sigma}_\nu)\mathrm{d}\boldsymbol{\nu}}. \tag{A-1}$$

A sample is drawn by first sampling $\boldsymbol{\nu}\in\mathbf{A}$, followed by sampling $\mathbf{x}$ from $\varphi_n(\tilde{\mathbf{x}};\boldsymbol{\mu}_{\tilde{x}|\nu},\boldsymbol{\Sigma}_{\tilde{x}|\nu})$.

Algorithm 1: Draw m samples from an S-GRF by Metropolis-Hastings MCMC.

Initialize $\boldsymbol{\nu}$ with a value in $\mathbf{A}$.
Iterate m times
  Select spatial block $b\subset\{1,\ldots,n\}$.
  Sample $\boldsymbol{\nu}'_{b|-b}$ sequentially from $q(\nu'_{b|-b})$.
  Accept $\boldsymbol{\nu}=[\boldsymbol{\nu}'_{b|-b},\boldsymbol{\nu}_{-b}]^T$ with probability α.
  Sample $\mathbf{x}\sim\varphi_n(\tilde{\mathbf{x}}|\boldsymbol{\nu})$.
End

Some clarifications of the algorithm follow:

1) Subindices b and $-b$ represent grid point sets $b\subset\{1,\ldots,n\}$ and $\{1,\ldots,n\}\setminus b$, respectively.
2) The proposal distribution is

$$q(\boldsymbol{\nu}_{b|-b}) = \prod_{i\in b} I(\nu_i\in A)\frac{\varphi_1(\nu_i|\boldsymbol{\nu}_{b_v},\boldsymbol{\nu}_{-b};\boldsymbol{\mu}_\nu,\boldsymbol{\Sigma}_\nu)}{\Phi_1(\nu_i\in A|\boldsymbol{\nu}_{b_v},\boldsymbol{\nu}_{-b};\boldsymbol{\mu}_\nu,\boldsymbol{\Sigma}_\nu)}, \tag{A-2}$$

with b_v denoting the set of indices in b already visited.
3) The Metropolis-Hastings acceptance probability α is

$$\alpha = \min\left\{1,\frac{p(\boldsymbol{\nu}'_{b|-b})}{p(\boldsymbol{\nu}_{b|-b})}\cdot\frac{q(\boldsymbol{\nu}_{b|-b})}{q(\boldsymbol{\nu}'_{b|-b})}\right\} = \min\left\{1,\prod_{i\in b}\frac{\Phi_1(\nu'_i\in A|\boldsymbol{\nu}'_{b_v},\boldsymbol{\nu}_{-b};\boldsymbol{\mu}_\nu,\boldsymbol{\Sigma}_\nu)}{\Phi_1(\nu_i\in A|\boldsymbol{\nu}_{b_v},\boldsymbol{\nu}_{-b};\boldsymbol{\mu}_\nu,\boldsymbol{\Sigma}_\nu)}\right\}. \tag{A-3}$$

The blocks and the order in which they are visited are predetermined, and the associated covariance matrices are precomputed to reduce the computational time.

APPENDIX B

THE POSTERIOR S-GRF

The discretized posterior S-GRF model is

$$
\begin{aligned}
p(\mathbf{r}|\mathbf{d}) &= p(\boldsymbol{\phi}, \mathbf{s}_w|\mathbf{d}) = p(\boldsymbol{\phi}|\tilde{\mathbf{s}}_w, \mathbf{d}) p(\tilde{\mathbf{s}}_w|\mathbf{d}, \boldsymbol{\nu} \in \mathbf{A}) \\
&= \varphi_{n_r}(\boldsymbol{\phi}; \boldsymbol{\mu}_{\phi|\tilde{s}_{w,d}}, \boldsymbol{\Sigma}_{\phi|\tilde{s}_{w,d}}) \frac{\Phi_{n_r}(\mathbf{A}; \boldsymbol{\mu}_{\nu|\tilde{s}_w,d}, \boldsymbol{\Sigma}_{\nu|\tilde{s}_w,d})}{\Phi_{n_r}(\mathbf{A}; \boldsymbol{\mu}_{\nu|d}, \boldsymbol{\Sigma}_{\nu|d})} \\
&\times \varphi_{n_r}(\tilde{\mathbf{s}}_w; \boldsymbol{\mu}_{\tilde{s}_w|d}, \boldsymbol{\Sigma}_{\tilde{s}_w|d}) \\
&= \frac{\Phi_{n_r}(\mathbf{A}; \boldsymbol{\mu}_{\nu|\tilde{\mathbf{r}},d}, \boldsymbol{\Sigma}_{\nu|\tilde{\mathbf{r}},d})}{\Phi_{n_r}(\mathbf{A}; \boldsymbol{\mu}_{\nu|d}, \boldsymbol{\Sigma}_{\nu|d})} \varphi_{2n_r}(\tilde{\mathbf{r}}; \boldsymbol{\mu}_{\tilde{r}|d}, \boldsymbol{\Sigma}_{\tilde{r}|d}). \qquad \text{(B-1)}
\end{aligned}
$$

The last equality follows because $\boldsymbol{\nu}$ is conditionally independent of $\boldsymbol{\phi}$ given $\tilde{\mathbf{s}}_w$; that is, $[\boldsymbol{\nu}|\boldsymbol{\phi}, \tilde{\mathbf{s}}_w, \mathbf{d}] = [\boldsymbol{\nu}|\tilde{\mathbf{s}}_w, \mathbf{d}]$.

The expressions for the parameters involved in the posterior model can be developed from the classic Gaussian theory; see Johnson and Wichern (2007). The parameters involved in the GRF basis prior model are

$$
\boldsymbol{\mu}_{\tilde{r}} = \begin{bmatrix} \mu_\phi \mathbf{i}_{n_r} \\ \mu_{\tilde{s}_w} \mathbf{i}_{n_r} \end{bmatrix}, \quad \boldsymbol{\Sigma}_{\tilde{r}} = \begin{bmatrix} \sigma_\phi^2 \boldsymbol{\Omega} & \rho_{\phi\tilde{s}_w} \sigma_\phi \sigma_{\tilde{s}_w} \boldsymbol{\Omega} \\ \rho_{\phi\tilde{s}_w} \sigma_\phi \sigma_{\tilde{s}_w} \boldsymbol{\Omega} & \sigma_{\tilde{s}_w}^2 \boldsymbol{\Omega} \end{bmatrix}. \qquad \text{(B-2)}
$$

It then follows that

$$
\begin{bmatrix} \boldsymbol{\mu}_{\tilde{r}|d} \\ \boldsymbol{\mu}_{\nu|d} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}_{\tilde{r}} \\ 0\mathbf{i}_{n_r} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\Sigma}_{\tilde{r}} \mathbf{G}^T \\ \boldsymbol{\Gamma}_{\nu\tilde{r}} \mathbf{G}^T \end{bmatrix} \boldsymbol{\Sigma}_d^{-1} (\mathbf{d} - \boldsymbol{\mu}_d), \qquad \text{(B-3)}
$$

and the conditional covariance matrices are

$$
\begin{bmatrix} \boldsymbol{\Sigma}_{\tilde{r}|d} & \boldsymbol{\Gamma}_{\tilde{r}\nu|d} \\ \boldsymbol{\Gamma}_{\nu\tilde{r}|d} & \boldsymbol{\Sigma}_{\nu|d} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Sigma}_{\tilde{r}} & \boldsymbol{\Gamma}_{\tilde{r}\nu} \\ \boldsymbol{\Gamma}_{\nu\tilde{r}} & \boldsymbol{\Sigma}_{\nu} \end{bmatrix} - \begin{bmatrix} \boldsymbol{\Sigma}_{\tilde{r}} \mathbf{G}^T \\ \boldsymbol{\Gamma}_{\nu\tilde{r}} \mathbf{G}^T \end{bmatrix} \boldsymbol{\Sigma}_d^{-1} \begin{bmatrix} \mathbf{G}\boldsymbol{\Sigma}_{\tilde{r}} & \mathbf{G}\boldsymbol{\Gamma}_{\tilde{r}\nu} \end{bmatrix}, \qquad \text{(B-4)}
$$

hence

$$
\begin{aligned}
\boldsymbol{\mu}_{\nu|\tilde{r},d} &= \boldsymbol{\mu}_{\nu|d} + \boldsymbol{\Gamma}_{\nu\tilde{r}|d} \boldsymbol{\Sigma}_{\tilde{r}|d}^{-1} (\tilde{\mathbf{r}} - \boldsymbol{\mu}_{\tilde{r}|d}), \\
\boldsymbol{\Sigma}_{\nu|\tilde{r},d} &= \boldsymbol{\Sigma}_{\nu|d} + \boldsymbol{\Gamma}_{\nu\tilde{r}|d} \boldsymbol{\Sigma}_{\tilde{r}|d}^{-1} \boldsymbol{\Gamma}_{\tilde{r}\nu|d}. \qquad \text{(B-5)}
\end{aligned}
$$

REFERENCES

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

Arellano-Valle, R. B., M. D. Branco, and M. G. Genton, 2006, A unified view on skewed distributions arising from selections: The Canadian Journal of Statistics, **34**, no. 4, 581–601.

Ayani, M., and D. Grana, 2020, Statistical rock physics inversion of elastic and electrical properties for CO2 sequestration studies: Geophysical Journal International, **223**, 707–724, doi: 10.1093/gji/ggaa346.

Besag, J., 1974, Spatial interaction and the statistical analysis of lattice systems: Journal of the Royal Statistical Society: Series B (Methodological), **36**, 192–225, doi: 10.1111/j.2517-6161.1974.tb00999.x.

Buland, A., O. Kolbjørnsen, R. Hauge, Ø. Skjæveland, and K. Duffaut, 2008, Bayesian lithology and fluid prediction from seismic prestack data: Geophysics, **73**, no. 3, C13–C21, doi: 10.1190/1.2842150.

Buland, A., O. Kolbjørnsen, and H. Omre, 2003, Rapid spatially coupled AVO inversion in the Fourier domain: Geophysics, **68**, 824–836, doi: 10.1190/1.1581035.

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

Casella, G., and R. L. Berger, 2001, Statistical inference: Cengage Learning.

Eidsvik, J., P. Avseth, H. Omre, T. Mukerji, and G. Mavko, 2004, Stochastic reservoir characterization using prestack seismic data: Geophysics, **69**, 978–993, doi: 10.1190/1.1778241.

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.

Hasselblad, V., 1966, Estimation of parameters for a mixture of normal distributions: Technometrics, **8**, 431–444, doi: 10.1080/00401706.1966.10490375.

Johnson, R. A., and D. W. Wichern, 2007, Applied multivariate statistical analysis, 6th ed.: Pearson Prentice Hall.

Landrø, M., 2001, Discrimination between pressure and fluid saturation changes from time lapse data: Geophysics, **66**, 838–844, doi: 10.1190/1.1444973.

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.

Matheron, G., H. Beucher, C. de Fouquet, and A. Galli, 1987, Conditional simulation of the geometry of fluvio-deltaic reservoirs: Annual Technical Conference and Exhibition, SPE, Extended Abstracts, Paper 16753.

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

Mosegaard, K., and A. Tarantola, 1995, Monte Carlo sampling of solutions to inverse problems: Journal of Geophysical Research, **100**, 12431–12447, doi: 10.1029/94JB03097.

Omre, H., and K. Rimstad, 2018, Bayesian spatial inversion and conjugate selection Gaussian prior models, https://arxiv.org/abs/1812.01882.

Rimstad, K., P. Avseth, and H. Omre, Hierarchical Bayesian lithology/fluid prediction: A North Sea case study: Geophysics, **77**, no. 1, B69–B85, doi: 10.1190/geo2011-0202.1..

Tarantola, A., 2005, Inverse problem theory and methods for model parameter estimation: SIAM.

Ulvmoen, M., H. Omre, and A. Buland, 2010, Improved resolution in Bayesian lithology/fluid inversion from prestack seismic data and well observations — Part 1: Methodology: Geophysics, **75**, no. 2, R21–R35, doi: 10.1190/1.3294570.

White, R., and R. Simm, 2003, Tutorial — Good practice in well ties: First Break, **10**, 75–83.

Zoeppritz, K. B., 1919, Vllb. On reflection and transmission of seismic waves by surfaces of discontinuity: Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse, 66–84.

Biographies and photographs of the authors are not available.

Bayesian Inversion of Time-Lapse Seismic AVO Data for Multimodal Reservoir Properties

Ole Bernhard Forberg, Dario Grana, and Henning Omre

Published in *IEEE Transactions on Geoscience and Remote Sensing*, 2021, vol. 59, no. 11.

DOI: 10.1109/TGRS.2020.3046102

URL: https://ieeexplore.ieee.org/document/9329085

Bayesian seismic AVO inversion using a laterally coupled multimodal prior model

Ole Bernhard Forberg, Øyvind Kjøsnes, and Henning Omre

To appear in *IEEE Transactions on Geoscience and Remote Sensing* with minor changes.

