Markov Chain Monte Carlo algorithms for target-oriented and interval-oriented amplitude versus angle inversion with non-parametric priors and non-linear forward modellings

Mattia Aleardi[1], Alessandro Salusti[1,2]

[1]University of Pisa, Earth Sciences Department, via S. Maria 53, 56126, Pisa, Italy

[2]University of Florence, Earth Sciences Department, via. G. La Pira 4, 50121, Florence, Italy

Corresponding author: Mattia Aleardi, mattia.aleardi@unipi.it

ABSTRACT

In geophysical inverse problems the posterior model can be analytically assessed only in case of linear forward operators, Gaussian, Gaussian mixture, or generalized Gaussian prior models, continuous model properties, and Gaussian-distributed noise contaminating the observed data. For this reason, one of the major challenges of seismic inversion is to derive reliable uncertainty appraisals in cases of complex prior models, non-linear forward operators and mixed discrete-continuous model parameters. We present two amplitude versus angle inversion strategies for the joint estimation of elastic properties and litho-fluid facies from pre-stack seismic data in case of non-parametric mixture prior distributions and non-linear forward modellings. The first strategy is a 2-dimensional target-oriented inversion that inverts the Amplitude Versus Angle responses of the target reflections by adopting the single-interface full Zoeppritz equations. The second is an interval-oriented approach that inverts the pre-stack seismic responses along a given time interval using a 1-dimensional convolutional forward modelling still based on the Zoeppritz equations. In both approaches the model vector includes the facies sequence and the elastic properties of P-wave velocity, S-wave velocity and density. The distribution of the elastic properties at each common-mid-point location (for the target-oriented approach) or at each time-sample position (for the time-interval approach) is assumed to be multimodal with as many modes as the number of litho-fluid facies considered. In this context, an

analytical expression of the posterior model is no more available. For this reason, we adopt a Markov Chain Monte Carlo algorithm to numerically evaluate the posterior uncertainties. With the aim of speeding up the convergence of the probabilistic sampling, we adopt a specific recipe that includes multiple chains, a parallel tempering strategy, a delayed rejection updating scheme and hybridizes the standard Metropolis-Hasting algorithm with the more advanced Differential Evolution Markov Chain method. For the lack of available field seismic data, we validate the two implemented algorithms by inverting synthetic seismic data derived on the basis of realistic subsurface models and actual well log data. The two approaches are also benchmarked against two analytical inversion approaches that assume Gaussian-mixture distributed elastic parameters. The final predictions and the convergence analysis of the two implemented methods proved that our approaches retrieve reliable estimations and accurate uncertainties quantifications with a reasonable computational effort.

Keywords

AVA; Bayesian inversion; Reservoir Characterization.

INTRODUCTION

One of the main objectives of reservoir characterization is to exploit the acquired seismic and well log data to infer the distribution of elastic parameters and litho-fluid facies around the investigated area. From the mathematical point of view this process is an ill-conditioned inverse problem (Tarantola, 2005; Aster et al. 2011) in which many models can fit the observed data equally well. For this reason, one goal of reservoir characterization studies is the quantification of the uncertainties affecting the recovered solution, which are expressed by the so-called posterior probability density function (*pdf*). The Bayesian approach is usually adopted to accurately propagate the uncertainties form the available seismic and geologic information to the estimated model. In this context it is crucial

including as much a-priori information as possible (e.g. expected lateral variability and mutual interdependence of the inverted parameters) to successfully constrain the inversion procedure.

One challenge of this inversion process concerns the simultaneous estimation of discrete (i.e. litho-fluid facies) and continuous (elastic properties) model parameters from the observed data (Gunning and Sams, 2018). Another challenge is related to the complexity of the property distribution and correlation. For example, the distribution of elastic properties is often multimodal due to the presence of multiple litho-fluid facies. Many inversion methods have been proposed to account for this multimodality (Grana and Della Rossa, 2010; Bosh et al. 2010; Aleardi et al. 2018a), but an analytical and computationally fast derivation of the posterior model is only possible in cases of linear forward operators, Gaussian, Gaussian-mixture, or generalized Gaussian distributed model parameters and Gaussian errors in the seismic data. However, the validity of the Gaussian or Gaussian-mixture assumptions is often case dependent because they could not be adequate to reliably capture the complex relations among elastic attributes and litho-fluid facies. At the same time, the linear forward model might not be sufficiently accurate to describe the relation between seismic data and elastic parameters in cases of strong elastic contrasts at the reflecting interface and far source-receiver offsets. In these cases, oversimplified prior models and/or forward modelling operators could provide unreliable or even biased model parameter estimations (Madsen and Hansen, 2018; Aleardi 2018).

For this reason, it is often advisable to numerically evaluate the posterior model through a Markov Chain Monte Carlo (MCMC) algorithm. From the one hand, MCMC methods have been successfully applied to solve many geophysical problems (Sambridge and Mosegaard, 2002; Malinverno, 2002; Bosh et al. 2007; Aleardi and Mazzotti, 2017) as they can theoretically assess the posterior uncertainties in cases of complex (i.e, non-parametric) prior distributions and non-linear forward modellings. From the other hand, these methods convert the inversion problem into a sampling problem and for this reason they require a much larger computational effort with respect to the analytical approach. Moreover, the use of non-parametric priors often complicates the inclusion of

geostatistical a-priori information (e.g. a semivariogram model) into the inversion procedure and for this reason the use of non-parametric models is not so common in geophysical inversions although some successful and significative examples can be found in Sabeti at al. (2017) and Grana (2018). Finally, classical MCMC methods, such as the Metropolis-Hastings algorithm, are known to mix slowly between the modes if the target distribution is multimodal (Holmes et al. 2017). To partially overcome this issue, multiple MCMC chains are usually employed so that the ability to exhaustively explore the high probability regions of the model space is enhanced.

The sampling density of MCMC is designed to be proportional to the posterior, so that the sampled models can be used to approximate the statistical properties of the target *pdf* (the so-called Bayesian integrals; Sen and Stoffa, 1996). In more detail, the first stage of the MCMC sampling (usually called the burn-in period) can be viewed as a global optimization that moves from a random starting model to a high-probability region of the model space. The second stage is often called the sampling stage in which the small fluctuation of the misfit value indicates that the MCMC algorithm reaches the stationary regime. Usually the samples accepted during the burn-in period do not accurately represent the posterior *pdf* and for this reason they are not considered in the computation of the posterior *pdf*. From a theoretical point of view, it is also known that for independent samples the approximation error of MCMC is proportional to $1/\sqrt{N}$ where N is the number of samples (MacKay, 2003). When samples are correlated, not only the convergence is typically slower, but also there is risk to derive biased *pdf* estimations. For this reason, not all the samples collected after the burn-in period are usually used to numerically estimate the posterior but several iterations of the algorithm are allowed to elapse in between successive samples: this number of iterations is the so-called lag value that can be set from the analysis of the autocorrelation function of the sampled models.

Determining the convergence of the MCMC algorithm to stable posterior distribution is a crucial aspect of any MCMC inversion. To this end several strategies can be employed (Gelman and Rubin, 1992). For example, in cases of multiple chains we can compute the potential scale reduction factor

(PSRF) for each model parameter. This number quantifies the difference between the "within-walk" and "between-walk" estimated variances. The PSRF decreases to 1 as the number of drawn samples tends to infinite. A high PSRF value indicates that the variance within the walks is small compared to that between the walks and that longer walks are needed to converge to a stable distribution. Usually, a PSRF lower than 1.2 for a given unknown proves that convergence has been achieved for that model parameter (Gelman et al. 1995).

In this work, we present two amplitude versus angle (AVA) inversion approaches that invert the pre-stack seismic response for the simultaneous estimation of elastic properties (P-wave velocity; *Vp*; S-wave velocity; *Vs*; and density ρ), litho-fluid facies and the related uncertainties. The first method performs a target-oriented inversion that considers the AVA responses of the target reflection extracted along a previously interpreted 2D stratigraphic horizon (Mazzotti and Zamboni 2003; Aleardi and Mazzotti 2014; Gongand and McMechan 2016). This target-oriented approach relies on two main assumptions: the data are plane waves, and the amplitude and phase effects produced during propagation through the overburden are adequately compensated during a tailored seismic processing sequence (Mazzotti and Ravagnan, 1995). From the one hand, the main advantage of the target-oriented approach relies on its limited computational cost with respect to the interval-oriented inversion. From the other hand, the target-oriented method requires a previous accurate geological interpretation phase in which the target reflection is identified and accurately mapped throughout the inverted 3D volume. In addition, a very delicate step is the extraction of the reflection coefficients from the observed reflected amplitudes and phases. A robust procedure to accomplish this task is described, for example, in Grion et al. (1998). Notwithstanding these limitations, the target-oriented approach has been investigated by many authors and successfully applied in different exploration areas (e.g., among many others, Adriansyah, and McMechan, 2001; Zhu, X., and McMechan, 2012; Aleardi et al. 2017). The second implemented algorithm inverts the 1D pre-stack seismic response extracted along a given time-interval and for this reason here we call this method an interval-oriented inversion (Buland and Omre, 2003). Both approaches solve the mixed continuous-discrete inversion

problem by adopting the non-linear Zoeppritz equations as the forward modelling and a non-parametric mixture prior model. Normal score transformation (Kennedy and Gentle 2018) is used to convert the non-parametric prior distribution into a Gaussian model so that geostatistical a-priori information about the lateral (for the target-oriented inversion) or vertical (for the interval-oriented inversion) variability of elastic parameters can be easily included into the optimization procedure by means of a variogram model and kriging interpolation (Doyen, 2007). In addition, a first-order Markov chain model is included to preserve the lateral or vertical continuity of the discrete facies distribution.

As previously mentioned, in the context of MCMC algorithms, the simultaneous presence of continuous and discrete parameter could result in persistent rejections of models, very low acceptance ratio and slow convergence of the sampling. For this reason, we adopt a specific MCMC recipe to speed up the convergence of the sampling. Our implementation is inspired by the method proposed by Holmes et al. (2017) but also includes multiple chains, a parallel tempering strategy (Falcioni and Deem, 1999; Sambridge, 2013), a delayed rejection updating scheme (Tierney and Mira, 1999) and hybridizes the standard Metropolis-Hasting algorithm with the more advanced Differential Evolution Markov Chain (DEMC) method (ter Braak and Vrugt, 2008).

The parallel tempering (PT) strategy is a meta-algorithm in which multiple interactive chains run at different temperature levels. The introduction of the temperature parameter (T) rescales the *pdf* to be sampled: High-temperature chains ensure a wide exploration of the model space, whereas low-temperature chains ensure exploitation of high-probability regions. According to stochastic criteria, swaps of models are allowed between chains at different temperatures, so that high temperature chains ensure that low-temperature chains access all the high probability regions while maintaining an efficient exploitation capability. The PT approach revealed to be particularly useful to explore highly multimodal posterior *pdfs* (Dosso et al. 2012). In addition, we include a delayed rejection scheme to additionally decrease the computational cost. The basic idea of this method is to learn from the

previous moves by adapting the proposal distribution to the local shape of the target *pdf*. Finally, the DEMC approach combines a standard MCMC algorithm with principles coming from the differential evolution optimization algorithm with the aim to promote the mixing between different chains and expedite the convergence of the MCMC sampling to the stationary regime.

The outcomes yielded by the two proposed MCMC inversion algorithms are compared against those provided by analytical inversion approaches that assume Gaussian-mixture distributed model parameters and linearized approximations of the Zoeppritz equations as forward modellings. For the lack of available field data, we focus our experiments on synthetic data derived on the basis of realistic subsurface models and actual well-log information.

THE METHODS

In this section we give a theoretical overview of the different methods we employ. We start by briefly describing the target-oriented and interval-oriented analytical approaches. Then, we discuss in more detail the two implemented Markov Chain Monte Carlo (MCMC) algorithms.

Target-oriented analytical inversion

In this case the forward modelling is based on the three-term Aki and Richards (1980) equation:

$$Rpp(\theta) = \frac{1}{2}(1 + tan^2(\theta)\,)\frac{\Delta Vp}{\overline{Vp}} - 4\frac{\overline{Vs}^2}{\overline{Vp}^2}sin^2(\theta)\frac{\Delta Vs}{\overline{Vs}} + \frac{1}{2}\left(1 - 4\frac{\overline{Vs}^2}{\overline{Vp}^2}sin^2(\theta)\right)\frac{\Delta\rho}{\bar{\rho}}$$

$$= \alpha_{Vp}(\theta)\frac{\Delta Vp}{\overline{Vp}} + \alpha_{Vs}(t,\theta)\frac{\Delta Vs}{\overline{Vs}} + \alpha_{\rho}(t,\theta)\frac{\Delta\rho}{\bar{\rho}} \quad (1)$$

where *Rpp* is the P-P wave reflection coefficient, θ is the incidence angle, Δx indicates the contrast in the elastic property *x* across the reflecting interface, whereas $\bar{x}$ is the average value of the property *x* over the reflecting interface. The terms $\frac{\Delta Vp}{\overline{Vp}}$, $\frac{\Delta Vs}{\overline{Vs}}$ and $\frac{\Delta\rho}{\bar{\rho}}$ are also known as the *Vp*, *Vs* and density

reflectivities and indicated with R_{Vp}, R_{Vs}, and R_{ρ} respectively. Following a matrix formalism, equation (1) can be written as:

$$\mathbf{d} = \mathbf{G}\mathbf{e}, \quad (2)$$

where **d** is the observed data vector, **e** is the vector of elastic parameters and **G** is the forward modelling matrix operator. More in detail:

$$\mathbf{d} = [Rpp(\theta_1), \dots, Rpp(\theta_N)]^T, \quad (3)$$

$$\mathbf{e} = [R_{Vp}, R_{Vs}, R_{\rho}]^T, \quad (4)$$

$$\mathbf{G} = \begin{bmatrix} \frac{1}{2}(1 + tan^2(\theta_1)) & -4\frac{\overline{Vs}^2}{\overline{Vp}^2}sin^2(\theta_1) & \frac{1}{2}\left(1 - 4\frac{\overline{Vs}^2}{\overline{Vp}^2}sin^2(\theta_1)\right) \\ \frac{1}{2}(1 + tan^2(\theta_2)) & -4\frac{\overline{Vs}^2}{\overline{Vp}^2}sin^2(\theta_2) & \frac{1}{2}\left(1 - 4\frac{\overline{Vs}^2}{\overline{Vp}^2}sin^2(\theta_2)\right) \\ \vdots & \vdots & \vdots \\ \frac{1}{2}(1 + tan^2(\theta_N)) & -4\frac{\overline{Vs}^2}{\overline{Vp}^2}sin^2(\theta_N) & \frac{1}{2}\left(1 - 4\frac{\overline{Vs}^2}{\overline{Vp}^2}sin^2(\theta_N)\right) \end{bmatrix}, \quad (5)$$

where *N* is the number of data points, whereas the $\frac{\overline{Vs}^2}{\overline{Vp}^2}$ value needed to define the **G** matrix can be derived from the available information about the investigated area (i.e. borehole data). In case of a target-oriented amplitude versus angle (AVA) inversion, the data are the AVA responses of the target reflection extracted for each common-mid-point (CMP) gather along the in-line and cross-line directions, whereas the model parameters are the three R_{Vp}, R_{Vs}, and R_{ρ}terms pertaining to each CMP gather.