DOI: 10.1109/TGRS.2021.3113865

URL: `https://ieeexplore.ieee.org/document/9570317`

Bayesian seismic AVO inversion using a laterally coupled multimodal prior model

Ole Bernhard Forberg, Øyvind Kjøsnes, and Henning Omre

Abstract

A Bayesian seismic amplitude versus offset (AVO) inversion scheme with a laterally coupled prior model for porosity, water saturation, and volume of clay is proposed. A 2D section and a 3D volume of an oil reservoir at the initial state is studied, which entails gravitationally induced bimodality in the water saturations along vertical traces. A selection Gaussian random field (S-GRF) prior model for porosity, water saturation, and volume of clay is specified, which is capable of representing this bimodality. The S-GRF is specified to have lateral correlation, which may reduce the impact of trace unique signal errors in the seismic AVO data. The likelihood model is linear and Gaussian, for which the S-GRF prior model is conjugate; hence, the posterior model is also an S-GRF. The form of the posterior distribution is therefore known and its parameter values can be analytically computed. Real seismic AVO data from the 2D section and the 3D volume is inverted and the results appear to be reliable along validation wells and represent a geologically plausible reservoir design. Furthermore, a notable variance reduction in the laterally coupled posterior model relative to an alternative posterior model without lateral coupling is achieved.

Introduction

The seismic properties of a reservoir are related to its petrophysical properties, which are the basis for reservoir characterization. Porosity and water saturation are informative about the storage capacity and fluid content of a reservoir and can be adequate for reservoir characterization in reservoirs that have a relatively homogeneous lithology. Within a fixed lithotype the seismic velocites and density tend to be negatively correlated with porosity, and the seismic velocities are higher in pores filled with water than in pores filled with hydrocarbon (Mavko et al., 2009). However, lithological heterogeneity may induce effects in the seismic properties that can not be accounted for by porosity and water saturation alone; hence, a lithological variable may be necessary to adequately characterize such reservoirs (Grana and Della Rossa, 2010; Rimstad et al., 2012). Seismic amplitude versus offset (AVO) data is often collected in order to characterize reservoirs. This data can be related to petrophysical properties by seismic models based on the Zoeppritz equations (Zoeppritz, 1919) and rock physics models. The seismic model can either be approximated or be based on full waveform, and similarly the rock physics model can either be approximated from empirical data or based on poroelasticity theory (Mavko et al., 2009).

In practice, the seismic AVO data are known while the petrophysical properties are unknown. The prediction of the petrophysical properties from the seismic AVO data can be formulated as an inverse problem (Tarantola, 2005). Seismic inverse problems are often approached probabilistically, in a Bayesian framework (Buland and Omre, 2003; Gunning and Glinsky, 2004; Larsen et al., 2006). This framework requires the definition of a likelihood model and a prior model, which together determine the form of the solution (Wang, 2006). The likelihood model is a probabilistic forward model for data acquisition and is often comprised of an approximate forward function and a stochastic error term (Buland and Omre, 2003; Grana et al., 2017; Forberg et al., 2021b). The prior model is a probability distribution representing experience and beliefs about the properties to be predicted. The solution to the Bayesian inverse problem is the posterior model, which is a probability distribution proportional to the product of the likelihood model and the prior model. A Gauss-linear likelihood model, defined by a linear forward model with Gaussian error terms, and a Gaussian random field (GRF) prior model, are computationally advantageous in a Bayesian seismic inversion framework (Buland et al., 2003). These model assumptions are convenient because the Gaussian prior model is conjugate with respect to Gauss-

linear likelihood models, i.e., that the posterior model is of the same form as the prior model. Consequently, the posterior model is Gaussian, which is extremely computationally advantageous in high dimensional settings since predictive quantities can be analytically computed. Should these assumptions be unjustifiable, one may use Markov chain Monte Carlo (McMC) methods (Mosegaard and Tarantola, 1995; Sen and Stoffa, 1996; Eidsvik et al., 2004) to assess the posterior distribution. These techniques are based on iterative algorithms that usually generate samples from the posterior model by proposing from the prior model and accepting based on the Metropolis rule applied to the likelihood of the proposed and current sample. The McMC algorithms tend to converge slowly for models that are defined on large spatial grids with strong spatial coupling; hence, the approach may be problematic in such cases.

Porosity and water saturation, and other continuous petrophysical or elastic properties, seldom appear as Gaussian due to underlying lithology and fluid classes. Rather, they appear as multimodal (Grana and Della Rossa, 2010); hence, a Gaussian prior model is usually not appropriate. The mixture Gaussian model (Hasselblad, 1966) offers support for multimodality and can be used in spatial settings. A spatial definition of the mixture Gaussian model requires a spatially defined mode indicator, for which a Markov random field prior model has been proposed (Rimstad et al., 2012; Fjeldstad and Grana, 2018). An advantage of this approach is that the mixture Gaussian model is a conjugate prior model with respect to Gauss-linear likelihood models. However, evaluation of the posterior distribution is computationally demanding and relies on McMC algorithms. An alternative approach relies on a S-GRF prior model. The concept of selection probability distributions (Arellano-Valle et al., 2006) has been developed and extended to spatial settings (Allard and Naveau, 2007; Omre and Rimstad, 2021), and has been applied to seismic inversion (Karimi et al., 2010; Rimstad and Omre, 2014; Forberg et al., 2021b). This approach is further explored in the current study.

We extend the study in Forberg et al. (2021b) to invert a 2D section and a 3D volume of a reservoir and use data from the same oil and gas discovery in the North Sea. However, whereas porosity and water saturation could provide adequate reservoir characterization in Forberg et al. (2021b), the chosen inversion zones in the current study are more lithologically heterogeneous and contain a mix of sand and shale. Therefore, we are also interested in characterizing the clay content in the inversion zones. We operate in a Bayesian framework. The forward model consists of an empirically approximated rock physics model combined with the linearized seismic AVO formulation used in (Buland and Omre, 2003).

The seismic model is a wavelet convolution of the linear Aki and Richards approximation (Aki and Richards, 1980) of the Zoeppritz equations (Zoeppritz, 1919). From this model the seismic AVO responses are predicted from the elastic properties P-wave velocity, S-wave velocity, and density. The rock physics model is approximated by multiple linear regressions of the logarithmic elastic properties on porosity, water saturation, and volume of clay (Landrø, 2001; Grana et al., 2017; Forberg et al., 2021b). We associate Gaussian error terms with the forward function; hence the likelihood model is Gauss-linear. The prior model is an S-GRF, which provides support for multimodality and is conjugate with respect to the Gauss-linear likelihood model (Omre and Rimstad, 2021). Moreover, the S-GRF prior model is laterally coupled with support for volume of clay. To facilitate the extensions made in the current study, a feasible sampling scheme for higher dimensions, which honors lateral coupling and supports different anisotropy for porosity and water saturation, is developed. The methodology is demonstrated on real 2D and 3D data from the Kneler field in the Alvheim oil and gas field in the North Sea.

Definitions and Notation

The reservoir zone under study is of spatial dimension m and it is discretized into the reservoir grid $\mathcal{G}_r$, consisting of n_r grid points. The petrophysical properties of interest in the reservoir zone are referred to as the reservoir properties. In the current study the reservoir properties are porosity, water saturation, and volume of clay, which are contained in the n_r-dimensional vectors $\boldsymbol{\phi}$, $\mathbf{s}_w$, and $\mathbf{v}_c$, respectively. Hence, the reservoir properties are represented by the $3n_r$-dimensional vector $\mathbf{r} = [\boldsymbol{\phi}, \mathbf{s}_w, \mathbf{v}_c]$. The following shorthand notation is useful

$$\mathbf{r}_k = \begin{cases} \boldsymbol{\phi}, & \text{for } k = 1, \\ \mathbf{s}_w, & \text{for } k = 2, \\ \mathbf{v}_c, & \text{for } k = 3. \end{cases} \tag{1}$$

Seismic AVO data informative about the reservoir properties is available on the seismic grid $\mathcal{G}_d$, consisting of n_d grid points. This data is represented by the $n_\theta n_d$-dimensional vector $\mathbf{d}$, where n_θ is the number of offset angles. The elastic properties on logarithmic form is contained in the $3n_r$-dimensional vector $\mathbf{m} = [log(\mathbf{v}_p), log(\mathbf{v}_s), log(\boldsymbol{\rho})]$, where the three n_r-dimensional vectors $\mathbf{v}_p$, $\mathbf{v}_s$, and $\boldsymbol{\rho}$ contain the P-wave velocities, S-wave velocities, and densities, respectively. All vectors, defined and yet to be defined, are column vectors unless otherwise stated. An $(n_1 \times n_2)$ matrix is a matrix with n_1 rows and n_2 columns, and $\mathbf{I}_n$ is the $(n \times n)$ identity matrix. The superscript T indicates the transpose.

The reservoir properties are assumed to be stochastic; hence, $\mathbf{r}$ is a random vector, implying the existence of an associated probability density function (pdf) and a corresponding cumulative density function (cdf), denoted by $p(\mathbf{r})$ and $P(\mathbf{r})$, respectively. Moreover, since each element of $\mathbf{r}$ is considered a random variable associated with a specific location in the reservoir grid $\mathcal{G}_r$, $\mathbf{r}$ is a so-called discretized random field (RF).

The RF $\mathbf{r}$ is a discretized GRF if its pdf is Gaussian, i.e., of the form

$$p(\mathbf{r}) = (2\pi)^{-\frac{3n_r}{2}} |\boldsymbol{\Sigma}_r|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2} (\mathbf{r} - \boldsymbol{\mu}_r)^T \boldsymbol{\Sigma}_r^{-1} (\mathbf{r} - \boldsymbol{\mu}_r) \right\}, \tag{2}$$

where $\boldsymbol{\mu}_r$ is the expectation vector and $\boldsymbol{\Sigma}_r$ is the covariance matrix. This pdf is denoted by $\varphi_{3n_r}(\mathbf{r}; \boldsymbol{\mu}_r, \boldsymbol{\Sigma}_r)$. The probability that $\mathbf{r}$ belongs to a subset $\mathbf{Q}$ of the $3n_r$-dimensional space of real numbers is denoted by $\Phi_{3n_r}(\mathbf{Q}; \boldsymbol{\mu}_r, \boldsymbol{\Sigma}_r)$ and given by the integral

$$\begin{aligned} \Phi_{3n_r}(\mathbf{Q}; \boldsymbol{\mu}_r, \boldsymbol{\Sigma}_r) &= \text{Prob}(\mathbf{r} \in \mathbf{Q}) \\ &= \int_{\mathbb{R}^{3n_r}} I(\mathbf{r} \in \mathbf{Q}) \varphi_{3n_r}(\mathbf{r}; \boldsymbol{\mu}_r, \boldsymbol{\Sigma}_r) \ \mathrm{d}\mathbf{r}, \end{aligned} \tag{3}$$

where $I(\cdot)$ is the indicator function, which is equal to 1 if the argument is true and equal to 0 otherwise.

The RF $\mathbf{r}$ is a discretized S-GRF if its pdf is selection Gaussian, i.e., of the form demonstrated in (Arellano-Valle et al., 2006; Omre and Rimstad, 2021),

$$\begin{aligned} p(\mathbf{r}) &= p(\tilde{\mathbf{r}}|\boldsymbol{\nu} \in \mathbf{A}) = \frac{p(\boldsymbol{\nu} \in \mathbf{A}|\tilde{\mathbf{r}})}{p(\boldsymbol{\nu} \in \mathbf{A})} p(\tilde{\mathbf{r}}) \\ &= \frac{\Phi_{n_\nu}\left(\mathbf{A}; \boldsymbol{\mu}_{\nu|\tilde{r}}, \boldsymbol{\Sigma}_{\nu|\tilde{r}}\right) \varphi_{3n_r}\left(\tilde{\mathbf{r}}; \boldsymbol{\mu}_{\tilde{r}}, \boldsymbol{\Sigma}_{\tilde{r}}\right)}{\Phi_{n_\nu}\left(\mathbf{A}; \boldsymbol{\mu}_{\nu}, \boldsymbol{\Sigma}_{\nu}\right)}, \end{aligned} \tag{4}$$

where $\tilde{\mathbf{r}}$ is the $3n_r$-dimensional basis variable, $\boldsymbol{\nu}$ is the n_ν-dimensional auxiliary variable, and $\mathbf{A}$ is the n_ν-dimensional selection set, see Forberg et al. (2021b) for details. The conditional parameters involved in the pdf can be computed as

$$\begin{aligned} \boldsymbol{\mu}_{\nu|\tilde{r}} &= \boldsymbol{\mu}_\nu + \boldsymbol{\Gamma}_{\tilde{r}\nu} \boldsymbol{\Sigma}_{\tilde{r}}^{-1} \left(\tilde{\mathbf{r}} - \boldsymbol{\mu}_{\tilde{r}}\right), \\ \boldsymbol{\Sigma}_{\nu|\tilde{r}} &= \boldsymbol{\Sigma}_\nu - \boldsymbol{\Gamma}_{\tilde{r}\nu} \boldsymbol{\Sigma}_{\tilde{r}}^{-1} \boldsymbol{\Gamma}_{\nu\tilde{r}}, \end{aligned} \tag{5}$$

where the $(3n_r \times n_\nu)$ matrix $\boldsymbol{\Gamma}_{\tilde{r}\nu}$ contains the covariances between $\tilde{\mathbf{r}}$ and $\boldsymbol{\nu}$.