Under the assumption of a Gaussian-mixture prior model, the a-priori information can be written as (Grana and Della Rossa, 2010):

$$p(\mathbf{e}) = \sum_{k=1}^{K} \omega_i N(\mathbf{e}; \boldsymbol{\mu}_{\mathbf{e}}^k, \boldsymbol{\Sigma}_{\mathbf{e}}^k), \quad (6)$$

where K is the total number of components of the Gaussian-mixture distribution, ω_i is the prior weight of the i-th component, N indicates the Gaussian distribution with mean vector $\mathbf{\mu}_\mathbf{e}^k$ and covariance matrix $\mathbf{\Sigma}_\mathbf{e}^k$, whereas the superscript k indicates that the mean and the covariance are facies dependent. In practice, the elastic reflectivities are assumed to be Gaussian distributed within each facies. The number of facies and the statistical properties of each Gaussian component can be determined from available geological information about the investigated area. In particular, the coefficients ω_i define the a-priori probability of the i-th facies.

In case of a linear forward operator the posterior model is again a Gaussian-mixture and is given by:

$$p(\mathbf{e}|\mathbf{d}) = \sum_{k=1}^{K} \lambda_i N\left(\mathbf{e}; \mathbf{\mu}_{\mathbf{e}|\mathbf{d}}^k, \mathbf{\Sigma}_{\mathbf{e}|\mathbf{d}}^k\right), \quad (7)$$

where the posterior weights λ_i, and the a-posteriori mean vector and covariances matrix ($\mathbf{\mu}_{\mathbf{e}|\mathbf{d}}^k$ and $\mathbf{\Sigma}_{\mathbf{e}|\mathbf{d}}^k$, respectively) can be derived as follows:

$$\mathbf{\mu}_{\mathbf{e}|\mathbf{d}}^k = \mathbf{\mu}_\mathbf{e}^k + (\mathbf{G}^T\mathbf{\Sigma}_\mathbf{d}^{-1}\mathbf{G} + (\mathbf{\Sigma}_\mathbf{e}^k)^{-1})^{-1}\mathbf{G}^T\mathbf{\Sigma}_\mathbf{d}^{-1}(\mathbf{d} - \mathbf{G}\mathbf{\mu}_\mathbf{e}^k), \quad (8)$$

$$\mathbf{\Sigma}_{\mathbf{e}|\mathbf{d}}^k = (\mathbf{G}^T\mathbf{\Sigma}_\mathbf{d}^{-1}\mathbf{G} + (\mathbf{\Sigma}_\mathbf{e}^k)^{-1})^{-1}, \quad (9)$$

$$\lambda_i = \frac{\omega_i N(\mathbf{d}; \mathbf{\mu}_\mathbf{d}^k, \mathbf{\Sigma}_\mathbf{d}^k)}{\sum_{i=1}^{K} \omega_i N(\mathbf{d}; \mathbf{\mu}_\mathbf{d}^k, \mathbf{\Sigma}_\mathbf{d}^k)}. \quad (10)$$

The coefficients λ_i represent the point-wise posterior probability of facies, where point-wise means that the coefficients λ_i at each spatial position are independent from those estimated at the neighboring positions. Finally, the a-posteriori mean elastic model can be derived as a weighted summation over the posterior Gaussian components (de Figueiredo et al. 2017):

$$\mathbf{e}_{est} = \sum_{k=1}^{K} \lambda_i \mathbf{\mu}_{\mathbf{e}|\mathbf{d}}^k. \quad (11)$$

In the context of a target-oriented approach, the inclusion of a 2D geostatistical model to laterally constraints the inverted elastic parameters will significantly increase the computational cost of the algorithm (Aleardi et al. 2018b). For this reason, no lateral constraints are included into the implemented target-oriented analytical approach. In other words, each AVA response extracted from each CMP gather is inverted independently.

Interval-oriented analytical inversion

In this case, the forward modelling is given by the extension along a time interval of the single-interface reflection coefficients of equation (1) (see Buland and Omre, 2003). If we consider the convolutional modelling and we adopt the matrix formalism, the observed CMP gather **d** can be computed as:

$$\mathbf{d} = \mathbf{SADe} = \mathbf{Ge}, \quad (12)$$

where **S** is the wavelet matrix, **A** contains the numerical coefficients $\alpha_{Vp}(t)$, $\alpha_{Vs}(t)$ and $\alpha_{\rho}(t)$ of equation (1), **D** is the first order numerical derivative operator and **e** contains discrete time samples of the natural logarithm of Vp, Vs, and ρ. We again assume a Gaussian-mixture prior model for **e**, that is we assume the natural logarithm of elastic properties to be Gaussian-distributed within each litho-fluid facies. The inversion approach we adopt is a modification of that proposed by de Figueiredo et al. (2018) to which we refer the reader for additional details. In this case the statistical properties of the posterior model can be derived as:

$$\boldsymbol{\mu}_{\mathbf{e}|\mathbf{d},\mathbf{f}}^{k} = \boldsymbol{\mu}_{\mathbf{e}}^{k} + \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{e}}^{k} \\ \mathbf{G}\boldsymbol{\Sigma}_{\mathbf{e}}^{k} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{e}}^{k} + \boldsymbol{\Sigma}_{\mathrm{f}} & \boldsymbol{\Sigma}_{\mathbf{e}}^{k}\mathbf{G}^{T} \\ \boldsymbol{\Sigma}_{\mathbf{e}}^{k}\mathbf{G} & \mathbf{G}^{T}\boldsymbol{\Sigma}_{\mathbf{d}}^{-1}\mathbf{G} \end{bmatrix}^{-1} \left(\begin{bmatrix} \mathbf{f} \\ \mathbf{d} \end{bmatrix} - \begin{bmatrix} \boldsymbol{\mu}_{\mathbf{e}}^{k} \\ \mathbf{G}\boldsymbol{\mu}_{\mathbf{e}}^{k} \end{bmatrix} \right), \quad (13)$$

$$\boldsymbol{\Sigma}_{\mathbf{e}|\mathbf{d},\mathbf{f}}^{k} = \boldsymbol{\Sigma}_{\mathbf{e}}^{k} - \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{e}}^{k} \\ \mathbf{G}\boldsymbol{\Sigma}_{\mathbf{e}}^{k} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{e}}^{k} + \boldsymbol{\Sigma}_{\mathrm{f}} & \boldsymbol{\Sigma}_{\mathbf{e}}^{k}\mathbf{G}^{T} \\ \boldsymbol{\Sigma}_{\mathbf{e}}^{k}\mathbf{G} & \mathbf{G}^{T}\boldsymbol{\Sigma}_{\mathbf{d}}^{-1}\mathbf{G} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{e}}^{k} \\ \mathbf{G}\boldsymbol{\Sigma}_{\mathbf{e}}^{k} \end{bmatrix}^{T}, \quad (14)$$

$$\lambda_i = \frac{\omega_i N(\begin{bmatrix}\mathbf{f}\\ \mathbf{d}\end{bmatrix}; \begin{bmatrix}\boldsymbol{\mu}_{\mathbf{e}}^{k}\\ \mathbf{G}\boldsymbol{\mu}_{\mathbf{e}}^{k}\end{bmatrix}, \begin{bmatrix}\boldsymbol{\Sigma}_{\mathbf{e}}^{k}+\boldsymbol{\Sigma}_{\mathrm{f}} & \boldsymbol{\Sigma}_{\mathbf{e}}^{k}\mathbf{G}^{T}\\ \boldsymbol{\Sigma}_{\mathbf{e}}^{k}\mathbf{G} & \mathbf{G}^{T}\boldsymbol{\Sigma}_{\mathbf{d}}^{-1}\mathbf{G}\end{bmatrix})}{\sum_{i=1}^{K}\omega_i N(\begin{bmatrix}\mathbf{f}\\ \mathbf{d}\end{bmatrix}; \begin{bmatrix}\boldsymbol{\mu}_{\mathbf{e}}^{k}\\ \mathbf{G}\boldsymbol{\mu}_{\mathbf{e}}^{k}\end{bmatrix}, \begin{bmatrix}\boldsymbol{\Sigma}_{\mathbf{e}}^{k}+\boldsymbol{\Sigma}_{\mathrm{f}} & \boldsymbol{\Sigma}_{\mathbf{e}}^{k}\mathbf{G}^{T}\\ \boldsymbol{\Sigma}_{\mathbf{e}}^{k}\mathbf{G} & \mathbf{G}^{T}\boldsymbol{\Sigma}_{\mathbf{d}}^{-1}\mathbf{G}\end{bmatrix})} \quad . \quad (15)$$

Similarly to the target-oriented approach, the coefficients $\boldsymbol{\lambda_i}$ express the point-wise posterior probability density function *(pdf)* of facies at each time sample, whereas the vector **f** is the low frequency elastic back-ground model with uncertainties expressed by the matrix $\boldsymbol{\Sigma_f}$. In this case the a-priori model covariance matrix $\boldsymbol{\Sigma_e^k}$ expresses both the mutual correlation of elastic properties and their vertical variability. In particular, their mutual correlation is given by a stationary covariance matrix, while the vertical correlation is obtained by multiplying (Kronecker product) the stationary covariance matrix by a first-order exponential function (Buland and Omre, 2003). Again, the a-posteriori mean elastic model can be derived by means of equation (11).

The implemented Markov Cain Monte Carlo inversions

In this section we describe the two implemented MCMC algorithms. The MCMC recipes for the target- and interval-oriented inversions are exactly the same because the two methods differ only for specific details related to the different forward modelling operators (i.e. the single interface Zoeppritz equation for the target-oriented approach and the Zoeppritz equations combined with a 1D convolutional operator for the interval-oriented inversion) and for the inclusions of geostatistical constraints (2D lateral constraints for the target-oriented inversion and 1D vertical constraints for the interval-oriented algorithm). For this reason, we first present our MCMC implementation before discussing in more details the differences between the two approaches.

Generally speaking, the MCMC algorithms generates subsurface models that honor the prior information and accept or reject these models on the basis of their likelihood value. Being **m** the current model and **m′** the proposed (perturbed) model, the probability for the chain to move from **m** to **m'** can be computed from the Metropolis-Hasting rule:

$$\alpha = p(\mathbf{m}'|\mathbf{m}) = \min\ [1, \text{prior ratio } \times \text{likelihood ratio } \times \text{proposal ratio }]$$

$$= \min\left[1, \frac{p(\mathbf{m}')}{p(\mathbf{m})} \times \frac{p(\mathbf{d}|\mathbf{m}')}{p(\mathbf{d}|\mathbf{m})} \times \frac{q(\mathbf{m}|\mathbf{m}')}{q(\mathbf{m}'|\mathbf{m})}\right], \quad (16)$$

where $q()$ is the proposal distribution that defines the new model $\mathbf{m}'$ as a random deviate from a probability distribution $q(\mathbf{m}'|\mathbf{m})$ conditioned only on the current model $\mathbf{m}$. Note that the proposal ratio term in equation (16) vanishes if symmetric proposals (for example a Gaussian proposal) are employed. If $\mathbf{m}'$ is accepted $\mathbf{m} = \mathbf{m}'$ and another model is generated as a random perturbation of $\mathbf{m}$. The ensemble of accepted models after the burn-in period is used to numerically compute the posterior *pdf*.

To derive a reliable posterior model we adopt multiple chains that start from different initial points defined on the basis of the a-priori information. The use of multiple chains has several desirable advantages, particularly when dealing with complex posterior distributions involving long tails, correlated parameters, multi-modality, and numerous local optima. In addition, to speed up the convergence of the MCMC sampling we adopt a parallel tempering strategy and a delayed rejection updating scheme.

The parallel tempering strategy runs multiple and interactive chains simultaneously at different temperature levels $T = [T_1, T_2, \dots, T_{max}]$. In this context, for symmetric proposal distributions the Metropolis-Hasting rule becomes:

$$p(\mathbf{m}'|\mathbf{m}) = \min\left[1, \frac{p(\mathbf{d}|\mathbf{m}')^{1/T}}{p(\mathbf{d}|\mathbf{m})^{1/T}} \frac{p(\mathbf{m}')}{p(\mathbf{m})}\right] = \min\left[1, \exp\left(-\frac{\varphi(\mathbf{m}') - \varphi(\mathbf{m})}{2T}\right)\frac{p(\mathbf{m}')}{p(\mathbf{m})}\right], (17)$$

where $\varphi(\mathbf{m}) = ||\mathbf{C_d}^{-\frac{1}{2}}(\mathbf{d} - \mathbf{d_{pre}})||_2^2$ is the weighted L2 norm difference between observed and predicted data, whereas $\mathbf{C_d}$ is the data covariance matrix. In practice the sampling becomes easier as the T increases because the likelihood function becomes flatter with increasing temperature. Note that for $T \to \infty$ the likelihood becomes uninformative and the sampling follows the prior distribution. In the approach to parallel tempering applied here, at each MCMC step the algorithm randomly chooses

two chains *i* and *j* for a proposed swap and these chains are allowed to exchange their current models (or equivalently their current temperature levels) with a probability equal to:

$$p(i,j) = \min\left[1, \frac{\exp\left(-\frac{\varphi(\mathbf{m}_j)}{T_i} - \frac{\varphi(\mathbf{m}_i)}{T_j}\right)}{\exp\left(-\frac{\varphi(\mathbf{m}_i)}{T_i} - \frac{\varphi(\mathbf{m}_j)}{T_j}\right)}\right]$$

$$= \min\left[1, \exp\left[\left(\frac{1}{T_i} - \frac{1}{T_j}\right)\left(\varphi(\mathbf{m}_i) - \varphi(\mathbf{m}_j)\right)\right]\right]. \quad (18)$$

If the swap is accepted:

$$(\mathbf{m}_i, T_i), (\mathbf{m}_j, T_j) \rightarrow (\mathbf{m}_j, T_i), (\mathbf{m}_i, T_j). \quad (19)$$

Only the models collected at $T = 1$ are considered in the computation of the posterior model because the models collected at $T > 1$ sample a biased posterior *pdf*.

To further increase the efficiency of the implemented MCMC algorithm we employ a delayed rejection strategy. Theoretically the characteristics of the proposal distribution does not influence the final estimated posterior *pdf* if an infinite number of MCMC steps is employed. However, the variance of the proposal distribution should be set in accordance to the expected spread of the target *pdf*: a suboptimal variance of the proposal generates persistent rejections of models, thus resulting in an increased computational cost and in a slower convergence rate. For this reason, it is advisable changing the variance of the proposal distribution when the target *pdf* is expected to have different spreads across different model space dimensions (Bodin and Sambridge, 2009). In the case of AVA inversion, we expect the posterior *pdf* having different spreads along the *Vp*, *Vs* and density directions due to the different resolvability of these parameters. In our implementation after a rejected model $\mathbf{m}'$, a second attempt is made with a different proposal q_2 that is independent from the previously rejected model generated according to the proposal q_1. In this context the acceptance probability can be derived as:

$$p(\mathbf{m}''|\mathbf{m}) = \min\left[1, \frac{p(\mathbf{m}''|\mathbf{d})}{p(\mathbf{m}|\mathbf{d})} \frac{q_1(\mathbf{m}'|\mathbf{m}'')}{q_1(\mathbf{m}'|\mathbf{m}')} \frac{(1-p(\mathbf{m}'|\mathbf{m}''))}{(1-p(\mathbf{m}'|\mathbf{m}))}\right]. \quad (20)$$

In practice we derive $\mathbf{m}''$ by perturbing $\mathbf{m}$ according to a second proposal distribution $\boldsymbol{q_2}$ characterized by a reduced variance with respect to the first proposal $\boldsymbol{q_1}$.

Let $\boldsymbol{\pi}$ and $\mathbf{e}$ be the facies and the elastic properties. In our case of a mixed discrete-continuous inverse problem, the posterior *pdf* can be written as:

$$p(\boldsymbol{\pi}, \mathbf{e}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{e}, \boldsymbol{\pi})\boldsymbol{p}(\mathbf{e}|\boldsymbol{\pi})\boldsymbol{p}(\boldsymbol{\pi})}{p(\mathbf{d})} = \frac{p(\mathbf{d}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{d})}, \quad (21)$$

where $\mathbf{m} = [\mathbf{e}, \boldsymbol{\pi}]$. Before the MCMC inversion, we exploit the available borehole data and/or the available geological information about the investigated area to define the $p(\boldsymbol{\pi})$ and $p(\mathbf{e}|\boldsymbol{\pi})$ distributions; The former is the prior distribution of facies, whereas the latter is the prior distribution of the elastic properties $\mathbf{e}$ within each considered facies. In our implementation $p(\mathbf{e}|\boldsymbol{\pi})$ is a non-parametric mixture distribution that is directly derived from the available data (e.g. well log data) by means of the kernel density estimation algorithm. Then, we apply a normal score transformation to convert each non-parametric component of the prior to a Gaussian model, thus deriving the $p(\mathbf{z}|\boldsymbol{\pi})$ distribution where $\mathbf{z}$ represents the normal-score transformed elastic properties. After this transformation the prior $p(\mathbf{z}|\boldsymbol{\pi})$ is a Gaussian mixture model from which we extract the mean vector and the covariance matrix of each component and the variogram model expressing the expected lateral or vertical variability of elastic parameters. From the one hand, the transformation to a Gaussian mixture model allows for an easy inclusion of geostatistical constraints into the MCMC sampling in the form of a variogram model (Doyen, 2007). From the other hand, this additional geostatistical a-priori information is used to attenuate the ill-conditioning of the AVA inversion. In other terms, the aim of our strategy is two-fold: preserving the geostatistical constraints in each model sampled during the MCMC iterations and generating the starting models using common and computationally fast geostatistical simulation techniques such as the truncated Gaussian simulation (Matheron et al. 1987) or the fast-Fourier-transform moving average (Ravalec et al. 2000).

The steps of our MCMC recipe are the following:

1. Select a given CMP position (for the target-oriented inversion) or time position (for the interval-oriented approach). To speed up the convergence of the algorithm this position is defined on the basis of the current fit between observed and predicted data. More in detail, let us consider the 1D inversion and let $\mathbf{d_{pre}}$ and $\mathbf{d_{obs}}$ be the predicted and observed CMP gathers (with offset distances converted into incidence angles). The probability of selecting the *i*-th time sample is given by:

$$p(i) = \frac{\sum_{q=1}^{Q}(\mathbf{d_{obs},}_{q} - \mathbf{d_{pre},}_{q})^2}{\sum_{l=1}^{L}\sum_{q=1}^{Q}(\mathbf{d_{obs},}_{q,l} - \mathbf{d_{pre},}_{q,l})^2}, \quad (22)$$

   where *Q* and L are the total numbers of incidence angles and time samples considered, respectively;

2. For even iteration numbers:

2.1 Draw a random number κ uniformly distributed over [0,1] and set the ξ parameter defined over [0,1]. Then;

2.2 *Facies Move*. If $\boldsymbol{\kappa \leq \xi}$, perturb the current facies model $\boldsymbol{\pi}$ at the selected position *i*, thus deriving $\boldsymbol{\pi}'$. For the interval-oriented and the target-oriented inversions the probability of replacing the current facies at the selected time position depends on the previously defined vertical and lateral transition matrices (see discussion below). Then, the new elastic properties at the *i*-th position are random realizations from the Gaussian distribution $p(\mathbf{z}|\,\boldsymbol{\pi}')$.

2.3 *Elastic Move*. If $\boldsymbol{\kappa > \xi}$, preserve the current facies at the *i*-th position (so that $\boldsymbol{\pi}' = \boldsymbol{\pi}$) and perturb only the elastic properties $\mathbf{z}$ at the selected position. This perturbation follows a Gaussian proposal with a zero mean and a previously defined covariance matrix $N(\mathbf{z}; 0, \boldsymbol{\sigma_o})$.

3. For odd iteration numbers:

3.1 *DEMC move*. Let *p* be the index of the current chain and *i* the time sample or the spatial position to be perturbed. The algorithm selects other two chains with indexes *a* and *b* that

share with the chain *p* the same facies at the *i*-th position: Then, the perturbed elastic model is defined as follows:

$$\mathbf{z'}_{p,i} = \mathbf{z}_{p,i} + \beta\left(\mathbf{z}_{a,i} - \mathbf{z}_{b,i}\right) + \delta N(\mathbf{z}; 0, \boldsymbol{\sigma}_{\mathbf{o}}) \quad with\ a \neq b \neq p, \qquad (23)$$

where β is the jump rate (see Vrugt, 2016) and δ is a random number <1. If there are no chains *a* and *b* that share with the chain *p* the same facies at the *i*-th position, move to step 2.1;

4. Use a previously defined variogram model and a kriging interpolation to propagate the perturbation of the elastic properties at the *i*-th position over the neighboring CMP positions or time samples (Doyen, 2007). This step is crucial to ensure that all the sampled elastic models honor the a-priori geostatistical constraints.

5. Apply the inverse normal score transformation to $\mathbf{z}'$ to derive $\mathbf{e}'$;

6. Use the Zoeppritz equations to compute the data predicted on the proposed elastic model $\mathbf{e}'$. Accept or reject the proposed elastic model according to the Metropolis-Hasting rule. In particular, the *facies move* results in an acceptance probability equal to:

$$p(\mathbf{e}'|\mathbf{e}) = \min\left[1, \frac{p(\mathbf{d}|\mathbf{e}')^{1/T_p}}{p(\mathbf{d}|\mathbf{e})^{1/T_p}}\right], \qquad (24)$$

where $\boldsymbol{T_p}$ is the temperature level of the considered *p*-th chain. The elastic move and the DEMC move, give an acceptance probability equal to:

$$p(\mathbf{e}'|\mathbf{e}) = \min\left[1, \frac{p(\mathbf{d}|\mathbf{e}')^{1/T_p}}{p(\mathbf{d}|\mathbf{e})^{1/T_p}} \frac{p(\mathbf{e}')}{p(\mathbf{e})}\right], \qquad (25)$$

7. If the perturbed elastic model is rejected come back to the current model $\mathbf{e}$ and apply the delayed rejection scheme to derive $\mathbf{e}''$ (see equation (20));

8. If the perturbed model is accepted set $\mathbf{e} = \mathbf{e}'$ (or $\mathbf{e} = \mathbf{e}''$) and $\boldsymbol{\pi} = \boldsymbol{\pi}'$.

9. Collect $\mathbf{m} = [\mathbf{e}, \boldsymbol{\pi}]$ from the chains with T=1;

10. Select two chains with indices p and q and apply the parallel tempering strategy (equations (18) and (19)):

11. Apply the normal score transformation to the current elastic model $\mathbf{e}$ and derive $\mathbf{z}$;

12. Repeat from 1 to 11 until the maximum number of iteration is reached;

13. Considering the selected burn-in phase and the lag value, use the ensemble of accepted models to compute the statistical properties of the posterior *pdf* (e.g. maximum-a-posteriori "MAP" solution or mean model) or to visualize the marginal posterior *pdf* along different model dimensions.

All the user-defined parameters (e.g. $\boldsymbol{\sigma}_{\mathbf{0}}$, ξ, β) are properly set in order to ensure an acceptance ratio between 0.2-0.4 for the chains at T=1. In particular ξ should be usually lower than 0.5 because the sampling space of a continuous property is larger than the sampling space of the underlying discrete property (Holmes et al. 2017).

To decrease the correlation between consecutively sampled models we can consecutively perturb the properties at different CMP positions (for the target-oriented inversion) or time positions (for the interval-oriented approach), that is we can iterate the steps from 1 to 4 before the normal score transformation and the likelihood evaluation (steps 5 and 6; see Aleardi et al. 2018a). This strategy increases the likelihood variation between the current and the proposed model with respect to the strategy described previously in which only one CMP or time position is perturbed at a time. This approach permits the reduction of the lag parameter, but at the expense of a decreasing of the acceptance rate.

To promote the lateral continuity of the discrete property, we employ a first order Markov model simulation during the MCMC sampling. In particular, for the target-oriented inversion the probability of a transition from facies b to facies a at the spatial position with horizontal coordinates x and y is given by:

$$p\left(\boldsymbol{\pi}_{x,y}^{a}\middle|\boldsymbol{\pi}_{x-1,y}^{h},\boldsymbol{\pi}_{x+1,y}^{h},\boldsymbol{\pi}_{x,y+1}^{h},\boldsymbol{\pi}_{x,y-1}^{h}\right) = p\left(\boldsymbol{\pi}_{x-1,y}^{h}\middle|\boldsymbol{\pi}_{x,y}^{a}\right)p\left(\boldsymbol{\pi}_{x+1,y}^{h}\middle|\boldsymbol{\pi}_{x,y}^{a}\right)p\left(\boldsymbol{\pi}_{x,y-1}^{h}\middle|\boldsymbol{\pi}_{x,y}^{a}\right)p\left(\boldsymbol{\pi}_{x,y+1}^{h}\middle|\boldsymbol{\pi}_{x,y}^{a}\right), \quad (26)$$

where the coordinates $(x-1,y)$, $(x+1,y)$, $(x,y-1)$ and $(x,y+1)$ correspond to the first four neighboring CMPs that surround the considered CMP, whereas the subscript h indicates the actual facies configuration at each neighboring CMP. The probability to move from facies b at one neighboring position to the facies a at position (x,y) can be derived from the b-th row and a-th column of the lateral transition matrix $\mathbf{T}^{l}$. For example:

$$p\left(\boldsymbol{\pi}_{x-1,y}^{b}\middle|\boldsymbol{\pi}_{x,y}^{a}\right) = \mathbf{T}_{b,a}^{l}. \quad (27)$$

Similarly, the 1D interval-oriented inversion includes a first-order Markov model to define the probability of a transition from facies b at the time sample i-1 to facies a at the i-th time position. This probability is given by the vertical transition matrix $\mathbf{T}^{v}$:

$$p(\boldsymbol{\pi}_{i}^{a}|\boldsymbol{\pi}_{i-1}^{b}) = \mathbf{T}_{b,a}^{v}. \quad (28)$$

As previously mentioned, determining the convergence of the sampling to a stable posterior model is a crucial aspect of any MCMC inversion. We discuss this aspect in more details in the following examples where we analyse the potential scale reduction factor (PSRF) factor and the autocorrelation of the sampled models.