The RF $\mathbf{r}$ is said to be stationary if the pdf of any subset of random variables in $\mathbf{r}$ is shift invariant; hence, the pdf must depend only on the distances between the selected random variables in the reservoir grid and not on their specific locations. For the GRF $\mathbf{r}$, this holds if the locationwise expectation and variance are constant, and the correlation structure is defined through shift invariant and positive definite correlation functions. This entails that the expectation vector and covariance matrix has the following form

$$\boldsymbol{\mu}_{r_k} = \mu_{r_k} \mathbf{i}_{n_r}, \quad \text{and} \quad \boldsymbol{\Sigma}_{r_k} = \sigma_{r_k}^2 \boldsymbol{\Omega}_{r_k}; \quad k = 1, 2, 3 \tag{6}$$

where μ_{r_k} and $\sigma_{r_k}^2$ are the locationwise expectation and variance, respectively, and $\boldsymbol{\Omega}_{r_k}$ is the spatial correlation matrix. This spatial correlation matrix is defined through the shift invariant and positive definite correlation function $\rho_{r_k}(\boldsymbol{\tau})$, with $\boldsymbol{\tau}$ being the distance between two grid points. Lastly, $\mathbf{i}_{n_r}$ is the n_r-dimensional vector of ones. The S-GRF $\mathbf{r}$ is said to be stationary if the criteria for a stationary GRF hold and the selection set $\mathbf{A}$ is identical in all locations, that is, if it can be expressed as the cartesian product $\mathbf{A} = A^{n_\nu}$, with A a subset of the real numbers.

Methodology

The prediction of the reservoir properties $\mathbf{r}$ from the seismic AVO data $\mathbf{d}$ is approached in a Bayesian framework. The solution to the inverse problem is the posterior pdf $p(\mathbf{r}|\mathbf{d})$, defined by Bayes' rule

$$p(\mathbf{r}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{r})p(\mathbf{r})}{p(\mathbf{d})} \propto p(\mathbf{d}|\mathbf{r})p(\mathbf{r}), \tag{7}$$

with $p(\mathbf{d}|\mathbf{r})$ and $p(\mathbf{r})$ the likelihood and prior models, respectively. The denominator $p(\mathbf{d})$ is the normalizing constant, which is generally hard to compute. The methodology is based on model assumptions for the likelihood and prior models that make analytical computation of $p(\mathbf{d})$ feasible.

Likelihood Model

The relationship between the seismic responses and the reservoir properties represents the seismic data acquisition procedure and is described by the likelihood model $p(\mathbf{d}|\mathbf{r})$. This model can be decomposed into a seismic likelihood model $p(\mathbf{d}|\mathbf{m})$ and a rock physics model $p(\mathbf{m}|\mathbf{r})$, see Forberg et al. (2021b) for details.

The seismic model relates the elastic properties to the seismic responses. In the current study this model is based on a convolution of the linear Aki and Richards (Aki and Richards, 1980) approximation to the Zoeppritz equations, and is of the form $[\mathbf{d}|\mathbf{m}] = \mathbf{WADm} + \mathbf{e}_{d|m} = \mathbf{W}\left(\mathbf{c} + \mathbf{e}_{c|m}\right) + \mathbf{e}_{d|c}$ (Buland and Omre, 2003), where the $n_\theta n_r$-dimensional vector $\mathbf{c}$ contains the computed reflectivity coefficients. The $(n_\theta n_r \times 3n_r)$ matrix $\mathbf{A}$ and $(3n_r \times 3n_r)$ matrix $\mathbf{D}$ represents the Aki and Richards approximation, with $\mathbf{A}$ containing angle dependent coefficients and $\mathbf{D}$ being a first order differential operator. Convolution is introduced through the $(n_\theta n_d \times n_\theta n_r)$ matrix $\mathbf{W}$, which contains discretizations of the seismic wavelet. Further, the $n_\theta n_d$-dimensional vector $\mathbf{e}_{d|m}$ is an error term assumed to be Gaussian, with expectation zero and $(n_\theta n_d \times n_\theta n_d)$ covariance matrix $\mathbf{\Sigma}_{d|m} = \mathbf{W}\sigma^2_{c|m}\mathbf{I}_{n_\theta n_r}\mathbf{W}^T + \sigma^2_{d|c}\mathbf{I}_{n_\theta n_d}$. The variance $\sigma^2_{c|m}$ is associated with model error from the Aki and Richards approximation, while the variance $\sigma^2_{d|c}$ is associated with observation error. Note that $\mathbf{e}_{d|m}$ contains both white and coloured noise since the model error is wavelet convolved. The seismic likelihood model is Gauss-linear and can be expressed as

$$p(\mathbf{d}|\mathbf{m}) = \varphi_{n_\theta n_d}\left(\mathbf{d}; \mathbf{WADm}, \mathbf{\Sigma}_{d|m}\right). \tag{8}$$

The rock physics model relates the reservoir properties to the elastic properties. As in previous studies (Forberg et al. (2021b), Forberg et al. (2021a)), the model is inspired by Landrø (2001) and formulated as a multiple linear regression, $[\mathbf{m}|\mathbf{r}] = \mathbf{B}\mathbf{r} + \mathbf{e}_{m|r}$. Regression coefficients from multiple linear regressions of the logarithmic elastic properties on the reservoir properties are contained in the $(3n_r \times 3n_r)$ matrix $\mathbf{B}$, while the $3n_r$-dimensional vector $\mathbf{e}_{m|r}$ represents the error associated with the model. The error term $\mathbf{e}_{m|r}$ is assumed to be Gaussian with expectation zero and $(3n_r \times 3n_r)$ covariance matrix $\mathbf{\Sigma}_{m|r}$. Hence, the rock physics likelihood model is Gauss-linear and can be expressed as

$$p(\mathbf{m}|\mathbf{r}) = \varphi_{3n_r}\left(\mathbf{m}; \mathbf{B}\mathbf{r}, \mathbf{\Sigma}_{m|r}\right). \tag{9}$$

The seismic and rock physics model applied in succession define the forward model, which takes the form $[\mathbf{d}|\mathbf{r}] = \mathbf{G}\mathbf{r} + \mathbf{e}_{d|r}$. Here, the $(n_\theta n_d \times 3n_r)$ matrix $\mathbf{G} = \mathbf{WADB}$ is the forward operator and the $n_\theta n_d$-dimensional vector $\mathbf{e}_{d|r} = \mathbf{WADe}_{m|r} + \mathbf{e}_{d|m}$ is an error term assumed to be Gaussian with expectation zero and $(n_\theta n_d \times n_\theta n_d)$ covariance matrix $\mathbf{\Sigma}_{d|r} = \mathbf{WAD}\mathbf{\Sigma}_{m|r}(\mathbf{WAD})^T + \mathbf{\Sigma}_{d|m}$. Hence, the likelihood model is Gauss-linear and can be expressed as

$$p(\mathbf{d}|\mathbf{r}) = \varphi_{n_\theta n_d}\left(\mathbf{d}; \mathbf{G}\mathbf{r}, \mathbf{\Sigma}_{d|r}\right). \tag{10}$$

The Gauss-linearity of the likelihood model entails that for certain parametric prior models the posterior model will be of the same form as the prior model (Forberg et al., 2021a).

We use the signal-to-noise ratio (SNR) as a measure of the magnitude of error in the seismic data. The SNR is defined as

$$\text{SNR} = \frac{\text{Trace}\left(\text{Var}\left(\mathbf{G}\mathbf{r}\right)\right)}{\text{Trace}\left(\text{Var}\left(\mathbf{d}|\mathbf{r}\right)\right)} = \frac{\text{Trace}\left(\mathbf{G}\mathbf{\Sigma}_r\mathbf{G}^T\right)}{\text{Trace}\left(\mathbf{\Sigma}_{d|r}\right)}, \tag{11}$$

with Trace$(\cdot)$ a function returning the sum of the diagonal elements of its matrix argument. The SNR is dependent on the prior model specification because $\mathbf{\Sigma}_r$ is involved in the expression.

Prior Model

The prior model $p(\mathbf{r})$ represents our understanding, experience, and beliefs about the reservoir properties to be predicted. We operate in a setting where porosity is assumed to be predominantly spatially smoothly varying and to have a locationwise skewed and unimodal pdf. Water saturation and volume of clay are assumed to exhibit abrupt spatial transitions and to have locationwise bimodal pdfs. We base the model construction on the decomposition

$$p(\tilde{\mathbf{r}}, \boldsymbol{\nu}) = p(\boldsymbol{\nu}|\tilde{\mathbf{r}})p(\tilde{\mathbf{r}}). \tag{12}$$

The basis variable $\tilde{\mathbf{r}}$ is assumed to be a stationary GRF,

$$p(\tilde{\mathbf{r}}) = \varphi_{3n_r}(\tilde{\mathbf{r}}; \boldsymbol{\mu}_{\tilde{r}}, \boldsymbol{\Sigma}_{\tilde{r}}), \tag{13}$$

with

$$\begin{aligned} \boldsymbol{\mu}_{\tilde{r}} &= \begin{bmatrix} \mu_{\tilde{\phi}}\mathbf{i}_{n_r} \\ \mu_{\tilde{s}_w}\mathbf{i}_{n_r} \\ \mu_{\tilde{v}_c}\mathbf{i}_{n_r} \end{bmatrix} \\ \boldsymbol{\Sigma}_{\tilde{r}} &= \begin{bmatrix} \sigma^2_{\tilde{\phi}}\boldsymbol{\Omega}_s & \lambda_1\sigma_{\tilde{\phi}}\sigma_{\tilde{s}_w}\boldsymbol{\Omega}_x & \lambda_2\sigma_{\tilde{\phi}}\sigma_{\tilde{v}_c}\boldsymbol{\Omega}_s \\ \lambda_1\sigma_{\tilde{\phi}}\sigma_{\tilde{s}_w}\boldsymbol{\Omega}_x & \sigma^2_{\tilde{s}_w}\boldsymbol{\Omega}_h & \lambda_3\sigma_{\tilde{s}_w}\sigma_{\tilde{v}_c}\boldsymbol{\Omega}_x \\ \lambda_2\sigma_{\tilde{\phi}}\sigma_{\tilde{v}_c}\boldsymbol{\Omega}_s & \lambda_3\sigma_{\tilde{s}_w}\sigma_{\tilde{v}_c}\boldsymbol{\Omega}_x & \sigma^2_{\tilde{v}_c}\boldsymbol{\Omega}_s \end{bmatrix} \end{aligned} \tag{14}$$

Here, $(\mu_{\tilde{\phi}}, \mu_{\tilde{s}_w}, \mu_{\tilde{v}_c})$ and $(\sigma^2_{\tilde{\phi}}, \sigma^2_{\tilde{s}_w}, \sigma^2_{\tilde{v}_c})$ are the locationwise expectations and variances of $\tilde{\boldsymbol{\phi}}$, $\tilde{\mathbf{s}}_w$, and $\tilde{\mathbf{v}}_c$, respectively. Moreover, $(\lambda_1, \lambda_2, \lambda_3)$ are the locationwise correlations between $\tilde{\phi}$ and $\tilde{s}_w$, between $\tilde{\phi}$ and $\tilde{v}_c$, and between $\tilde{s}_w$ and $\tilde{v}_c$, respectively. Lastly, $(\boldsymbol{\Omega}_s, \boldsymbol{\Omega}_h, \boldsymbol{\Omega}_x)$ are $(n_r \times n_r)$ correlation matrices containing structural spatial correlation, horizontal spatial correlation, and a mix between structural and horizontal spatial correlation, respectively. Different spatial correlation structures are needed because porosity and volume of clay are defined by previous sedimentation whereas water saturation is governed largely by current gravitation, which entails that their anisotropies are likely to differ. The correlation matrices are defined through shift invariant and positive definite correlation functions $\rho_{r_k}(\boldsymbol{\tau}; \boldsymbol{\alpha}_k)$. The correlation structure between water saturation and the other reservoir properties is defined to be $\boldsymbol{\Omega}_x = \frac{1}{2}\boldsymbol{\Omega}_s + \frac{1}{2}\boldsymbol{\Omega}_h$.

The auxiliary GRF $\boldsymbol{\nu}$ is defined on the reservoir grid $\mathcal{G}_r$ and we define one auxiliary variable for each basis variable, i.e., $\boldsymbol{\nu} = [\boldsymbol{\nu}_1, \boldsymbol{\nu}_2, \boldsymbol{\nu}_3]$. Hence, $n_\nu = 3n_r$. The conditional relation between the auxiliary variable $\boldsymbol{\nu}$ and

the basis variable $\tilde{\mathbf{r}}$ is

$$p(\boldsymbol{\nu}|\tilde{\mathbf{r}}) = \prod_{k=1}^{3} p(\boldsymbol{\nu}_k|\tilde{\mathbf{r}}_k) = \prod_{k=1}^{3} \varphi_{n_r}(\boldsymbol{\nu}_k; \boldsymbol{\mu}_{\nu_k|\tilde{r}_k}, \boldsymbol{\Sigma}_{\nu_k|\tilde{r}_k}) \tag{15}$$
$$= \prod_{k=1}^{3} \left(\prod_{i=1}^{n_r} \varphi_1(\nu_{k,i}; \mu_{\nu_{k,i}|\tilde{r}_k}, (1-\gamma_k^2)) \right).$$

Here, γ_k is the locationwise correlation between $\tilde{r}_k$ and ν_k, and the conditional expectation is $\boldsymbol{\mu}_{\nu_k|\tilde{r}_k} = 0\mathbf{i}_{n_r} + \boldsymbol{\Gamma}_{\tilde{r}_k\nu_k}\boldsymbol{\Sigma}_{\tilde{r}_k}^{-1}\left(\tilde{\mathbf{r}}_k - \boldsymbol{\mu}_{\tilde{r}_k}\right)$, $k = 1, 2, 3$. The $(n_r \times n_r)$ cross covariance matrix $\boldsymbol{\Gamma}_{\tilde{r}_k\nu_k}$ contains the covariances between $\tilde{\mathbf{r}}_k$ and $\boldsymbol{\nu}_k$ and can be found on the diagonal of the $(3n_r \times 3n_r)$ cross covariance matrix $\boldsymbol{\Gamma}_{\tilde{r}\nu}$,

$$\boldsymbol{\Gamma}_{\tilde{r}\nu} = \begin{bmatrix} \gamma_1\sigma_{\tilde{\phi}}\boldsymbol{\Omega}_s & \lambda_1\gamma_2\sigma_{\tilde{\phi}}\boldsymbol{\Omega}_x & \lambda_2\gamma_3\sigma_{\tilde{\phi}}\boldsymbol{\Omega}_s \\ \lambda_1\gamma_1\sigma_{\tilde{s}_w}\boldsymbol{\Omega}_x & \gamma_2\sigma_{\tilde{s}_w}\boldsymbol{\Omega}_h & \lambda_3\gamma_3\sigma_{\tilde{s}_w}\boldsymbol{\Omega}_x \\ \lambda_2\gamma_1\sigma_{\tilde{v}_c}\boldsymbol{\Omega}_s & \lambda_3\gamma_2\sigma_{\tilde{v}_c}\boldsymbol{\Omega}_x & \gamma_3\sigma_{\tilde{v}_c}\boldsymbol{\Omega}_s \end{bmatrix}. \tag{16}$$

Note that the expectation $\boldsymbol{\mu}_{\nu_k|\tilde{r}_k}$ is linear in $\tilde{\mathbf{r}}$ and that the conditional pdf $p(\boldsymbol{\nu}|\tilde{\mathbf{r}})$ in Equation 15 is Gaussian; hence, the auxiliary GRF is defined to be Gauss-linearly related to the basis GRF, which entails that their joint pdf $p(\tilde{\mathbf{r}}, \boldsymbol{\nu})$ is Gaussian. Moreover, locationwise conditional independence for $[\boldsymbol{\nu}|\tilde{\mathbf{r}}]$ is assumed.

The $(3n_r \times 3n_r)$ covariance matrix associated with the auxiliary GRF $\boldsymbol{\nu}$ is

$$\boldsymbol{\Sigma}_\nu = \begin{bmatrix} \boldsymbol{\Sigma}_{\nu_1} & \lambda_1\gamma_1\gamma_2\boldsymbol{\Omega}_x & \lambda_2\gamma_1\gamma_3\boldsymbol{\Omega}_s \\ \lambda_1\gamma_1\gamma_2\boldsymbol{\Omega}_x & \boldsymbol{\Sigma}_{\nu_2} & \lambda_3\gamma_2\gamma_3\boldsymbol{\Omega}_x \\ \lambda_2\gamma_1\gamma_3\boldsymbol{\Omega}_s & \lambda_3\gamma_2\gamma_3\boldsymbol{\Omega}_x & \boldsymbol{\Sigma}_{\nu_3} \end{bmatrix}, \tag{17}$$

where

$$\boldsymbol{\Sigma}_{\nu_k} = \begin{cases} \gamma_1^2\boldsymbol{\Omega}_s + (1-\gamma_1^2)\mathbf{I}_{n_r}, & k = 1, \\ \gamma_2^2\boldsymbol{\Omega}_h + (1-\gamma_2^2)\mathbf{I}_{n_r}, & k = 2, \\ \gamma_3^2\boldsymbol{\Omega}_s + (1-\gamma_3^2)\mathbf{I}_{n_r}, & k = 3. \end{cases} \tag{18}$$

Lastly, the $3n_r$-dimensional selection set $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3]$ consists of the selection sets $\mathbf{A}_k$ corresponding to $\boldsymbol{\nu}_k$, $k = 1, 2, 3$. These selection sets are of the form $\mathbf{A}_k = A_k^{n_r}$, where $A_k = \bigcup_{i=1}^{n_{A_k}} [a_{k,i}, b_{k,i}]$ are subsets of the real numbers, with n_{A_k} being the number of disjoint subsets that selection set A_k consists of. The prior model for the reservoir variables, $p(\mathbf{r})$, is the trivariate discretized stationary S-GRF

$$p(\mathbf{r}) = p(\tilde{\mathbf{r}}|\boldsymbol{\nu} \in \mathbf{A}) \tag{19}$$
$$= \frac{\prod_{k=1}^{3}\prod_{i=1}^{n_r}\varphi_1(A_k;\mu_{\nu_{k,i}|\tilde{r}_k},(1-\gamma_k^2))}{\Phi_{3n_r}(\mathbf{A};0\mathbf{i}_{3n_r},\boldsymbol{\Sigma}_\nu)}\varphi_{3n_r}(\tilde{\mathbf{r}};\boldsymbol{\mu}_{\tilde{r}},\boldsymbol{\Sigma}_{\tilde{r}}).$$

This prior model is parameterized by $\boldsymbol{\Theta}_p^{SG} = \left\{[\mu_{\tilde{r}_k},\sigma_{\tilde{r}_k}^2,\lambda_k,\gamma_k,A_k,\rho_{r_k}(\cdot)];\quad k=1,2,3\right\}$, where the first five parameters for each k are primarily related to the locationwise trivariate selection Gaussian pdf, whereas the last parameter for each k primarily relates to the spatial correlation structure.

The prior model $p(\mathbf{r})$ has support for values outside the physical range $[0,1]$ of the reservoir properties. In line with Forberg et al. (2021b), we correct for this effect in the predictor and prediction intervals, which is described in the next section.

The spatial correlation functions are for pairs of grid points defined as

$$\rho_{r_k}(\boldsymbol{\tau};\boldsymbol{\alpha}_k) = \exp\left\{-\frac{\tau_v}{\alpha_{v,k}}\right\} \times \exp\left\{-\frac{\tau_h}{\alpha_{h,k}}\right\};\quad k=1,2,3, \tag{20}$$

where $\boldsymbol{\tau} = [\tau_v,\tau_h]$ are the grid distances, with $\tau_v \geq 0$ being the vertical distance and $\tau_h \geq 0$ being the lateral distance between the grid points. In $m=3$ spatial dimensions, τ_h is the cartesian distance; hence, the lateral correlation is isotropic. Lastly, $\boldsymbol{\alpha}_k = [\alpha_{v,k},\alpha_{h,k}]$ are the range parameters, being positive. These correlation functions define the spatial correlation matrices $\boldsymbol{\Omega}_s$ and $\boldsymbol{\Omega}_h$. Different anisotropies can be represented by transformation of the reservoir grid, which will be explained later.

Posterior Model

Because the S-GRF prior model $p(\mathbf{r})$ is conjugate with respect to the Gauss-linear likelihood model $p(\mathbf{d}|\mathbf{r})$, the posterior model $p(\mathbf{r}|\mathbf{d})$ is also an S-GRF, see Omre and Rimstad (2021),

$$p(\mathbf{r}|\mathbf{d}) = p(\tilde{\mathbf{r}}|\boldsymbol{\nu} \in \mathbf{A},\mathbf{d}) \tag{21}$$
$$= \frac{\prod_{k=1}^{3}\Phi_{n_r}\left(\mathbf{A}_k;\boldsymbol{\mu}_{\nu_k|\tilde{r}_k,d},\boldsymbol{\Sigma}_{\nu_k|\tilde{r}_k,d}\right)}{\Phi_{3n_r}\left(\mathbf{A};\boldsymbol{\mu}_{\nu|d},\boldsymbol{\Sigma}_{\nu|d}\right)} \times \varphi_{3n_r}\left(\tilde{\mathbf{r}};\boldsymbol{\mu}_{\tilde{r}|d},\boldsymbol{\Sigma}_{\tilde{r}|d}\right).$$

The conditional parameters of this posterior model are developed in Forberg et al. (2021b). We use a block-independent Metropolis-Hastings

(M-H) McMC algorithm (Rimstad and Omre, 2014) to simulate from the posterior model. These simulations form the basis for prediction and uncertainty assessment. Simulation is done either trace-by-trace or by conditioning on neighboring traces, depending on whether lateral correlation is present or not. The McMC algorithms are presented in Appendix A.

The reservoir properties are predicted by the marginal maximum posterior (MMAP). This predictor identifies the values of the reservoir properties with the highest marginal posterior probability density at each grid point in $\mathcal{G}_r$, and is defined as

$$\hat{\mathbf{r}}_{\text{MMAP}} = \left\{ \underset{r_{k,i}}{\text{argmax}}\{p(r_{k,i}|\mathbf{d})\}; \quad k = 1,2,3, \ i = 1,...,n_r \right\}. \tag{22}$$

Prediction intervals in the form of $(1-\alpha)\times 100\%$ highest posterior density intervals (HPDI) are used to reflect the variability in the locationwise posterior distributions (Forberg et al., 2021b). These intervals may consist of a collection of disjoint regions, thereby reflecting multimodality in the posterior distribution. The MMAP is located within one of the HPDI regions. The coverage level $(1-\alpha)\times 100\%$ is user specified and should be assigned based on considerations about the application and the shape of the distribution. Since the prior model can support non-physical values, so can the MMAP predictions and HPDI. If they exceed $[0,1]$, they are truncated to the appropriate limits of this range.

The MMAP prediction and accompanying $(1-\alpha)\times 100\%$ HPDI can be quantitatively evaluated if observations of the reservoir properties, $\mathbf{r}_{\text{obs}}$, are available in some locations. The root-mean-square error (RMSE) is a measure of predictor deviation from the observations; hence, the lower the value, the better. The prediction intervals are evaluated by empirical coverage, i.e., by computing the percentage of observations within the prediction intervals. Empirical coverage close to the chosen coverage level of the prediction intervals is favourable.

Case: Real seismic AVO data from a 2D section

We apply the proposed Bayesian seismic inversion methodology to a 2D section of the Kneler field in the Alvheim oil and gas field, in the North Sea. The turbidite Alvheim reservoir is challenging to characterize due to complex geological depositions. The oil reservoir under study mostly consists of porous sandstone and is underlain by an aquifer, but its upper part is composed of unconsolidated and interbedded sand and shale. The chosen 2D section, with the location of two nearby exploration wells indicated, is displayed in Figure 1. Note that the true aspect ratio of the 2D section is altered for visualization purposes. The section has a much bigger lateral extent than what it appears to; hence, the apparent curvature in the section is much less severe. The 2D section follows the interpreted top reservoir, ranging in depth from 1972 ms to 2052 ms in two-way time, for 4600 m along a cross line. Vertically, the section starts two seismic sample points above the interpreted top horizon, presumably in a top shale layer, and ends in the underlying aquifer. The 2D section contains 93 traces with a regular inter-trace spacing of 50 m, each of which covers a depth range of 52 ms in two-way time. The reservoir grid $\mathcal{G}_r$ consists of $n_r = 2511$ grid points which are regularly spaced laterally and with trace-unique vertical gridding, while the seismic grid $\mathcal{G}_d$ consists of $n_d = 1302$ grid points with similar grid structure. The trace-unique vertical gridding consists of 27 points in the reservoir grid and 14 points in the seismic grid. The data and inversion results are linearly interpolated to a finer and regular grid for visualization. This artifical increase in resolution produces somewhat smoother figures, but all computations are performed on the irregular grids $\mathcal{G}_r$ and $\mathcal{G}_d$.

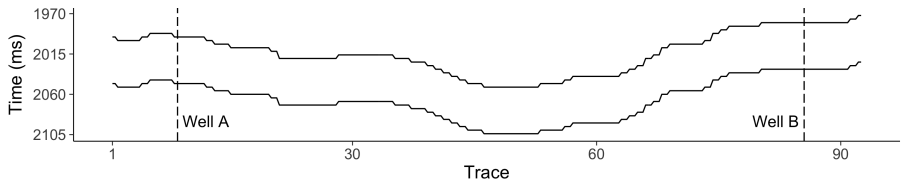

Figure 1: Geometry of the 2D section with wells marked.

The exploration wells can be used for model construction and validation. Well B is in immediate vicinity of the 2D section, whereas Well A is located a few cross lines away. Well B is the exploration well studied in Forberg et al. (2021b), and its location relative to the 2D section makes it the preferred basis for model construction. Measured well logs

from Well B are displayed in Figure 2. The top reservoir is at about 1990 ms, and the underlying aquifer starts at 2020 ms, as reflected by a clear oil-water contact (OWC) in the water saturation log. The characteristics of the logs above 1990 ms reflect an overlying shale layer. An increased variability in the logs can be seen between roughly 1990 ms and 2004 ms. This variability does not appear to be consistent with variations in the volume of clay alone; hence, there may be a notable contribution from other minerals in this region. Below 2004 ms, a more homogeneously cemented porous sandstone can be seen. Furthermore, note that the elastic properties reflect two anomalous regions, which are located at roughly 1998 ms and 2014 ms. These anomalies, and the above mentioned increased variability between 1990 ms and 2004 ms, are likely caused by intermittently occurring limestone, which has a small presence throughout the 2D section. Well A is used exclusively for validation of inversion results. The measured porosity, water saturation, and volume of clay along Well A are displayed in Figure 3. The oil reservoir zone appears to extend from the top of the depth interval to about 2030 ms where the underlying aquifer starts. The OWC is not as clearly reflected as in Well B, since the oil-water transition seems to occur in a region that has relatively large variability in the volume of clay. Moreover, the well logs indicate specks of limestone around 2005 ms, 2010 ms, and 2022 ms in the oil reservoir zone and reflect much variation in the mineral fractions above 2005 ms.