INVERSION RESULTS

Target-oriented inversions

In this case the reference 2D model simulates a slice from a stratigraphic grid and is derived from a geostatistical simulation driven by actual well log data and geological information pertaining to a clastic, gas saturated reservoir. The well log data will be also used to derive the a-priori model in the following inversion procedure.

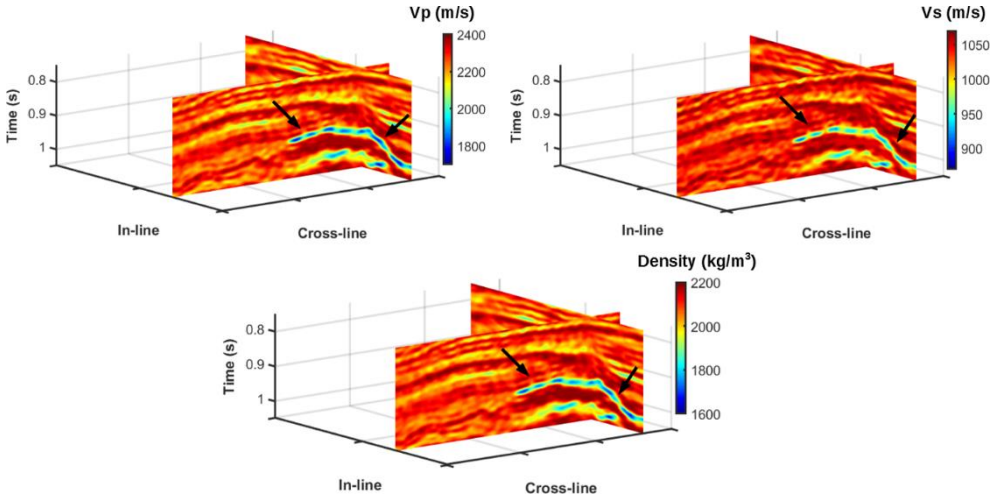

Figure 1: Example of in-line and cross-line sections extracted from the simulated 3D models of Vp, Vs and density. From this 3D model we extract the 2D reference model for the target-oriented inversion, which represents the elastic reflectivity contrasts at the top of the target, gas-saturated layer (indicated by the black arrows).

We employ the truncated Gaussian simulation algorithm (Matheron et al. 1987) to derive the 3D discrete facies model, whereas the actual well log data distribution and the fast-Fourier-transform moving average (Ravalec et al. 2000) are used to distribute the elastic properties within each facies. In Figure 1 we show an in-line and cross-line section extracted from the simulated 3D *Vp*, *Vs* and density models. In this figure the black arrows identify the considered gas-saturated sand layer that is characterized by significant decreases of *Vp*, *Vs* and density, with respect to the encasing shales. From the so derived 3D model we extract the elastic properties around the top of the considered reservoir interval. The elastic reflectivity contrasts along the top of the reservoir are represented in Figure 2. In this case the lowest and highest R_{Vp}, R_{Vs}, R_{ρ} values correspond to gas-saturated sand and shale, respectively. The elastic reflectivity contrasts at the target and equation (1) are used to

analytically compute the observed amplitude versus angle (AVA) responses to which we add Gaussian random noise to better simulate a field dataset. The so derived AVA responses constitute the observed data for the following inversion procedure.

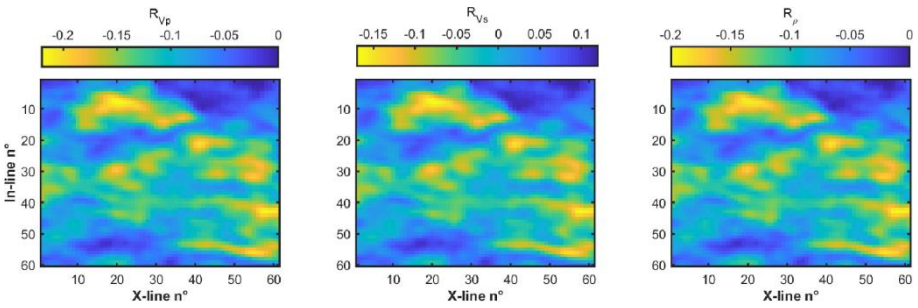

Figure 2: The reference model for the target-oriented inversion that represents the elastic reflectivity contrasts at the top of the considered reservoir interval (see also Figure 1). From left to right we represent R_{Vp}, R_{Vs}, R_{ρ}.

Figure 3 compares the a-priori marginal non-parametric and Gaussian mixture distributions for the elastic parameters (R_{Vp}, R_{Vs}, R_{ρ}). We observe a close similarity between the non-parametric and Gaussian-mixture distributions. Figure 4 shows normal probability plots for each elastic reflectivity and for each considered litho-fluid facies derived from the available well log information (used to simulate the model of Figure 1) and computed before and after the normal score transformation (Figure 4a and 4b, respectively). Figures 3 and 4 prove that, although some minor deviations from the theoretical Gaussian distribution (e.g. skewness) especially in the gas sand, the actual distribution of the elastic reflectivities within each facies is very close to a Gaussian model. In other terms, these figures prove that in this context the Gaussian-mixture assumption should constitute an acceptable compromise between the accuracy of the final predictions and the computational cost requested for a reliable assessment of the posterior model. For this reason, the aim of the following inversion tests is two-fold: demonstrate the suitability of the implemented Markov Chain Monte Carlo (MCMC) approach for target-oriented AVA inversion; prove that the lateral constraints included into the

MCMC inversion provide more realistic and stable predictions in case of low signal-to-noise (S/N) ratios.

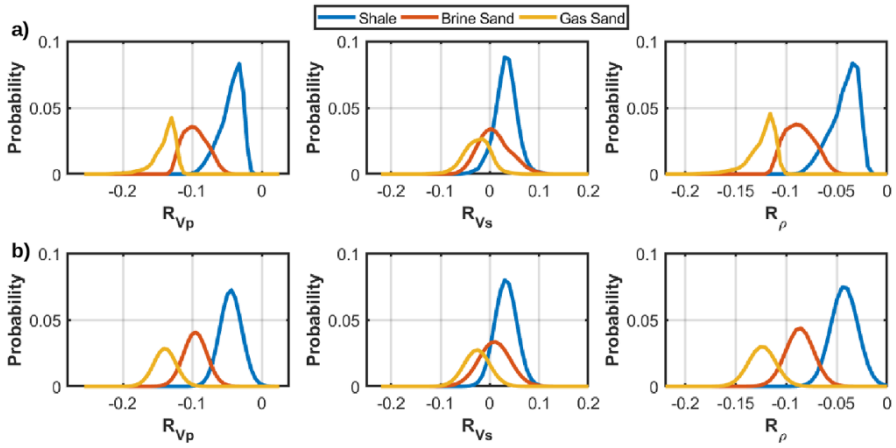

Figure 3: A-priori marginal distributions for the elastic reflectivities within each litho-fluid facies derived from the available well log information. a) The a-priori non-parametric distributions derived through the kernel density estimation algorithm. b) The Gaussian-mixture prior model.

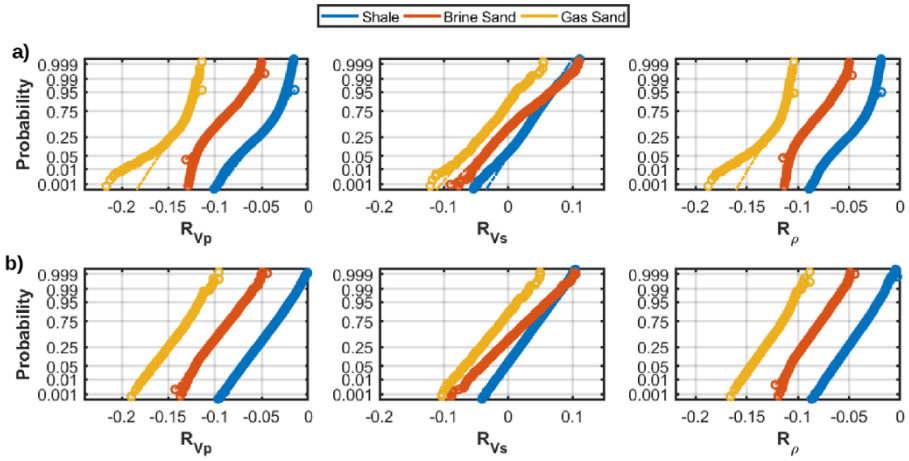

Figure 4: a) Normal probability plots for each elastic reflectivity and litho-fluid facies derived from the available well log information around the target. b) Normal probability plots derived after the normal-score transformation of a). In these plots the dotted lines represent the theoretical Gaussian distribution, whereas the circles represent the actual data.

Figure 5 illustrates the actual autocorrelation functions computed on the reference model along the in-line and cross-line directions, together with the theoretical autocorrelation functions computed by

assuming a Gaussian variogram model. The inclusion of this theoretical variogram model into the MCMC inversion imposes lateral constraints to the elastic models sampled by the MCMC algorithm and ensures more stable results in case of severe noise contamination of the observed data. In case of field data inversion, the lateral amplitude variability of the seismic data, integrated by the available well log information and by the geological knowledge of the investigated area, can be used to define the autocorrelation functions.

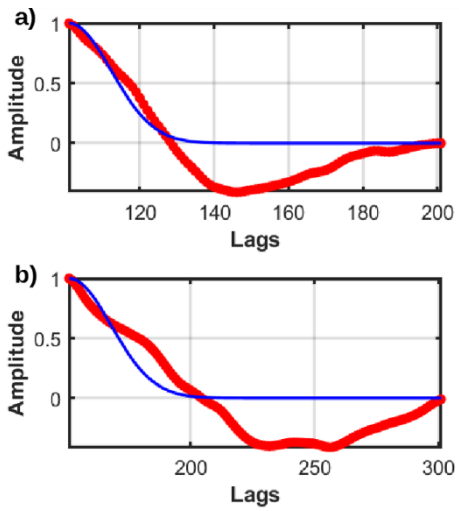

Figure 5: Actual (red lines) and theoretical (blue lines) lateral autocorrelation functions for the cross-line and in-line directions (parts a) and b), respectively)

For the MCMC inversion we use 40 parallel and interactive chains, and according to Dosso et al. (2012) we choose logarithmic temperature spacing between chains. In particular, we employ 20 chains at T=1 and the remainder with temperatures logarithmically distributed in the range $10 < T \leq 500$. The upper temperature limit is set to a value sufficiently large to ensure an acceptance ratio around to 0.5-0.6 for the corresponding chain. To decrease the correlation between the current and the proposed model we consecutively perturb the properties at 20 different common-mid-point (CMP) gather positions before the likelihood evaluation. The number of MCMC iterations is set to 400000 with a burn-in period of 200000 and a lag of 30. These values have been determined from the evolution of the likelihood values of the different chains, from the evolution of the potential scale

reduction factor (PSRF) for different model parameters and from the inspection of the autocorrelation function for different model parameters (see the next section).

For the comparability of the results, both the analytical and MCMC inversions consider the same angle range between 0 and 30 degrees. However, note that the Zoeppritz equations make it possible for the MCMC considering a more extended angle range. The lateral transition matrix employed by the MCMC algorithm along both the in-line and cross-line directions is the following:

$$\mathbf{T}^l = \begin{bmatrix} 0.6 & 0.2 & 0.2 \\ 0.2 & 0.6 & 0.2 \\ 0.2 & 0.2 & 0.6 \end{bmatrix}. \quad (29)$$

If we consider the *i*-th spatial position, the rows (from top to bottom) represent shale, brine sand and gas sand at one neighbouring position *i*-1, whereas the columns (from left to right) represent shale, brine sand and gas sand at the *i*-th position. The higher values along the main diagonal of $\mathbf{T}^l$ preserve the continuity of litho-fluid facies in the sampled models. This matrix can be inferred from available geological information and from the lateral variability of the available data. However, a quality control of the inversion results (i.e. comparison between inverted properties and well log data) and slight modifications of the so derived transition matrices are usually needed to obtain optimal predictions.

In the first inversion example we simulate an optimal signal to noise ratio equal to 50 in the observed data. In this favourable scenario we expect that the analytical and MCMC target-oriented inversions yield very similar predictions. Indeed, Figure 6 demonstrates that both approaches provide comparable elastic a-posteriori mean model. In addition, the comparison of Figure 7 and Figures 8 show that the analytical and MCMC inversion also achieve very similar maximum-a-posteriori (MAP) solutions and posterior probability density functions *(pdfs)* in the facies classification. However, even in this favourable scenario with low noise and almost Gaussian-mixture-distributed elastic reflectivities in the reference model, the MCMC approach provides superior facies classification results. In particular, note that the analytical approach erroneously predicts brine sand

between cross-line 40-50 and in-line 30-50 (green circles in Figures 7a-b). Differently, the MCMC approach correctly identifies some small and circular shale bodies at these spatial locations (green circles in Figures 8a-b).

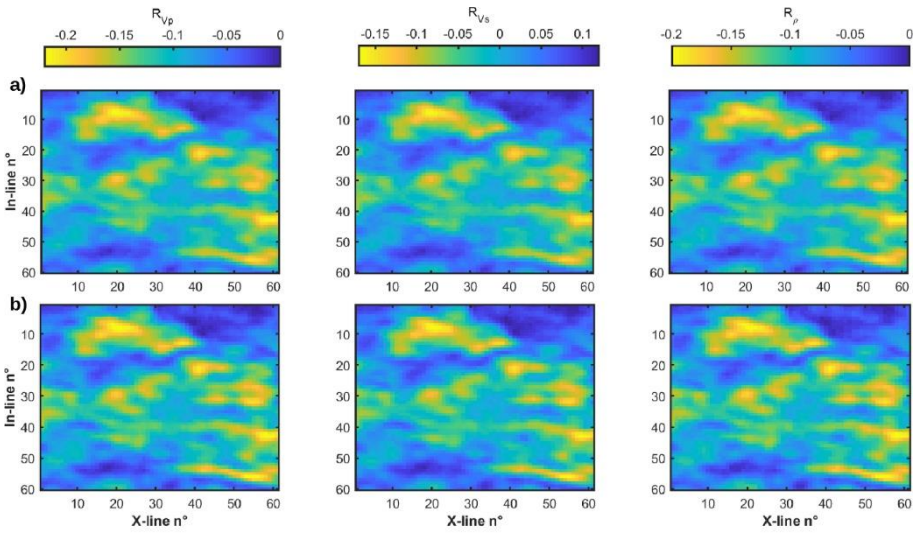

Figure 6: Results for a high S/N ratio equal to 50. a) Mean a-posteriori models predicted by the analytical inversion. b) Mean a-posteriori models predicted by the MCMC inversion. In a) and b) R_{Vp}, R_{Vs}, R_{ρ} *are represented from left to right.*

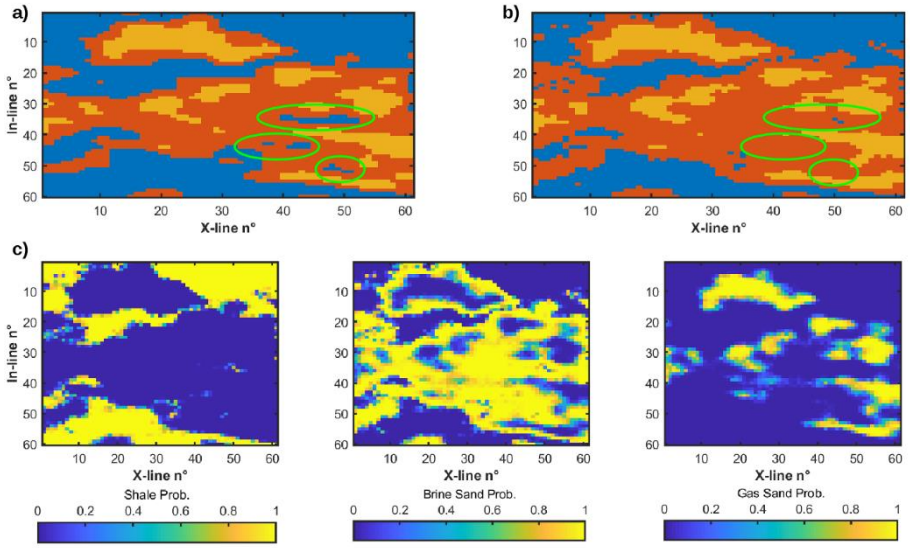

Figure 7: Facies classification results provided by the analytical inversion for a high S/N ratio equal to 50. a) True facies model. b) MAP facies solution. c) Posterior probability for shale (left), brine

sand (center) and gas sand (right). The circles in (a) and (b) highlight some examples of misclassifications.

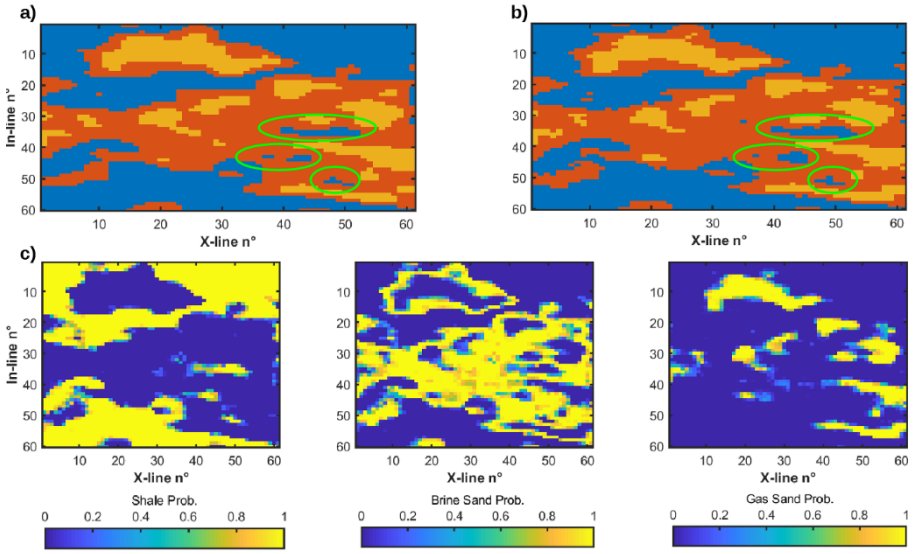

Figure 8: Facies classification results provided by the MCMC inversion for a high S/N ratio equal to 50. a) True facies model. b) MAP facies solution. c) Posterior probability for shale (left), brine sand (center) and gas sand (right). The circles in (a) and (b) highlight the same areas evidenced in Figure 7. Here, note the superior classification results provided by the MCMC method.

In the next example we decrease the S/N ratio to 2. In this less favourable scenario, we expect that the lateral constraints for both the continuous and the discrete properties employed by the MCMC algorithm will provide more stable and reliable results than those yielded by the analytical approach where the lateral parameter correlation is neglected. Figure 9a shows that the analytical approach now provides very scattered estimates of the elastic reflectivities, and final 2D models where the lateral formation boundaries are difficult to identify. This scattering is related to the pointwise nature of the inversion that does not account for the spatial correlation of the elastic parameters and facies distribution. Differently, the MCMC approach preserves the actual lateral variability of the R_{Vp}, R_{Vs}, R_{ρ} properties and retrieves stables estimates (Figure 9b). In Figures 10 and 11, we observe that the overall match between the true and the final facies models is decreased with respect to the previous inversion with high S/N ratio. However, the comparison of Figure 10 and Figure 11 again

confirms that the MCMC inversion provides much less scattered and more stable predictions in which the actual lateral continuity of the discrete property is preserved.

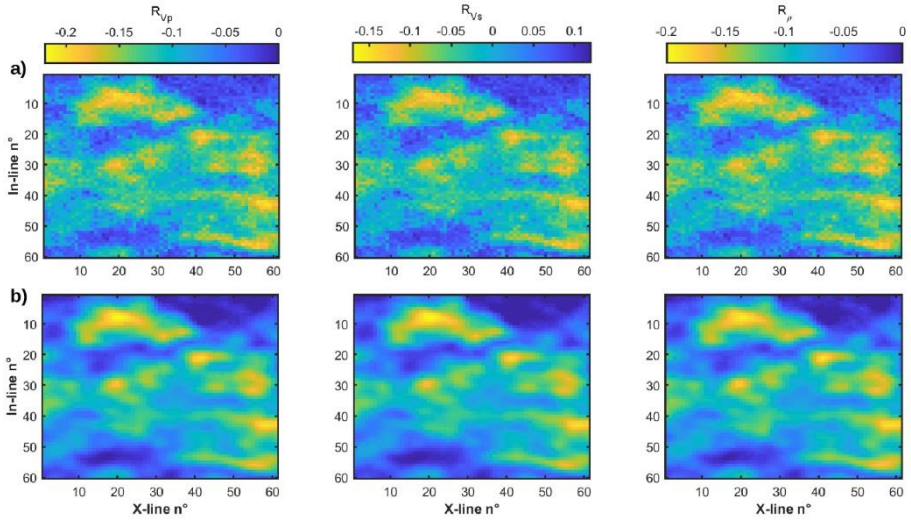

Figure 9: As in Figure 6 but for a S/N ratio equal to 2. a) and b) refer to the analytical and MCMC estimated mean models, respectively. The R_{Vp}, R_{Vs}, R_{ρ} are represented from left to right.

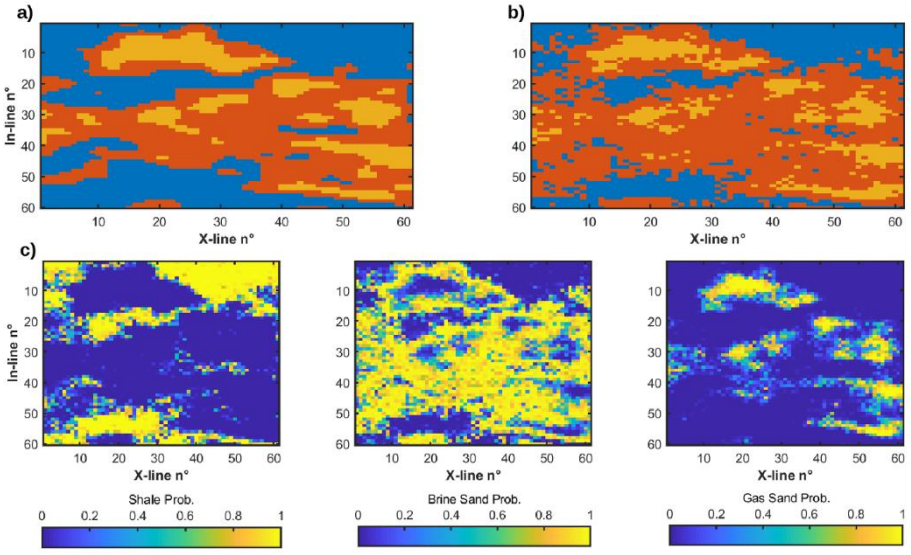

Figure 10: Facies classification results provided by the analytical inversion for a S/N ratio equal to 2. a) True facies model. b) MAP facies solution. c) Posterior probability for shale (left), brine sand (center) and gas sand (right).

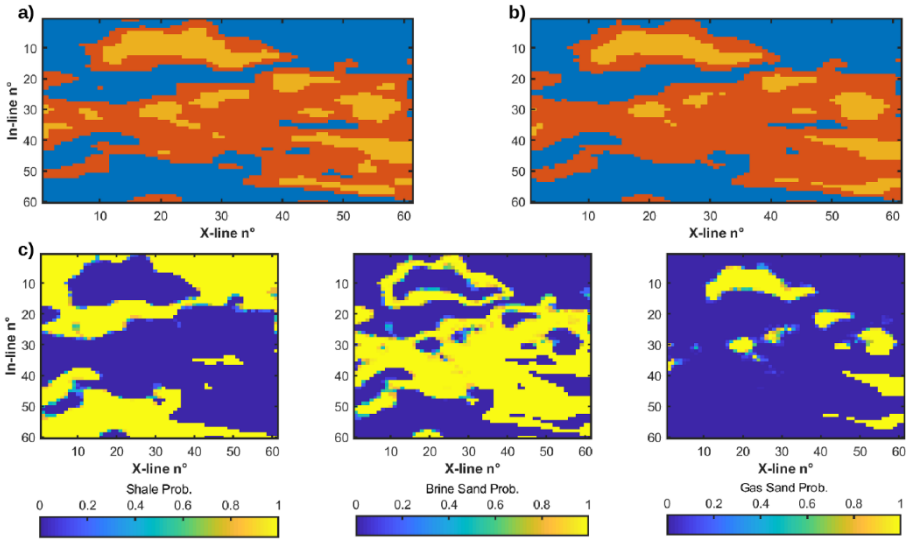

Figure 11: Facies classification results provided by the MCMC inversion for a S/N ratio equal to 2. a) True facies model. b) MAP facies solution. c) Posterior probability for shale (left), brine sand (center) and gas sand (right).

Figure 12 shows comparisons for two CMP gathers of the observed AVA responses and the AVA responses computed on the mean a-posteriori solutions provided by the analytical and the MCMC inversions. In both cases we note a satisfactory match between the predicted and the observed data. However, if we compute the L2 norm data misfit we get for the analytical approach values of 0.0661 (CMP on the left of Figure 12a) and 0.0742 (CMP on the right of Figure 12a), whereas from the MCMC predictions we obtain values of 0.0689 (CMP on the left of Figure 12b) and 0.0756 (CMP on the right of Figure 12b). The higher data misfits for the MCMC inversions are related to the inclusions of the lateral constraints that act as regularization terms that stabilize the solution at the expense of an increased data misfit value.

Finally, some considerations about the computational costs. The analytic inversion takes less than 5 seconds if we consider a serial Matlab codes running an Intel i7-7700HQ@2.8GHz with 16 Gb RAM. For the MCMC inversion we implement a parallel Matlab code that runs in approximately 40 minutes on two compute nodes equipped with two deca-core Intel E5-2630 @2.2 GHz (128 Gb RAM). It is

clear that the MCMC inversion retrieves highly accurate uncertainty appraisals but at the expense of an increased computational effort. We return to this aspect in the discussion section.

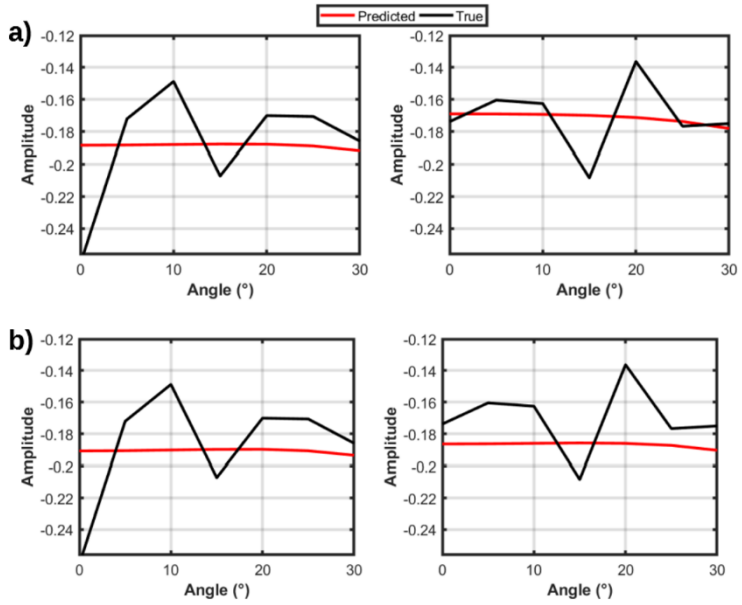

Figure 12: Examples of observed AVA responses (black lines) and predicted AVA responses (red lines) computed on the MAP solutions provided by the analytical inversion (a) and the MCMC inversion (b) in case of S/N=2.