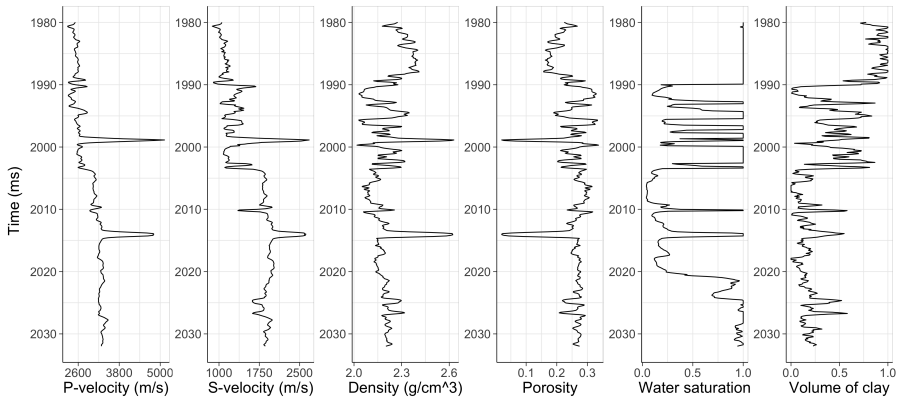

Figure 2: Well B logs, displayed column-wise from left to right: P-wave velocity, S-wave velocity, density, porosity, water saturation, and volume of clay.

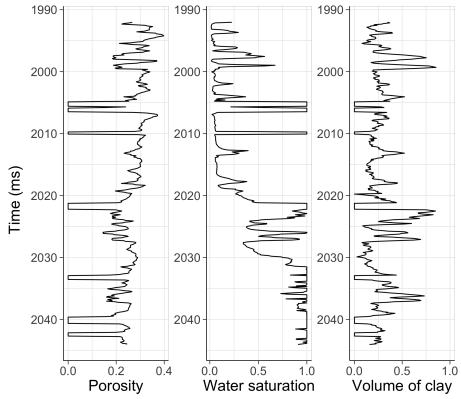

Figure 3: Well A logs, displayed column-wise from left to right: Porosity, water saturation, and volume of clay.

The seismic AVO data is presented in Figure 4, where the $n_\theta = 4$ angle stacks are displayed row-wise. The signal is relatively strong roughly between traces 20 and 35, especially in the 27° and 35° stacks, which in addition to fluid transitions may indicate a more abrupt change in lithological properties than elsewhere. The signal in the middle part of the 2D section, roughly between traces 40 and 65, is relatively weak in the 35° stack, indicating an absence of fluid transitions. Moreover, this region is located deeper than either side of it, which may indicate a higher content of more compressible lithologies here than elsewhere in the reservoir, for example a higher fractional content of shale. We hereafter refer to this region as Center Region (CR).

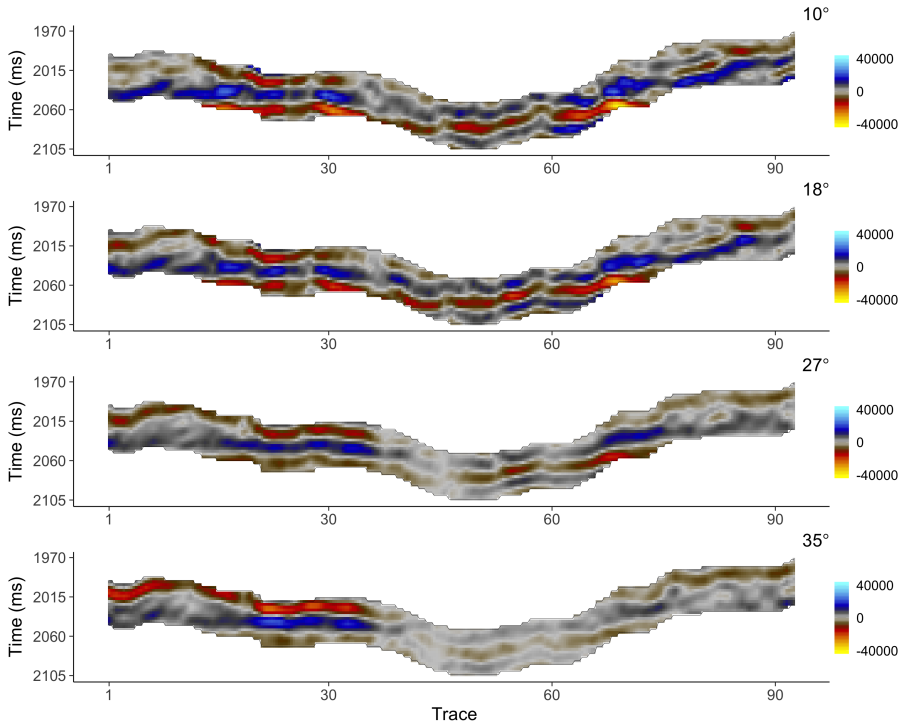

Figure 4: Seismic AVO data from the 2D section under study. The angle stacks $10°, 18°, 27°$, and $35°$ are displayed in order from the top to bottom row.

Model Construction

Model construction entails assigning parameter values to the likelihood model $p(\mathbf{d}|\mathbf{r})$ and prior model $p(\mathbf{r})$. Because the construction is based on well logs from Well B, the model may not be adequately representative in the entire 2D section. The effects of lateral correlation in the prior model on the inversion results are highlighted by defining an alternative prior model without lateral correlation, as in Forberg et al. (2021a), for comparison. We refer to these prior models as prior models 1 and 2, corresponding to the prior model with and without lateral correlation, respectively. The inversions are carried out using SNR $= 5$.

Likelihood model

The likelihood model consists of a seismic likelihood $p(\mathbf{d}|\mathbf{m})$ and a rock physics likelihood $p(\mathbf{m}|\mathbf{r})$. The matrices $\mathbf{A}$ and $\mathbf{D}$ in the seismic likelihood

model are known, while the convolution matrix $\mathbf{W}$ requires estimation. The wavelet function defining $\mathbf{W}$ is chosen to be the same as in Forberg et al. (2021b); hence, the wavelet is estimated by regressing the seismic data on synthetic reflection coefficients generated from observed elastic properties by $\mathbf{c}_{\text{mod}} = \mathbf{ADm}_{\text{obs}}$. The estimation is based on data from a depth interval of 300 m containing the reservoir zone and should be representative in the current study. The estimated wavelet is displayed in Figure 5.

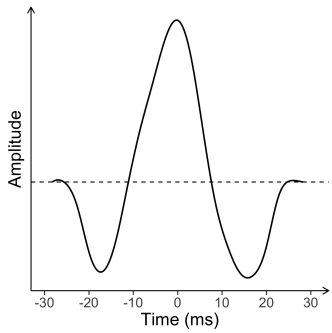

Figure 5: Estimated wavelet.

The rock physics likelihood model requires specification of the regression matrix $\mathbf{B}$ and the covariance matrix $\mathbf{\Sigma}_{m|r}$. The model is based on well logs from Well B after removal of the possible limestone influences. The three regression models all have an associated R^2 value of at least 0.9. In Table 1, the numerical values of the estimated regression coefficients and standard deviations associated with the Gauss-linear rock physics likelihood model are displayed. Moreover, the estimated inter-covariances between the elastic properties are $\hat{\xi}_{v_p v_s}$, $\hat{\xi}_{v_p \rho}$, and $\hat{\xi}_{v_s \rho}$, with corresponding correlations 0.656, 0.319, and 0.446. The performance of the rock physics model is represented graphically in Figure 6, where the measured and rock physics predicted logarithmic elastic properties are displayed together. Note that the rock physics model performs relatively poorly in the possibly limestone influenced interval between roughly 1900 ms and 2004 ms. The covariance matrix $\mathbf{\Sigma}_{m|r}$ has the variances associated with the regressions on the diagonal and the inter-covariances $\hat{\xi}_{v_p v_s}$, $\hat{\xi}_{v_p \rho}$, and $\hat{\xi}_{v_s \rho}$ on the off-diagonal.

Table 1: Estimated parameters in the rock physics likelihood model. The estimated intercept $\hat{\alpha}$, the slopes $(\hat{\beta}_\phi, \hat{\beta}_{s_w}, \hat{\beta}_{v_c})$ of porosity, water saturation, and volume of clay, respectively, are displayed together with the estimated standard deviation $\hat{\sigma}$ for each regression model in the top table. Each row of the table contains the parameters associated with the model for the elastic property specified in the leftmost column. The estimated cross-covariances between the elastic properties, $\hat{\xi}_{v_p v_s}$, $\hat{\xi}_{v_p \rho}$, and $\hat{\xi}_{v_s \rho}$, are shown in the lower table.

	$\hat{\alpha}$	$\hat{\beta}_\phi$	$\hat{\beta}_{s_w}$	$\hat{\beta}_{v_c}$	$\hat{\sigma}$
$\log(v_p)$	8.528	-1.567	0.085	-0.503	0.034
$\log(v_s)$	8.456	-2.892	0	-0.975	0.066
$\log(\rho)$	1.006	-0.910	0.035	-0.017	0.003

$\hat{\xi}_{v_p v_s}$	$\hat{\xi}_{v_p \rho}$	$\hat{\xi}_{v_s \rho}$
$1.50 \cdot 10^{-3}$	$3.15 \cdot 10^{-5}$	$8.50 \cdot 10^{-5}$

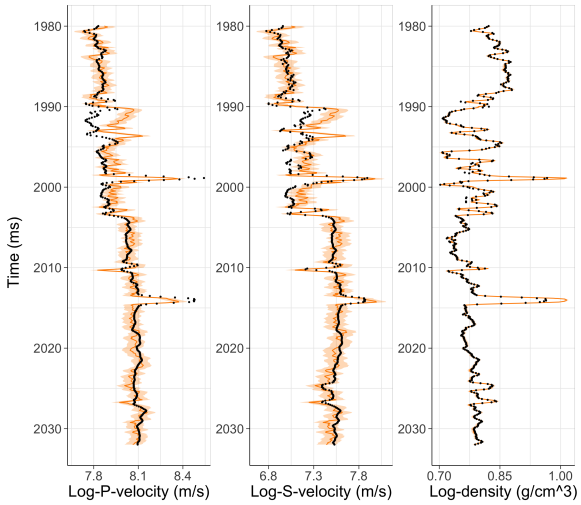

Figure 6: Predicted and measured logarithmic elastic properties, displayed column-wise from left to right: P-wave velocity, S-wave velocity, and density. The logarithm of the measured elastic properties (points) and the rock physics predicted reservoir properties (orange solid lines) with associated 90% prediction intervals (transparent orange regions) are displayed together.

Prior model

The parameter values used in prior model 1 are obtained by the heuristic parameter estimation approach described in Forberg et al. (2021b), and are listed in Table 2. Determining an appropriate value for the range parameter associated with lateral correlation is particularly difficult due to insufficient well data. The 2D section is assumed to be lithologically varying laterally because of different mineral fractions in CR. We therefore take a conservative approach and define a lateral correlation that in practice establishes lateral dependency only on the two nearest traces in each direction. It is also necessary to specify the anisotropies of porosity and water saturation. Because the 2D section follows the interpreted horizon of the reservoir, we assume lithological layering to follow the geometry of the section. This lithological layering is described by structural correlation in porosity and volume of shale, represented by $\mathbf{\Omega}_s$, and it is achieved by computing the spatial correlations on a rectangular version of the irregular 2D section. The fluid content in the reservoir zone is superimposed on the lithology and obeys gravitational effects. Therefore, the correlation structure of water saturation, represented by $\mathbf{\Omega}_h$, is defined to favour horizontal fluid contacts, which is achieved by computing the spatial correlations on the irregular 2D section.

Table 2: Parameter values used in the prior S-GRF. The parameters associated with the basis GRF model are shown in the top; the locationwise expectations $\mu_{\tilde{r}_k}$ in the leftmost block, the locationwise variances $\sigma^2_{\tilde{r}_k}$ in the middle block, and the inter-variable correlations λ_k to the right, $k = 1, 2, 3$. The parameters associated with the auxiliary GRF are shown in the middle; the marginal selection sets A_k to the left and the correlations with the basis GRF γ_k to the right. The range parameters associated with the spatial correlation structures, $\alpha_{k,i}$ for $k = 1, 2, 3$ and $i = 1, 2$, are listed in the bottom.

$\mu_{\tilde{\phi}}$	0.3	$\sigma^2_{\tilde{\phi}}$	$2.420 \cdot 10^{-3}$	λ_1	-0.25
$\mu_{\tilde{s}_w}$	0.5	$\sigma^2_{\tilde{s}_w}$	$6.561 \cdot 10^{-2}$	λ_2	-0.7
$\mu_{\tilde{v}_c}$	0.15	$\sigma^2_{\tilde{v}_c}$	$2.000 \cdot 10^{-1}$	λ_3	0.5

A_1	$(-\infty, -0]$	γ_1	0.9
A_2	$(-\infty, -1.2] \cup [1.45, \infty)$	γ_2	0.95
A_3	$[-0.15, 0.15] \cup [1.6, 2.2]$	γ_3	0.95

$\alpha_{v,1}$	1	$\alpha_{v,2}$	1	$\alpha_{v,3}$	1
$\alpha_{h,1}$	25	$\alpha_{h,2}$	25	$\alpha_{h,3}$	25

The parameter values used in prior model 2 are displayed in Table 2, except for the lateral correlation which is assigned $\alpha_{h,1} = \alpha_{h,2} = \alpha_{h,3} = 0$, that is, no lateral correlation.

The locationwise prior models for porosity, water saturation, and volume of clay, representing the marginal pdfs in the RF $\mathbf{r}$, are displayed in Figure 7. The prior models are superimposed on histograms of well log data, after outlier removal. The locationwise prior model for porosity is unimodal and notably skewed in sandy regions, whereas the locationwise prior models for water saturation and volume of clay are distinctly bimodal. The distinct bimodality in the water saturation prior model reflects the gravitational effects assumed to be present which cause abrupt spatial fluid transitions. Volume of clay has a notable effect on the prior models for porosity and water saturation. We see that the clay effect on the prior model for porosity is a shift of location and an inverse relationship between skewness and volume of clay, whereas the clay effect on the prior model for water saturation is the alteration of the oil and water probabilities, with high volume of clay corresponding to high probability of water.

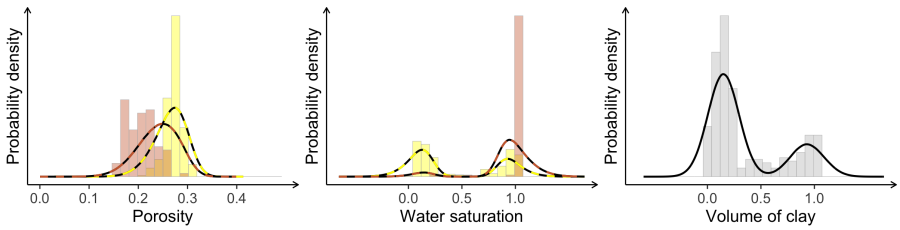

Figure 7: Locationwise prior models superimposed on histograms of well logs from Well B. The histograms of porosity and water saturation are colored yellow and brown, corresponding to data from sandy and presumably shaly regions, respectively. This classification is based on the bimodal distribution assigned to volume of clay. Moreover, because the prior models for porosity and water saturation are relatively strongly influenced by the volume of clay, a sandy prior model is displayed in hatched yellow and a shaly prior model is displayed in hatched brown. These models are based on the expected clay volume within each clay mode.

Results

We refer to the posterior models corresponding to prior models 1 and 2 simply as models 1 and 2, respectively. The convergence of the simulation based model assessments is ensured by running several M-H McMCs from random initial states until the assessed models can not be distinguished, and by evaluation of locationwise convergence plots Forberg et al. (2021b).

The MMAP predictions from models 1 and 2 are displayed in Figures 8 and 9, respectively. The predictions from model 1, see Figure 8, indicate two major high porosity oil zones between the overlying shale layer and the underlying aquifer, separated by CR where, as suspected, a considerable presence of high clay content is reflected in the predicted volumes of clay. Moreover, the geometry of the predicted oil zones can be seen to have a natural explanation in the predicted volumes of clay. A few specks of low water saturation are predicted outside the major oil zones, but these are most likely reflecting unsupported lithological hetereogeneity or noise in the seismic AVO data. The predicted water saturations exhibit strong continuity and smoothness, with distinctly defined regions of low values and high values. The predictions from model 2, see Figure 9, indicate high porosity oil zones on each side of CR, between the overlying shale layer and the underlying aquifer. The geometry of the major oil zones appear to have a natural explanation in the predicted volumes of clay. However, the major oil zone to the right of CR is not contained within the 2D section and appears to continue in the top horizon above CR. Moreover, several minor disconnected regions of low water saturation are predicted in CR, which based on the volume of clay predictions in CR, are overlaid by water. Despite a prior model without lateral correlation, the predictions clearly exhibit continuity, indicating strong continuity in the seismic AVO data. Both models predict the top of the reservoir zone to primarily have high water saturation and high clay content, which strongly indicates the presence of an overlying shale layer. The porosity predictions from the two models are very similar, but the predicted water saturations and volumes of clay differ notably in CR. The predictions from model 1 indicate that this region does not contain oil, whereas the predictions from model 2 indicate several irregular and disconnected regions of low water saturation and unphysical fluid transitions. Based on the seismic data in Figure 4, we do not expect oil in CR. The predicted volumes of clay from model 1 indicate a dominant presence of high clay content in CR, whereas model 2 indicates more of a mix between sandy regions and shaly regions. The clay volume predictions from model 1 seem geologically more plausible.

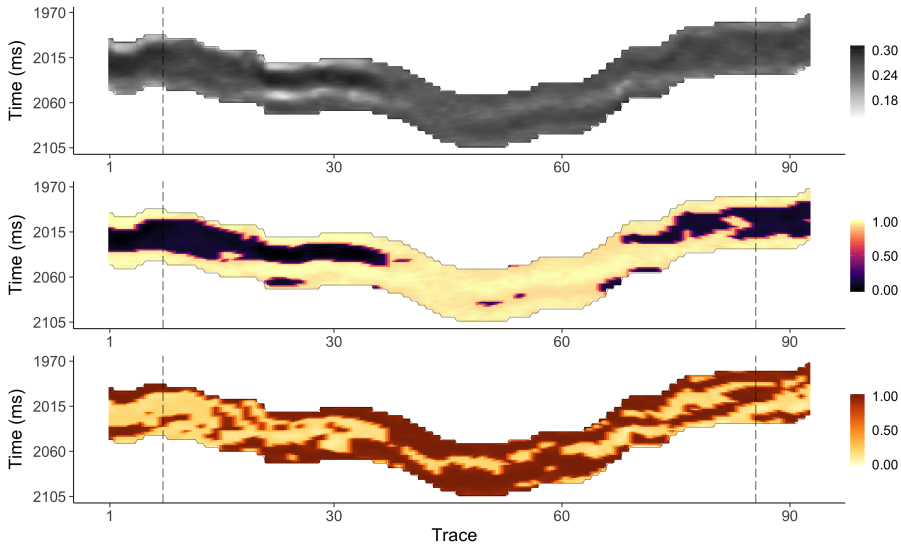

Figure 8: MMAP predictions from model 1. (Top) Porosity. (Middle) Water saturation. (Bottom) Volume of clay. The locations of Well A and Well B are marked.

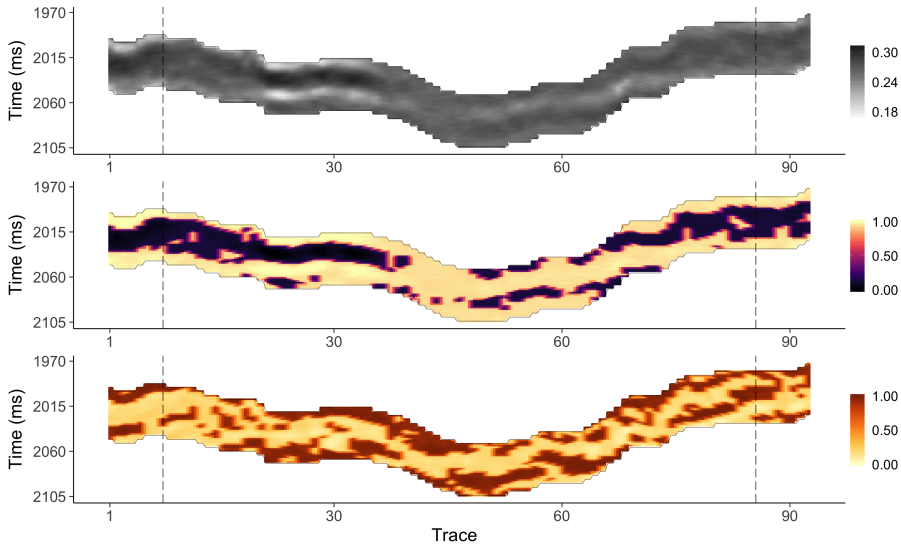

Figure 9: MMAP predictions from model 2. (Top) Porosity. (Middle) Water saturation. (Bottom) Volume of clay. The locations of Well A and Well B are marked.

The locationwise variances of models 1 and 2 are displayed in Figures 10 and 11, respectively. The variances associated with the porosity predictions from both models are very similar and relatively low, except for along the top and bottom of the 2D section and along a thin structure across CR. The variances of water saturation and volume of clay from model 1, see Figure 10, display a similar pattern with respect to their corresponding predictions: the variances appear relatively low except for at the borders separating low value and high value regions in the MMAP predictions, and this effect is particularly apparent for water saturation. The relatively strong variance contrast for water saturation is likely due to less mixing of low value and high value regions in the predictions than in the volume of clay predictions, which results in lower background variance. Moreover, the contrast between low and high variance is stronger for water saturation and volume of clay than for porosity because the locationwise posteriors of porosity are unimodal, while they are bimodal at the region borders for water saturation and volume of clay. Lastly, the variances appear to be relatively low to the left of CR, coinciding with the strongest signals in the seismic AVO data, see Figure 4. The variances of water saturation and volume of clay from model 2, see Figure 11, appear particularly high in CR and moderately high to the right of CR. To the left of CR, where the strongest signals in the seismic AVO data are, some identifiable regions can be discerned and the background variance appears relatively low. Upon comparison of the two models, the variances of model 1 appear to be more structured and the background variances seem to be lower. This is particularly evident for water saturation. Consequently, predictions from model 2 are more uncertain than predictions from model 1, as expected, because model 1 incorporates data from neighboring traces and thereby reduces uncertainty. The differences in the locationwise variances should generate apparent differences in realizations from the two posterior models, and for water saturation in particular.

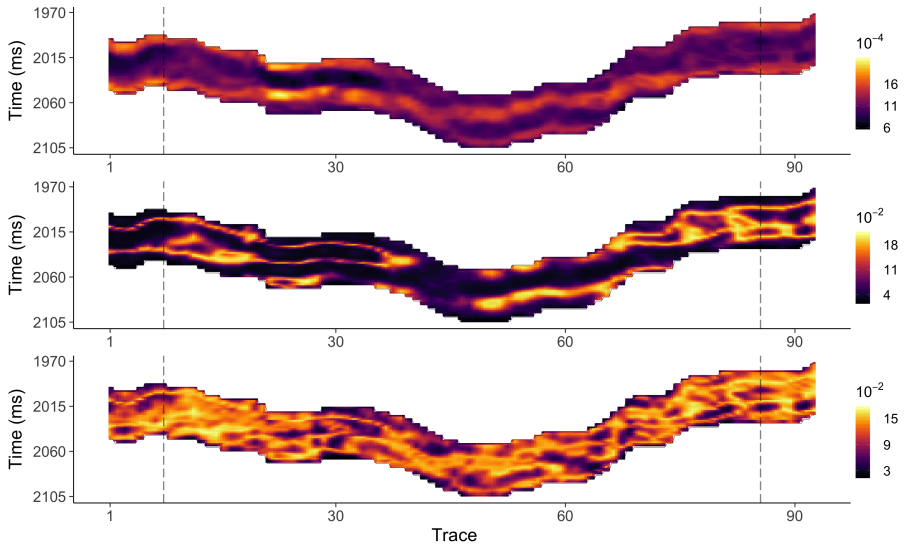

Figure 10: Locationwise variances of model 1. (Top) Porosity. (Middle) Water saturation. (Bottom) Volume of clay. The locations of Well A and Well B are marked.

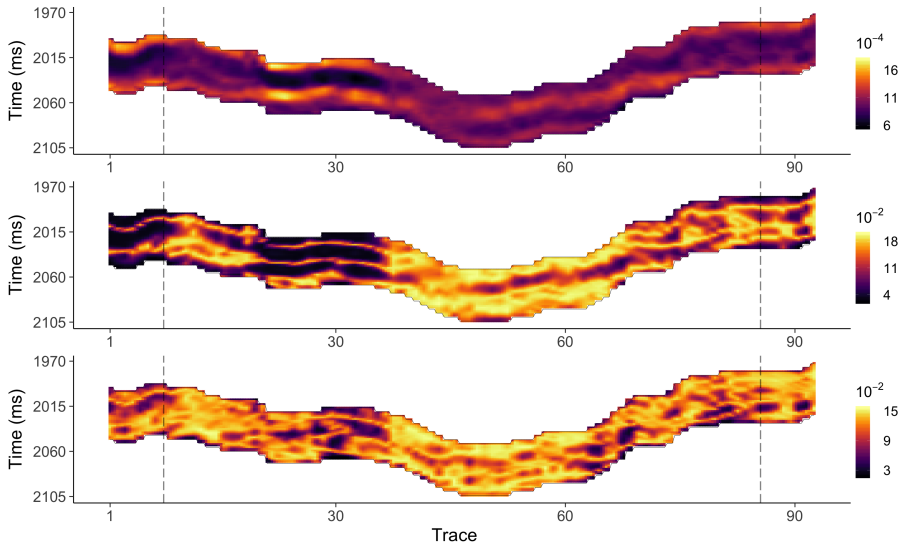

Figure 11: Locationwise variances of model 2. (Top) Porosity. (Middle) Water saturation. (Bottom) Volume of clay. The locations of Well A and Well B are marked.