Convergence diagnostic for the implemented target-oriented algorithm

We now discuss in more detail the convergence of the implemented MCMC algorithm for target-oriented AVA inversion. Here we limit to the second MCMC example with S/N ratio equal to 2. Figure 13 shows examples of evolutions of the negative log-likelihood for 5 out of 40 chains. We note that the combined use of parallel tempering and delayed rejection, guarantee optimal exploration and exploitation capabilities and makes the algorithm rapidly converge toward the most promising portions of the model space. Indeed, the exploitation is guaranteed by the chains at low temperature, while the exploration is performed by the chains at high temperature. In particular, the black arrows in Figure 13, highlight the beneficial effects produced by swaps of models between different chains. These swaps produce an instantaneous decrease of the negative log-likelihood value and allow the chains to escape from local optima of the posterior *pdf*. Note that all the considered chains attain a

stable negative log-likelihood value after approximately 200000 iterations. This value corresponds to the selected burn-in period.

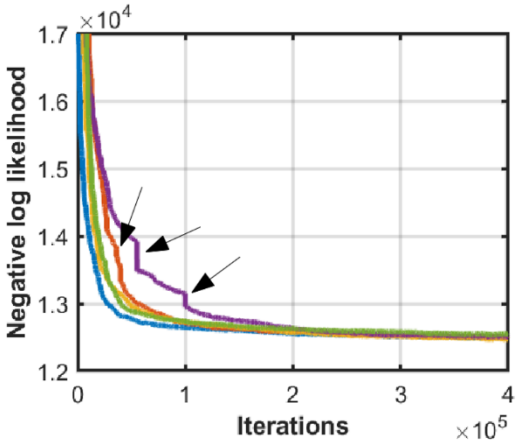

Figure 13: Examples of evolutions of the negative log-likelihood for 5 out of 40 interactive chains. The black arrows highlight the beneficial effect of the PT strategy. Different colors represent different chains.

To assess the convergence of the algorithm we monitored the evolution of the potential scale reduction factor (PSRF) for different model parameters (elastic reflectivity estimated at different CMP locations). Examples of PSRF values for 6 CMP positions are represented in Figure 14. We note that only 20000 iterations, approximately, are needed to attain reliable posterior *pdf* estimations for R_{Vp}, whereas 200000 iterations are usually requested to accurately estimate the posterior *pdf*s for R_{Vs} and R_{ρ}. These different convergence rates are obviously related to the different influence that the elastic parameters exert on the observed data. Indeed, the R_{Vp} plays the major role in determining the observed P-wave reflection coefficients, while R_{Vs} and R_{ρ} exert weaker influences.

Finally, Figure 15 represents some examples of normalized autocorrelation functions computed on the sampled R_{Vp}, R_{Vs}, and R_{ρ} models. To set the lag value of 30 used in the previous inversion tests, we simply count the number of consecutive models requested to obtain autocorrelation values below 0.5-0.4. This threshold value is assumed to be a good compromise between reliable posterior assessments and the overall computational cost of the inversion procedure. A lower threshold value

would guarantee more accurate *pdf* estimations but at the expense of an extra computational cost requested to sample a number of models sufficient enough to reliably evaluate the posterior. The analysis of the evolution of the negative log-likelihood value, of the PSRF, and of the autocorrelation function is crucial to determine the number of MCMC iterations requested to attain accurate posterior *pdf* estimations.

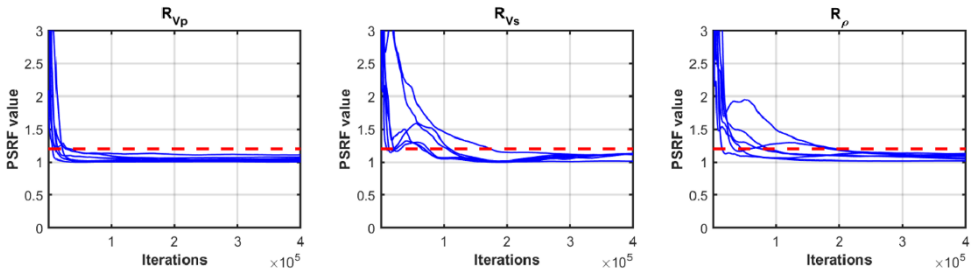

Figure 14: Example of evolution of the PSRF value for the elastic reflectivities pertaining to 6 CMP gathers. The dotted red lines show the desired PSRF value of 1.2. In each plot, the blue lines pertain to model parameters estimated at different CMP gather positions.

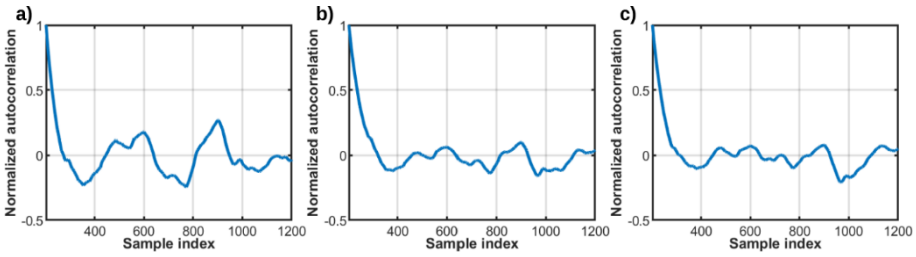

Figure 15: Close-up of the normalized autocorrelation functions derived from the R_{Vp} (a), R_{Vs} (b), and R_{ρ}(c) vales sampled at a given CMP gather position.

Interval-oriented inversions

In this case we consider actual well-log data pertaining to 7 wells investigating a gas-saturated clastic reservoir located in a shale-sand sequence. Figure 16 shows the a-priori non-parametric and the Gaussian-mixture marginal distributions for each elastic property derived from 5 out of 7 wells investigating the reservoir interval. The remaining two wells are used as blind tests in the following inversion examples. This characteristic makes the following inversion tests more realistic with respect

to the previous target-oriented examples. Indeed, now the prior is not directly derived from the information used to define the true model but is derived from a subset of the available well log data while the remaining well data are used as blind tests. The elastic properties pertaining to the two blind wells and the forward modelling of equation (12) are used to compute the CMP gathers. To better simulate a filed dataset, we contaminate the observed data with Gaussian uncorrelated noise. We employ a 55-Hz Ricker wavelet as the source signature that is assumed known during the inversion, whereas the angle range is [0, 30] degrees. Differently from the previous example (Figures 3 and 4), we now note some important differences between the two non-parametric and Gaussian-mixture distributions. In particular, the distributions are still very similar for shale, but significatively different for brine sand and gas sand where the non-parametric distribution shows skewness or even multimodalities. Obviously, the Gaussian-mixture model does not capture these characteristics and for this reason it constitutes an oversimplified statistical model in this context. For this reason, we expect the MCMC inversion outperforms the analytical approach. These considerations are confirmed by the normal probability plots of Figure 17a where we observe significative deviations from the Gaussian model for *Vp* and *Vs* in the brine sand and gas sand. These deviations disappear after the normal score transformation (Figure 17b). For comparability in both Figures 16 and 17 we plot the natural logarithm of the elastic parameters.

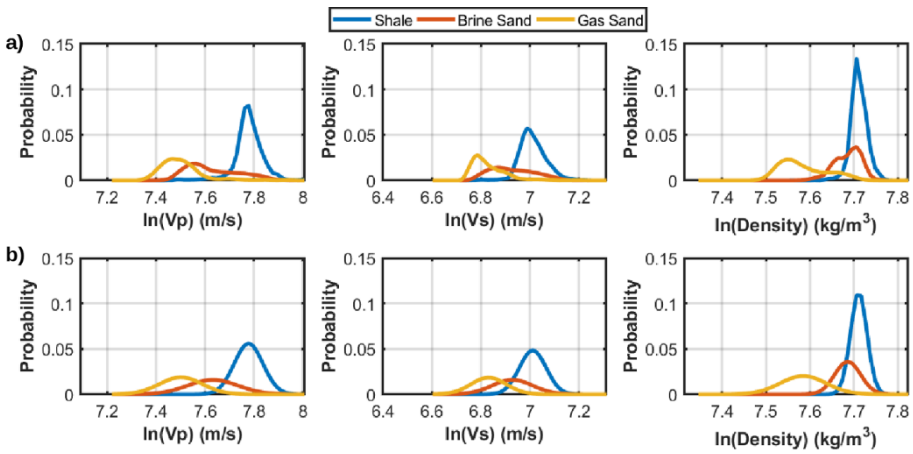

Figure 16: A-priori models for the elastic parameters and for each litho-fluid facies derived from 5 out of 7 available wells. a) The a-priori non-parametric model derived through the kernel density

estimation algorithm. b) The Gaussian-mixture prior model. In both a) and b) we consider the natural logarithm of the elastic properties.

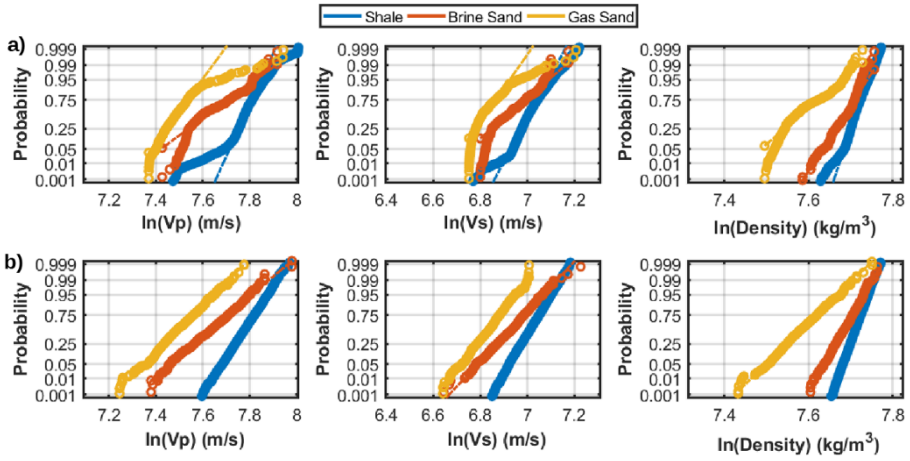

Figure 17: a) Normal probability plot derived from the actual well log data pertaining to 5 out 7 available wells. b) Normal probability plot derived on the normal score transformed actual well log data. In both a) and b) we consider the natural logarithm of the elastic properties. In these plots the dotted lines represent the theoretical Gaussian distribution, whereas the circles represent the actual well log data.

In the example on the first blind well we simulate a S/N ratio of 10 in the observed data, whereas this ratio is decreased to 2 in the second blind well example. In the MCMC inversion we use 40 different chains running for 10000 iterations each and with a burn-in period of 5000: 20 chains run at T=1, while the remainder at logarithmically spaced temperature values. We consecutively perturb the elastic properties at ten different time positions before the likelihood evaluation and we set a lag value of 20. The vertical transition matrix employed by the MCMC approach is the following:

$$\mathbf{T}^{v} = \begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0.3 & 0.7 & 0 \\ 0.1 & 0.4 & 0.5 \end{bmatrix}. \quad (30)$$

Note the null transition probability from a brine sand above to a gas sand below. As for the target-oriented approach this matrix can be defined on the basis of available well log data and properly modified to avoid unphysical transitions (from example from a brine sand above to a gas sand below).

Figures 18 and 19 compare the results provided by the analytical and MCMC inversions for the first blind well. We observe that both approaches correctly capture the vertical variability of the elastic parameters and yield a predicted data that closely matches the observed one. The good agreement between the final estimates provided by the two methods seems to confirm the suitability of the Gaussian-mixture assumption also in this case. For this reason, for a more quantitative assessment of the quality of the analytical and MCMC predictions we compute the coverage probability that is the actual probability that a given probability interval (in the following the 0.80 probability interval defined around the mean a-posteriori model) contains the true property value. The coverage probabilities for the three elastic properties provided by two inversion methods are given in Table 1. We observe that the MCMC method yields higher coverage probability values for *Vp* and density, thus confirming the superior quality of the prediction intervals provided by this method with respect to those yielded by the analytical approach. The facies classification results are again very similar, although the 1D first-order Markov model employed by the MCMC method avoids unphysical transitions in the MAP facies solution such as the transition from a brine to a gas saturated sand predicted by the analytical approach around 937 ms.

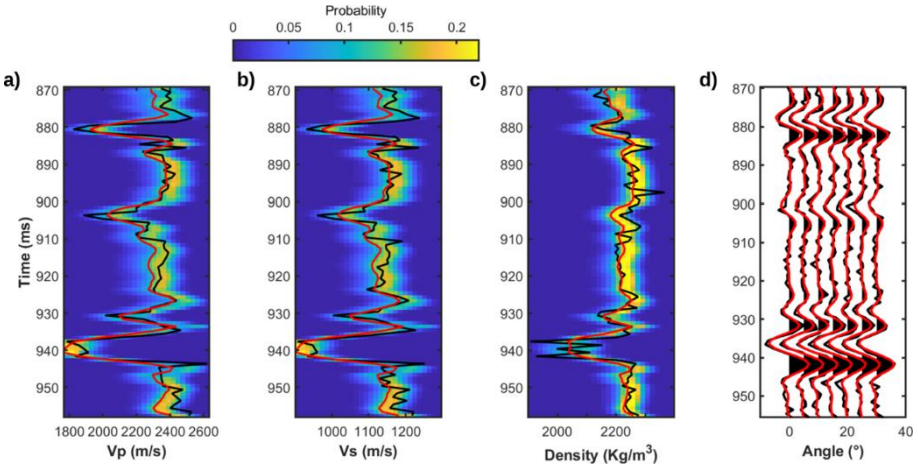

Figure 18: Results provided by the analytical inversion for the first blind well test. In a), b) and c) the black lines represent the true property values, the red lines are the estimated mean models, whereas the colormap codes the estimated posterior pdf. d) Comparison of observed (black) and predicted (red) seismic data computed on the a-posteriori mean model.

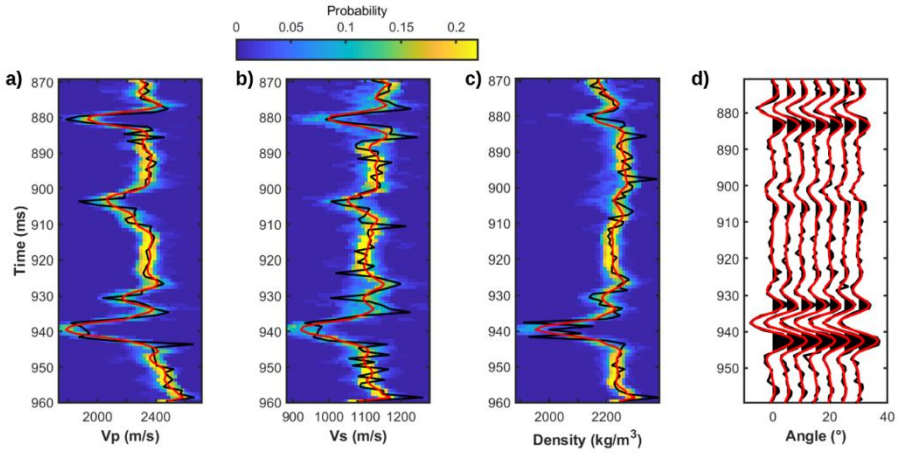

Figure 19: As in Figure 14 but for the MCMC inversion.

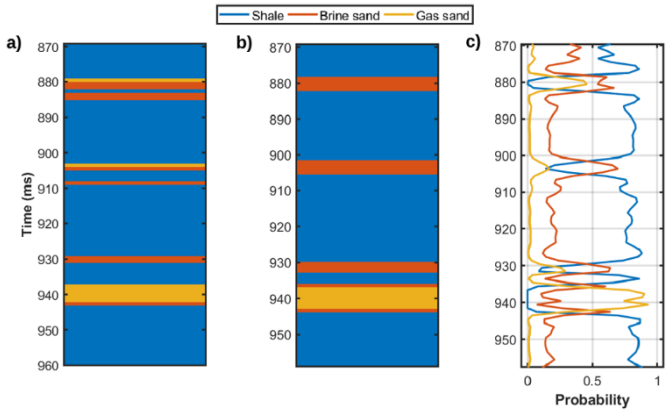

Figure 20: Facies classification results provided by the analytical inversion for the first blind test. a) True facies profile. b) MAP facies solution. c) Estimated posterior pdf of facies.

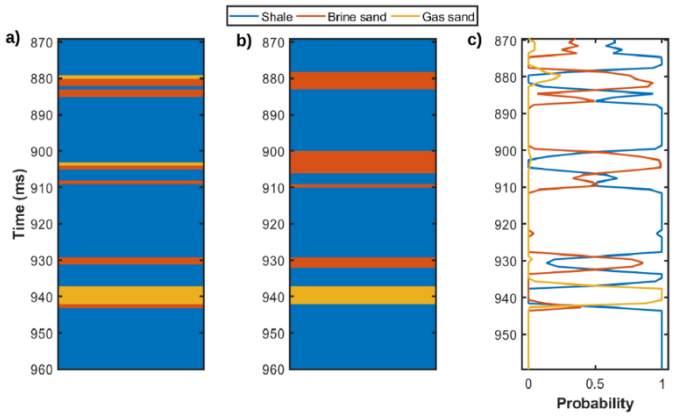

Figure 21: As in Figure 20 but for the MCMC algorithm.

Coverage probabilities (0.8)	*P-wave velocity*	*S-wave velocity*	*Density*
Analytical	0.916	0.912	0.926
MCMC	0.944	0.902	0.934

Table 1: Coverage probabilities derived from the posterior pdfs estimated by the analytical and MCMC algorithms for the first blind well.

The comparison of the MCMC and analytical approach is now extended to the second blind well where we impose a S/N ratio of 2. Again, the two approaches yield similar predicted elastic profiles (see Figures 22 and 23), although the MCMC inversion often provides slightly superior prediction intervals as demonstrated by the coverage probability values (see Table 2). In this case the differences between the outcomes of the two approaches can be clearly appreciated by comparing the facies classification results. Indeed, just a visual inspection of the estimated facies models and the associated posterior *pdfs* confirms that the MCMC method outperforms the analytical inversion as it estimates a MAP facies solution with a closer match with the actual facies profile especially below 935 ms, where the analytical approach erroneously interprets a finely layered shale-brine sand sequence as a gas saturated layer enclosed in a thick brine sand sequence. Finally, Figure 26 shows a comparison between the marginal distributions derived from the actual elastic properties of the second blind well and the marginal distributions derived from the ensemble of accepted MCMC models. This figure proves that the implemented MCMC algorithm successfully samples the actual multimodal distribution of the elastic parameters.

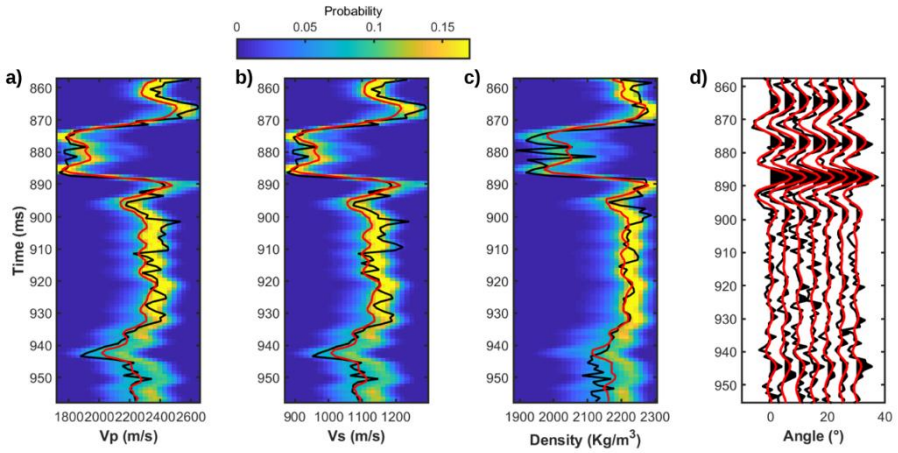

Figure 22: Results provided by the analytical inversion for the second blind well test. In a), b) and c) the black lines represent the true property values, the red line are the mean models, whereas the colormap codes the estimated posterior pdf. d) Comparison of observed (black) and predicted (red) seismic data computed on the a-posteriori mean model.

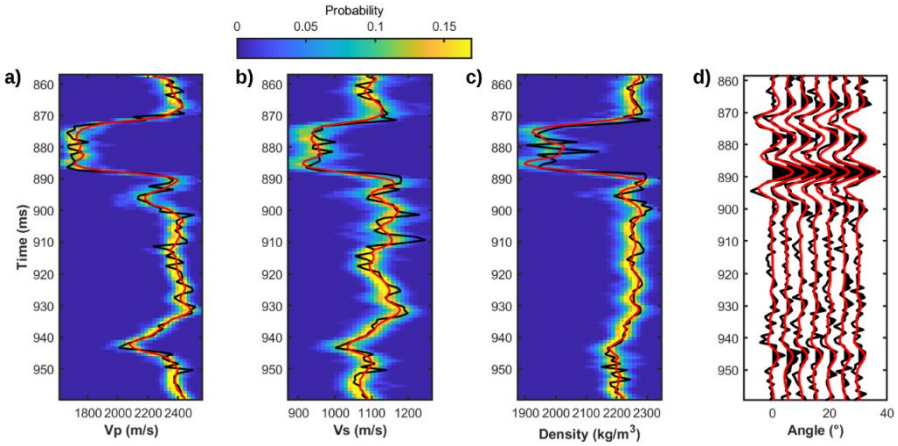

Figure 23: As in Figure 22 but for the MCMC inversion.

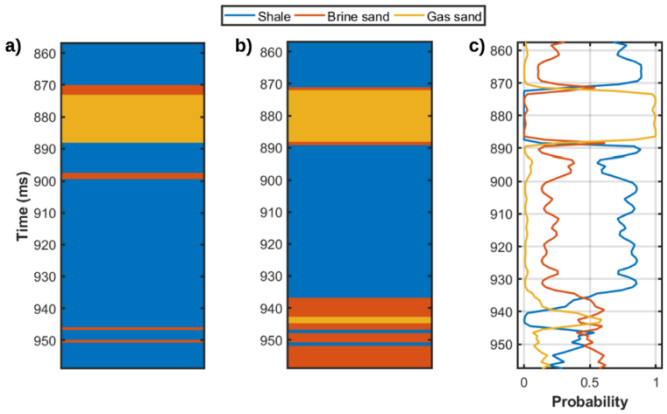

Figure 24: Facies classification results provided by the analytical inversion for the second blind test. a) True facies profile. b) MAP facies solution. c) Estimated posterior pdf of facies.

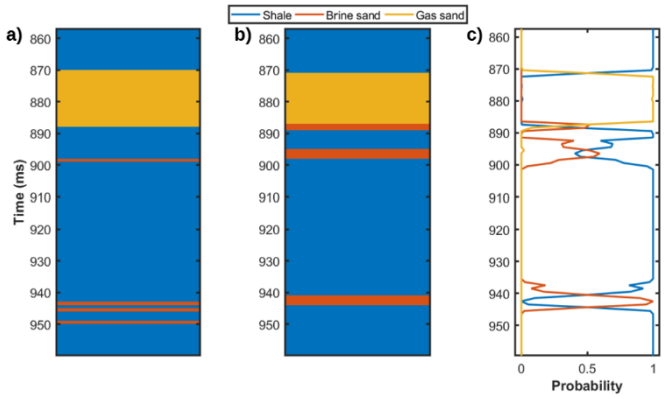

Figure 25: As Figure 20 but for the MCMC algorithm.

Coverage probabilities (0.8)	*P-wave velocity*	*S-wave velocity*	*Density*
Analytical	0.824	0.796	0.805
MCMC	0.924	0.865	0.906

Table 2: Coverage probabilities derived from the posterior pdfs provided by the analytical and MCMC algorithms for the second blind well.

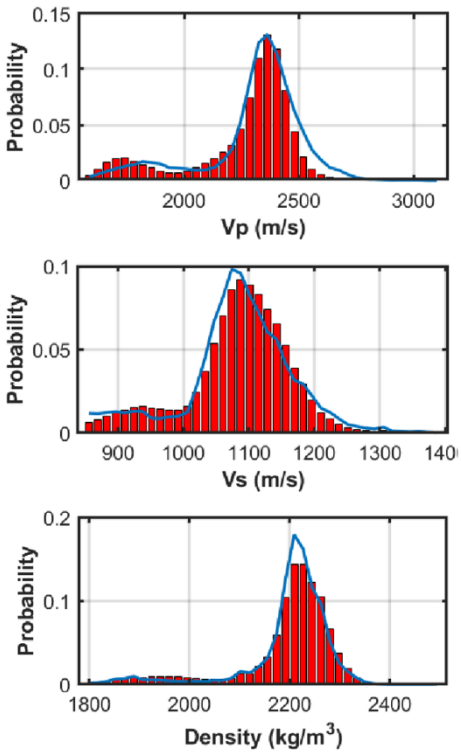

Figure 26: Comparison between the actual marginal distributions of elastic parameters along the second blind well (blue curves) and the marginal distributions derived from the accepted MCMC models (red bars).

This example proves that in the investigated area, the correct modelling of the facies dependency of the elastic properties is crucial to achieve accurate estimations and reliable prediction intervals. For what concern the computational costs, the analytical method runs in very few seconds if we consider a serial Matlab code, running on an Intel i7-7700HQ@2.8GHz with 16 Gb RAM. For the MCMC inversion we employ a parallel Matlab code that runs in approximately 12 minutes on two compute nodes equipped with two deca-core Intel E5-2630 @2.2 GHz (128 Gb RAM).

Convergence diagnostic for the interval-oriented algorithm

We now focus the attention on the second blind well example to analyse the convergence of the interval-oriented algorithm. The evolution of the log-likelihood values for 5 out of 40 chains (Figure 27) proves that the algorithm attains the stationary regime after 5000 iterations, approximately, after

which the likelihood values fluctuate around a stable value. Figure 28 shows examples of evolutions of the PSRF value for the *Vp*, *Vs*, and density values extracted from six different time positions. It emerges that 1000 iterations, approximately, are needed to achieve reliable uncertainty quantifications for the *Vp* parameter, whereas the convergence for the *Vs* and density is attained after 5000 iterations. These different convergence rates are again related to the different influence played by *Vp*, *Vs*, and density on the observed seismic amplitudes. The analysis of the normalized autocorrelation functions was used to set the lag-period value in the MCMC sampling (Figure 29). We observe that approximately 20 consecutive models are needed to obtain a correlation value below the selected threshold of 0.5-0.4.

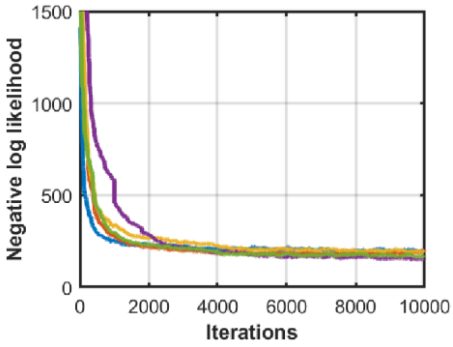

Figure 27: Example of evolutions of the negative log-likelihood for 5 out of 40 interactive chains.