The inversion results from models 1 and 2 along Well A and Well B are jointly presented in Figure 12. The two models yield very similar porosity results; the predicted porosity profiles are in good agreement with the well log observations and the associated HPDI look reasonable. The predicted water saturation profiles are mostly in good agreement with the well log observations, but a few notable exceptions occur. In Well A, the predictions from both models along the top reservoir suggest the presence of a shale layer which is not indicated by the well logs. Moreover, the two models predict different OWC locations, with model 1 predicting the OWC location roughly 4 ms deeper than model 2. Because Well A is not located exactly on the cross line under study it is unclear whether these well observations are representative; hence, we can not conclude anything about these apparent discrepancies. In Well B, neither model predicts the highly water saturated interval between 1998 ms and 2002 ms. However, this high water saturation occurrence is located within the interval in which the rock physics model performs relatively poorly. The predicted volume of clay profiles from models 1 and 2 are mostly in good agreement, but differ at several lithological transitions. In Well A, the two models yield very similar volume of clay results, but model 1 predicts the deepest high clay content interval to start higher than what model 2 does. In Figures 8 and 9, we see that model 1 predicts a connected high clay content layer from the left of the 2D section and through Well A, whereas model 2 predicts disconnected regions here. A connected region is geologically more likely. In Well B, the two models predict a different thickness of the overlying shale layer. Based on the well log observations, model 1 appears to predict the shale layer thickness accurately, whereas model 2 underpredicts. Both models seem to overpredict the thickness of the high clay content interval in the middle of the well, but model 2 predicts a slightly shorter interval than model 1. Note that the possibly limestone influenced interval is located within these predicted high clay content intervals. Lastly, model 2 predicts a slightly shorter interval of high clay content than model 1 near the deeper end of the well observations. Here, the predictions from model 2 appears to be more accurate than the predictions from model 1. As expected from the locationwise model variances displayed in Figures 10 and 11, the HPDI from model 2 reflect more bimodality in the locationwise posteriors of water saturation than the HPDI from model 1.

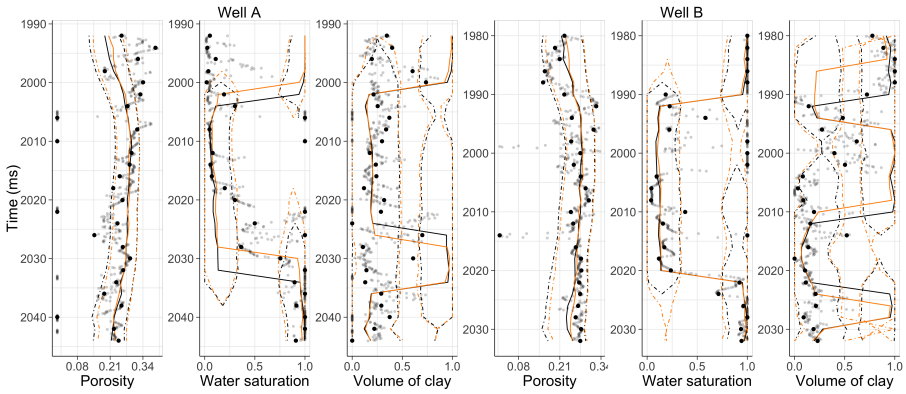

Figure 12: Inversion results from model 1 (black) and model 2 (orange) along Well A and Well B. The MMAP predictions (solid lines) are displayed with 90% HPDI (dot-dashed lines). Well observations are displayed on the reservoir grid (black points) and on the well log grid (transparent points).

One realization of water saturation from each model, truncated to the physical range $[0, 1]$, is displayed in Figure 13. These realizations should be compared with the corresponding MMAP predictions in Figures 8 and 9, respectively. The realization from model 1 exhibits most features of the MMAP prediction, with identifiable regions of low water saturation and with CR being a region of mostly high water saturation. In CR and to the right of CR, the realization appears more heterogeneous than the MMAP, as expected based on the locationwise posterior variances displayed in Figure 10. The realization from model 2 appears similar to the corresponding MMAP prediction to the left of CR, but it is otherwise difficult to recognize features of the MMAP due to very high hetereogeneity. The large locationwise posterior variances in Figure 11 support this predominant heterogeneity and lack of spatial structure. Realizations of porosity and volume of clay do not have as marked differences, but the clear differences observed in the water saturation realizations may still impact potential subsequent results, such as production forecasting by fluid flow simulation.

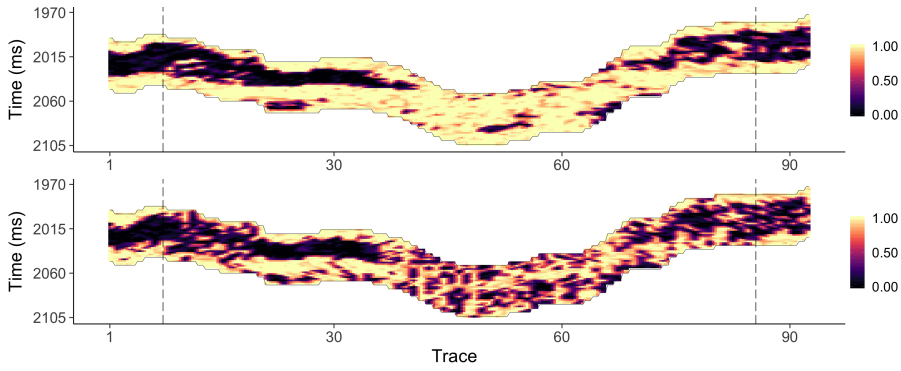

Figure 13: One realization of water saturation from each model, truncated to $[0, 1]$. (Top) Model 1. (Bottom) Model 2.

Quantitative performance measures, based on well log data on the reservoir grid, are presented for the inversion results along Well A and Well B in Table 3 for model 1 and in Table 4 for model 2. As expected, the RMSE and empirical coverage indicate that the results in Well B are overall better than in Well A. The difference in performance between the wells is mostly a consequence of the conflicting predictions and observations at the top of Well A. The models perform very similarly in terms of RMSE. Neither model achieves the desired empirical coverage of 0.9, which is most likely a result of partly unrepresentative data in Well A and a few poorly supported data points in Well B. Model 1 appears to have lower empirical coverage than model 2, as expected, due to variance reduction from lateral correlation. The most favourable feature of the predictions from model 1 is the relatively high homoegeneity compared to the predictions from model 2, see Figures 8 and 9, which strongly suggests that the lateral coupling in model 1 reduces the effects of trace unique signal errors in the seismic AVO data.

Table 3: RMSE and empirical coverage (EC) of the 90% HPDI along wells A and B for model 1.

	Model 1			
	Well A		Well B	
	RMSE	EC	RMSE	EC
$\hat{\boldsymbol{\phi}}$	0.121	0.630	0.059	0.852
$\hat{\mathbf{s}}_w$	0.606	0.519	0.397	0.741
$\hat{\mathbf{v}}_c$	0.392	0.852	0.492	0.741

Table 4: RMSE and empirical coverage (EC) of the 90% HPDI along wells A and B for model 2.

	Model 2			
	Well A		Well B	
	RMSE	EC	RMSE	EC
$\hat{\phi}$	0.118	0.630	0.059	0.852
$\hat{\mathbf{s}}_w$	0.545	0.630	0.390	0.778
$\hat{\mathbf{v}}_c$	0.394	0.889	0.495	0.815

Case: Real seismic AVO data from a 3D volume

We apply the proposed seismic inversion methodology to a 3D volume of the Kneler field in the Alvheim oil and gas field, in the North Sea. The 3D volume consists of five crosslines containing traces 70 through 90; hence, the region between traces 70 and 90 of the 2D section studied previously, see Figure 1, is contained. The center crossline in the 3D volume is referred to as C_{2D}, which is the crossline studied in the previous section. The 3D volume contains a total of 105 traces with a regular inter-trace spacing of 50 m, each of which covers a depth range of 52 ms in two-way time. That is, the distance between inlines and crosslines is 50 m.

The reservoir grid $\mathcal{G}_r$ consists of $n_r = 2835$ grid points which are regularly spaced laterally and with trace-unique vertical gridding, while the seismic grid $\mathcal{G}_d$ consists of $n_d = 1470$ grid points with similar grid structure. The trace-unique vertical gridding consists of 27 points in the reservoir grid and 14 points in the seismic grid.

Model construction

The model construction is based on Well B; hence, the likelihood model is identical to the likelihood model of the previous section with estimated wavelet shown in Figure 5 and rock physics parameters listed in Table 1. Moreover, the prior model parameter values are also identical and are displayed in Table 2. However, note that the prior model is laterally correlated in one additional spatial dimension compared to the laterally correlated prior model of the previous section.

Results

The MMAP predictions are displayed in Figure 14. The MMAP predictions from the previous section should be compared to the predictions on crossline C_{2D}, which appear very similar. However, a notable exception can be seen in the water saturations and volumes of clay: the oil region and the high clay content region in the middle are now apparently disconnected. The geometry of the oil zone and the high clay content regions slowly vary across crosslines, demonstrating lateral continuity perpendicular to the 2D section of the previous section.

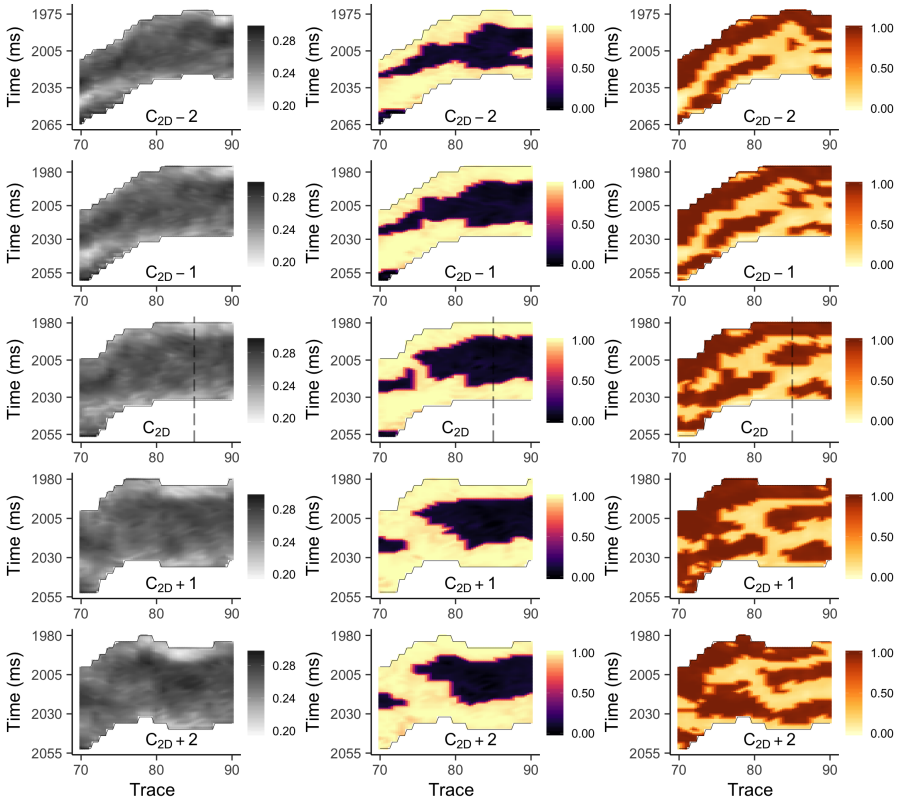

Figure 14: MMAP predictions. (Left column) Porosity. (Middle column) Water saturation. (Right column) Volume of clay. The location of Well B is marked.

The locationwise variance of the posterior model on crossline C_{2D} is displayed in Figure 15. Compared to the locationwise variances of the posterior models in the previous section, shown in Figures 10 and 11, the variance appears to be reduced. This is particularly evident in the variances of water saturation, which now have more structure due to the disappearance of some high variance borders inside the oil zone. The variance reduction is expected due to conditioning on additional neighboring traces perpendicular to the ones of the previous section in the sampling. Consequently, realizations from this posterior model should exhibit less heterogeneity than realizations from either of the posterior models in the previous section.

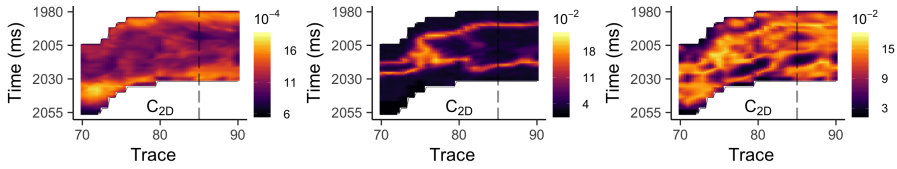

Figure 15: Locationwise variances of the posterior on C_{2D}. (Left) Porosity. (Middle) Water saturation. (Right) Volume of clay. The location of Well B is marked.

The inversion result is presented along Well B in Figure 16. The MMAP predictions are overall in good agreement with the observations for all reservoir properties. Furthermore, the HPDI appear reasonable. Compared to the inversion results along Well B in the previous section, the porosity result appears as very similar. The predicted water saturation profile is very similar to the previously predicted profiles, but the associated HPDI now reflect less bimodality in the posteriors. Lastly, the predicted volume of clay profile is consistent with the predicted volume of clay profile from the laterally correlated model in the previous section, but does not overpredict the thickness of the high clay content interval in the middle of the well as much. Moreover, the associated HPDI reflect less bimodality in the posteriors.

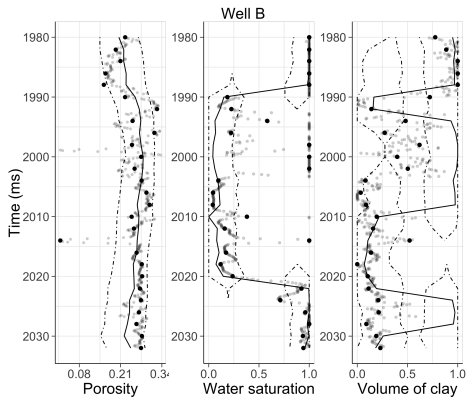

Figure 16: Inversion result along Well B. The MMAP predictions (solid lines) with 90% HPDI (dot-dashed lines). Well observations are displayed on the reservoir grid (black points) and on the well log grid (transparent points).

Quantitative performance measures, based on well log data on the reservoir grid, are presented for the inversion results along Well B in Table 5. Compared to the RMSE and empirical coverage of model 1 along Well B in the previous section, the RMSE has slightly improved for all reservoir properties. The empirical coverages have not changed.

Table 5: RMSE and empirical coverage (EC) of the 90% HPDI along Well B.

	RMSE	EC
$\hat{\boldsymbol{\phi}}$	0.057	0.852
$\hat{\mathbf{s}}_w$	0.392	0.741
$\hat{\mathbf{v}}_c$	0.483	0.741

Conclusion

A Bayesian seismic AVO inversion scheme for porosity, water saturation, and volume of clay is proposed. The prior model is a laterally correlated S-GRF, which can represent the gravitationally induced bimodality observed in the water saturations. The likelihood model is Gauss-linear, for which the S-GRF prior model is conjugate; hence, the posterior model is also an S-GRF. An alternative S-GRF prior model without lateral correlation is defined for comparison through seismic inversion of a 2D section of real seismic AVO data from the Kneler field in the Alvheim oil and gas field in the North Sea. Two major high porosity oil zones are identifiable in the inversion results from both models, indicating high lateral continuity in the seismic AVO data. However, the laterally correlated S-GRF provides predictions that are more homogeneous and with more clearly defined regions than the laterally uncorrelated S-GRF. Furthermore, a relative locationwise variance reduction is observed in the laterally correlated posterior S-GRF for water saturation and volume of clay, and this is reflected in realizations of water saturation from the posterior models. Inversion of a 3D volume containing a region of the 2D section is performed, and the inversion of the 2D region is compared between the 2D and 3D inversion. An improvement in the predictions and a reduction in the inversion uncertainty is observed in the 3D inversion. Hence, although the seismic AVO data appears to have strong lateral continuity, lateral correlation in the prior model for the reservoir properties appears to be favourable to reduce trace unique signal errors, even in conservative amounts as in this study.

Computationally, a laterally uncorrelated prior model is preferable because sampling from the posterior model can be done on a trace-by-trace basis. A laterally correlated prior model entails the involvement of neighboring traces in conditional sampling from the posterior. This requires more complex sampling schemes and higher computer demands. However, conditioning on neighboring traces appears to improve the inversion results and make the results less susceptible to trace unique signal errors. The laterally correlated S-GRF is a more complex prior model, but the benefits discussed above seem to justify the additional cost this entails.

Acknowledgements

This research is part of the Uncertainty in Reservoir Evaluation (URE) activity at the Norwegian University of Science and Technology (NTNU).

The authors thank Aker BP for data access and the right to publish. The authors acknowledge application of the R-packages *MASS* and *akima* in the computational algorithms and data visualization.

Appendix: Sampling from an m-dimensional S-GRF

To draw samples from the m-dimensional and n-variate S-GRF $\mathbf{x}$, we make use of the decomposition

$$\begin{aligned} p(\mathbf{x}) &= p\left(\tilde{\mathbf{x}}|\boldsymbol{\nu} \in \mathbf{A}\right) = p\left(\tilde{\mathbf{x}}|\boldsymbol{\nu}\right) p\left(\boldsymbol{\nu}|\boldsymbol{\nu} \in \mathbf{A}\right) \\ &= \varphi_n\left(\tilde{\mathbf{x}}; \boldsymbol{\mu}_{\tilde{x}|\nu}, \boldsymbol{\Sigma}_{\tilde{x}|\nu}\right) \times \frac{I(\boldsymbol{\nu} \in \mathbf{A})\varphi_n\left(\boldsymbol{\nu}; \boldsymbol{\mu}_{\nu}, \boldsymbol{\Sigma}_{\nu}\right)}{\int_{\mathbb{R}^n} I(\boldsymbol{\nu} \in \mathbf{A})\varphi_n\left(\boldsymbol{\nu}; \boldsymbol{\mu}_{\nu}, \boldsymbol{\Sigma}_{\nu}\right)\ \mathrm{d}\boldsymbol{\nu}}. \end{aligned} \tag{A-1}$$

A sample is drawn by first sampling $\boldsymbol{\nu} \in \mathbf{A}$, followed by sampling $\mathbf{x}$ from $\varphi_n\left(\tilde{\mathbf{x}}; \boldsymbol{\mu}_{\tilde{x}|\nu}, \boldsymbol{\Sigma}_{\tilde{x}|\nu}\right)$. The S-GRF is assumed to consist of $n = n_v \times n_h$ grid points, with n_v and n_h being the number of grid points vertically and laterally, respectively.

Algorithm 1: Draw k samples from $\mathbf{x}$ by M-H McMC, trace-by-trace.

Partition $\tilde{\mathbf{x}}$ and $\boldsymbol{\nu}$ into traces, i.e., $\tilde{\mathbf{x}} \in \mathbb{R}^{n_v}$ and $\boldsymbol{\nu} \in \mathbb{R}^{n_v}$.
Partition each trace into n_b vertical blocks
$\mathbf{b}_i \subset \{1, ..., n_v\}; i = 1, ..., n_b$, and define block neighborhoods
$\mathbf{b}_i^n = \{1, ..., n_v\} \backslash \mathbf{b}_i$.

Iterate n_h **times**
Initialize $\boldsymbol{\nu}$ with a value in $\mathbf{A}$.
  Iterate k **times**
    Iterate n_b **times**
      Select vertical block $\mathbf{b}_i$.
      Sample $\boldsymbol{\nu}'_{b_i|b_i^n}$ sequentially from $q(\boldsymbol{\nu}'_{b_i|b_i^n})$.
      Accept $\boldsymbol{\nu} = [\boldsymbol{\nu}'_{b_i|b_i^n}, \boldsymbol{\nu}_{b_i^n}]^T$ with probability α.
    End
  Sample $\mathbf{x} \sim \varphi_{n_v}\left(\tilde{\mathbf{x}}|\boldsymbol{\nu}\right)$.
  End
End

Algorithm 2: Draw k samples from $\mathbf{x}$ by M-H McMC.

Partition the field into n_b m-dimensional rectangular blocks, $\mathbf{b}_i \subset \{1, ..., n\}; i = 1, ..., n_b$, and define block neighborhoods $\mathbf{b}_i^n = (\mathbf{b}_{v,i}^n \times \mathbf{b}_{h,i}) \bigcup \mathbf{b}_{h,i}^n$, where $\mathbf{b}_{v,i}^n = \{1, ..., n_v\} \backslash \mathbf{b}_{v,i}$ is the vertical neighborhood, $\mathbf{b}_{h,i}$ are the traces in $\mathbf{b}_i$, and $\mathbf{b}_{h,i}^n$ consists of laterally neighboring traces of $\mathbf{b}_i$. Moreover, the block $\mathbf{b}_i$ has complementing grid points $\mathbf{b}_i^c = \{1, ..., n\} \backslash \mathbf{b}_i$.

Initialize $\boldsymbol{\nu}$ with a value in $\mathbf{A}$.
Iterate k **times**
  Iterate n_b **times**
    Select block $\mathbf{b}_i$.
    Sample $\boldsymbol{\nu}'_{b_i|b_i^n}$ sequentially from $q(\boldsymbol{\nu}'_{b_i|b_i^n})$.
    Accept $\boldsymbol{\nu} = [\boldsymbol{\nu}'_{b_i|b_i^n}, \boldsymbol{\nu}_{b_i^c}]^T$ with probability α.
  End
Sample $\mathbf{x} \sim \varphi_n\left(\tilde{\mathbf{x}}|\boldsymbol{\nu}\right)$.
End

Some clarifications of the algorithms follow:

- The proposal distribution is

$$q(\boldsymbol{\nu}_{b|b^n}) = \prod_{i \in b} I\left(\nu_i \in A\right) \frac{\varphi_1\left(\nu_i; \mu_{\nu_i|\nu_{b^{v,n}}}, \sigma^2_{\nu_i|\nu_{b^{v,n}}}\right)}{\Phi_1\left(A; \mu_{\nu_i|\nu_{b^{v,n}}}, \sigma^2_{\nu_i|\nu_{b^{v,n}}}\right)}, \tag{A-2}$$

  with subindex $b^{v,n}$ denoting the union of the set of indices in $\mathbf{b}$ already visited and the block neighborhood.

- The Metropolis-Hastings acceptance probability, α, is

$$\begin{aligned} \alpha &= \min\left\{1, \frac{p\left(\boldsymbol{\nu}'_{b|b^n}\right)}{p\left(\boldsymbol{\nu}_{b|b^n}\right)} \cdot \frac{q\left(\boldsymbol{\nu}_{b|b^n}\right)}{q\left(\boldsymbol{\nu}'_{b|b^n}\right)}\right\} \\ &= \min\left\{1, \prod_{i \in b} \frac{\Phi_1\left(A; \mu'_{\nu_i|\nu_{b^{v,n}}}, \sigma^2_{\nu_i|\nu_{b^{v,n}}}\right)}{\Phi_1\left(A; \mu_{\nu_i|\nu_{b^{v,n}}}, \sigma^2_{\nu_i|\nu_{b^{v,n}}}\right)}\right\}. \end{aligned} \tag{A-3}$$

The blocks and the order in which they are visited are predetermined, and the associated covariance matrices are precomputed to reduce the computational time.

References

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

Allard, D. and Naveau, P. (2007). A New Spatial Skew-Normal Random Field Model. *Communications in Statistics - Theory and Methods*, 36:1821–1834.

Arellano-Valle, R. B., Branco, M. D., and Genton, M. G. (2006). A Unified View on Skewed Distributions Arising from Selections. *The Canadian Journal of Statistics*, 34:581–601.

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

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

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

Fjeldstad, T. M. and Grana, D. (2018). Joint probabilistic petrophysics-seismic inversion based on Gaussian mixture and Markov chain prior models. *Geophysics*, 83:1JF–Z3.

Forberg, O. B., Grana, D., and Omre, H. (2021a). Bayesian inversion of time-lapse seismic AVO data for multimodal reservoir properties. *IEEE Transactions on Geoscience and Remote Sensing*.

Forberg, O. B., Kjøsnes, Ø., and Omre, H. (2021b). Bayesian seismic AVO inversion for reservoir variables with bimodal spatial histograms. *Geophysics*, 86.

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. M., and Omre, H. (2017). Bayesian Gaussian Mixture Linear Inversion for Geophysical Inverse Problems. *Mathematical Geosciences*, 49:493–515.

Gunning, J. and Glinsky, M. E. (2004). Delivery: an open-source model-based Bayesian seismic inversion program. *Computers & Geosciences*, 30(6):619–636.

Hasselblad, V. (1966). Estimation of Parameters for a Mixture of Normal Distributions. *Technometrics*, 8:431–444.

Karimi, O., Omre, H., and Mohammadzadeh, M. (2010). Bayesian closed-skew Gaussian inversion of seismic AVO data for elastic material properties. *Geophysics*, 75:R1–R11.

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

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.

Mavko, G., Mukerji, T., and Dvorkin, J. (2009). *The rock physics handbook*. Cambridge University Press.

Mosegaard, K. and Tarantola, A. (1995). Monte Carlo sampling of solutions to inverse problems. *Journal of Geophysical Research*, 100:12431–12447.

Omre, H. and Rimstad, K. (2021). Bayesian Spatial Inversion and Conjugate Selection Gaussian Prior Models. *Journal of Uncertainty Quantification*, To appear.

Rimstad, K., Avseth, P., and Omre, H. (2012). Hierarchical Bayesian lithology/fluid prediction: A North Sea case study. *Geophysics*, 77:B69–B85.

Rimstad, K. and Omre, H. (2014). Skew-Gaussian Random Fields. *Spatial Statistics*, 10:43–62.

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

Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. SIAM.

Wang, Y. (2006). *Seismic Inversion: Theory and Applications*. Wiley Blackwell.

Zoeppritz, K. B. (1919). Vllb. On reflection and transmission of seismic waves by surfaces of discontinuity. *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-physikalische Klasse*, pages 66–84.

ISBN 978-82-326-6231-9 (printed ver.)
ISBN 978-82-326-5505-2 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

NTNU

Norwegian University of
Science and Technology