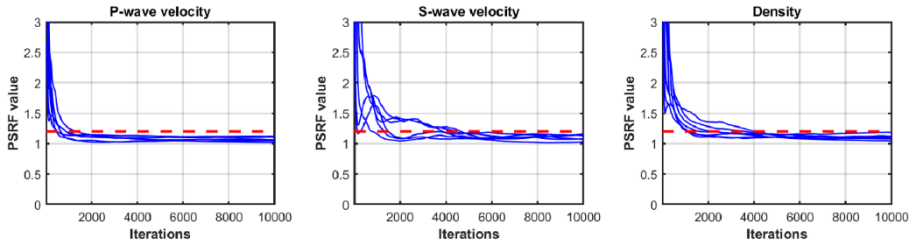

Figure 28: Examples of evolution of the PSRF value for the elastic properties estimated at 6 different time positions. In Each plot the dotted red lines show the desired PSRF value of 1.2, whereas the blue lines pertain to the elastic properties estimated at different time samples.

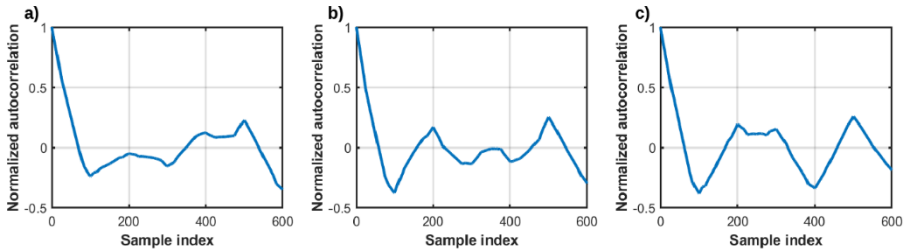

Figure 29: Close-up of the normalized autocorrelation functions for the Vp (a), Vs (b), and density (c) models sampled at a given time position.

DISCUSSION

The aim of the implemented Markov Chain Monte Carlo (MCMC) inversions is to accurately estimate the multi-modal posterior distribution of facies and elastic properties in a high-dimensional parameter space. In our inversion examples, we assume the noise to be uncorrelated and Gaussian-distributed. Possible improvements could be including the noise spatial correlation (Madsen et al. 2017) and/or the noise standard deviation as additional unknowns to be sampled during the inversion (Bodin et al. 2012; Sambridge 2014). In other words, this strategy let the algorithm infer the appropriate level of fitting and prevents to overfit or underfit the observed data. We are now working on these research topics.

From a mathematical point of view, the 1D interval approach can be easily extended to 2D or 3D cases by adding lateral constrains for the continuous and discrete model parameters. One of the challenges in field case applications, is the convergence of the algorithm, which can be very slow in high dimensional spaces due to the long time necessary to reach the stationary regime. In these cases, the posterior probability density function *(pdf)* tends to be highly localized within each model dimension and this results in a low acceptance ratio and a very slow convergence of the chain. In other terms, the chain tends to converge to local optima of the target *pdf* represented by a given facies configuration. To partially attenuate this situation, we suggest using high ξ value (around 0.8-0.9) at

the beginning of the inversion in order to promote jump between different facies configurations. Then, the ξ value should be decreased (i.e. to 0.2-0.3) during the sampling to get a robust estimate of the posterior *pdf* of the continuous property. In addition, the use of the a-priori as the proposal usually reduces the acceptance ratio of the chain and slows down the convergence. For this reason, the algorithm can be improved by sampling from an approximate posterior distribution rather than from the prior model. To this end many strategies could be adopted, for example implementing the adaptive Metropolis algorithm (Haario et al. 2001), or sampling from an approximated estimate of the posterior covariance matrix computed from a local approximation of the Jacobian matrix (see Dosso et al. 2014).

In a 2D interval-oriented inversion of field seismic data we performed (not shown here for confidentiality reasons), we found also particularly useful deriving the starting model for the MCMC sampling from the results of a previous Bayesian linear inversion. In this way, the burn-in period that is needed to attain the stationary regime is drastically reduced. However, if we consider the current implementation there is still room for a further reduction of the computational cost, for example by employing a more efficient and scalable parallel code. However, it is the author' opinion that all these improvements are not enough to make the algorithm applicable in 3D inversions. For this reason, we are now working on more advanced MCMC algorithms that incorporate the principles of Hamiltonian dynamics into the standard Metropolis-Hasting method (Betancourt, 2017). This approach exploits the derivative information of the objective function (the negative log-likelihood) to speed up the converge of the MCMC sampling toward the stationary regime. An additional, outstanding benefit of this method is that it produces largely independent models so that it allows for a substantial reduction of the lag parameter. This reflects in more accurate posterior estimations and in a significant reduction of the computational effort.

CONCLUSIONS

We presented two Markov Chain Monte Carlo (MCMC) inversion algorithms for target- and interval-oriented amplitude versus angle (AVA) inversion. The main advantage of the implemented MCMC recipe is that it is suitable for mixed discrete-continuous inverse problems, non-linear forward modellings and multimodal, non-parametric prior distributions. In other terms, our approach does not require any assumptions (i.e. Gaussianity) about the distribution of the continuous properties in a given facies. The first algorithm performs a 2D target-oriented inversion where only the AVA responses of the target reflection are considered. Differently, the second approach performs a 1D interval-oriented inversion in which a given time range around the target reflection is inverted. Both methods include geostatistical constraints for the elastic parameters, a 1D Markov prior models for the facies distribution, and use the exact non-linear Zoeppritz equations as the forward modelling. Our implemented MCMC recipe is especially aimed at decreasing the computational effort and it includes multiple chains, a parallel tempering strategy, a delayed rejection updating scheme and hybridize the standard Metropolis-Hastings algorithm with the Differential Evolution Markov Chain method. Our inversion results and the convergence analysis of the MCMC sampling demonstrated that two implemented algorithms efficiently sample from a multimodal non-parametric mixture distribution with a reasonable computational effort. The outcomes of the two implemented methods have been assessed by comparison with the results provided by analytical inversions that consider Gaussian-mixture distributed elastic properties.

Our examples showed that the proposed target-oriented MCMC approach yields final predictions very similar to the analytical inversion in case of optimal signal-to-noise ratio and when the actual model parameter distribution can be reasonably approximated to a Gaussian-mixture model. Differently, the lateral constraints included in the target-oriented MCMC method ensured more stable and more reliable predictions in cases of poor signal-to-noise ratio. The experiments on the 1D interval-oriented approach demonstrated the importance of correctly modelling the multimodal

behavior of the elastic properties to retrieve accurate predictions. Indeed, although the analytical inversion algorithm achieved satisfactory results, the non-parametric prior considered by the MCMC approaches guaranteed superior solutions and more accurate uncertainty quantifications.

Obviously, for the applications of the proposed methods all the requirements for AVA inversion must be met (controlled amplitude processing, good data quality). In addition, the local 1D assumption for the subsurface model must be valid. For this reason, in case of lateral velocity variations the seismic data input for the inversion should be accurately pre-stack time migrated.

Data Availability Statement

Data available on request due to privacy/ethical restrictions

References

Adriansyah, and McMechan, G. A. (2001). AVA analysis and interpretation of a carbonate reservoir: northwest Java basin, Indonesia. Geophysics, 66(3), 744-754.

Aleardi, M., Ciabarri, F., and Gukov, T. (2018a). A two-step inversion approach for seismic-reservoir characterization and a comparison with a single-loop Markov-chain Monte Carlo algorithm. Geophysics, 83(3), R227-R244.

Aleardi, M. (2018). Analysis of Different Statistical Models in Probabilistic Joint Estimation of Porosity and Litho-Fluid Facies from Acoustic Impedance Values. Geosciences, 8, 388. doi:10.3390/geosciences8110388

Aleardi, M., and Mazzotti, A. (2014). A feasibility study on the expected seismic AVA signatures of deep fractured geothermal reservoirs in an intrusive basement. Journal of Geophysics and Engineering, 11(6), 065008.

Aleardi, M., Ciabarri, F., and Mazzotti, A. (2017). Probabilistic estimation of reservoir properties by means of wide-angle AVA inversion and a petrophysical reformulation of the Zoeppritz equations. Journal of Applied Geophysics, 147, 28-41.

Aleardi, M., and Mazzotti, A. (2017). 1D elastic full-waveform inversion and uncertainty estimation by means of a hybrid genetic algorithm–Gibbs sampler approach. Geophysical Prospecting, 65(1), 64-85.

Aleardi, M., Ciabarri, F., and Gukov, T. (2018b). Reservoir Characterization Through Target-Oriented AVA-Petrophysical Inversions with Spatial Constraints. Pure and Applied Geophysics, 176 (2), 901–924.

Aki, K., and Richards, P. G. (1980). Quantative seismology: Theory and methods. New York, 801.

Aster, R. C., Borchers, B. and Thurber C. H. (2011). Parameter Estimation and Inverse Problem. Elsevier.

Betancourt, M. (2017). A conceptual introduction to Hamiltonian Monte Carlo. arXiv preprint arXiv:1701.02434.

Bodin, T., and Sambridge, M. (2009). Seismic tomography with the reversible jump algorithm. Geophysical Journal International, 178(3), 1411-1436.

Bodin, T., Sambridge, M., Tkalčić, H., Arroucau, P., Gallagher, K., and Rawlinson, N. (2012). Transdimensional inversion of receiver functions and surface wave dispersion. Journal of Geophysical Research: Solid Earth, 117, B02301.

Bosch, M., Cara, L., Rodrigues, J., Navarro, A., and Díaz, M. (2007). A Monte Carlo approach to the joint estimation of reservoir and elastic parameters from seismic amplitudes. Geophysics, 72(6), O29-O39.

Bosch, M., Mukerji, T., and Gonzalez, E.F. (2010). Seismic inversion for reservoir properties combining statistical rock physics and geostatistics: A review. Geophysics, 75(5), 165–176.

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

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

de Figueiredo, P.L., Grana, D., Luis Bordignon, F., Santos, M., Roisenberg, M., and Rodrigues, B. B. (2018). Joint Bayesian inversion based on rock-physics prior modeling for the estimation of spatially correlated reservoir properties. Geophysics, 83(5), 1-53.

Dosso, S. E., Holland, C. W., and Sambridge, M. (2012). Parallel tempering for strongly nonlinear geoacoustic inversion. The Journal of the Acoustical Society of America, 132(5), 3030-3040.

Dosso, S. E., Dettmer, J., Steininger, G., and Holland, C. W. (2014). Efficient trans-dimensional Bayesian inversion for geoacoustic profile estimation. Inverse Problems, 30(11), 114018.

Doyen, P. (2007). Seismic reservoir characterization: An earth modelling perspective. Houten: EAGE publications.

Falcioni, M., and Deem, M. W. (1999). A biased Monte Carlo scheme for zeolite structure solution. The Journal of chemical physics, 110(3), 1754-1766.

Gelman, A., and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. Statistical science, 7(4), 457-472.

Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. (1995). Bayesian Data Analysis, Chapman & Hall, London.

Gong, T., and McMechan, G. A. (2016). Target-oriented linear least squares and nonlinear, trust-region Newton inversions of plane waves using AVA and PVA data for elastic model parameters. Geophysics, 81(5), R325-R338.

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

Grana, D. (2018). Joint facies and reservoir properties inversion. Geophysics, *83*(3), M15-M24.

Grion, S., Mazzotti, A., and Spagnolini, U. (1998). Joint estimation of AVO and kinematic parameters. Geophysical Prospecting, 46(4), 405-422.

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

Haario, H., Saksman, E., and Tamminen, J. (2001). An adaptive Metropolis algorithm. Bernoulli, 7(2), 223-242.

Holmes, C., Krzysztof, L. and Pompe, E. (2017). Adaptive MCMC for multimodal distributions. Technical report.

Kennedy, W. J., and Gentle, J. E. (2018). Statistical computing. Routledge.

MacKay, D.J. (2003). Information Theory, Inference and Learning Algorithms, Cambridge University Press.

Madsen, R. B., Zunino, A., and Hansen, T. M. (2017). On inferring the noise in probabilistic seismic AVO inversion using hierarchical Bayes. In SEG Technical Program Expanded Abstracts 2017 (pp. 601-606). Society of Exploration Geophysicists.

Madsen, R. B., and Hansen, T. M. (2018). Estimation and accounting for the modeling error in probabilistic linearized amplitude variation with offset inversion. Geophysics, 83(2), N15-N30.

Malinverno, A. (2002). Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem. Geophysical Journal International, 151(3), 675-688.

Matheron, G., Beucher, H., De Fouquet, C., Galli, A., Guerillot, D., and Ravenne, C. (1987). Conditional simulation of the geometry of fluvio-deltaic reservoirs: SPE Annual Technical Conference and Exhibition, SPE, Extended Abstracts, SPE-16753–MS.

Mazzotti, A., and Zamboni, E. (2003). Petrophysical inversion of AVA data. Geophysical Prospecting, 51(6), 517-530.

Mazzotti, A., and Ravagnan, G. (1995). Impact of processing on the amplitude versus offset response of a marine seismic data set. Geophysical prospecting, 43(3), 263-281.

Le Ravalec, M., Noetinger, B., and Hu, L. Y. (2000). The FFT moving average (FFT-MA) generator: An efficient numerical method for generating and conditioning Gaussian simulations. Mathematical Geology, 32(6), 701-723.

Sambridge, M. (2013). A parallel tempering algorithm for probabilistic sampling and multimodal optimization. Geophysical Journal International, 196(1), 357-374.

Sabeti, H., Moradzadeh, A., Ardejani, F. D., Azevedo, L., Soares, A., Pereira, P., and Nunes, R. (2017). Geostatistical seismic inversion for non-stationary patterns using direct sequential simulation and co-simulation. Geophysical Prospecting, 65, 25-48.

Sambridge, M., and Mosegaard, K. (2002). Monte Carlo methods in geophysical inverse problems. Reviews of Geophysics, 40(3), 3-1.

Sambridge, M. (2014). A parallel tempering algorithm for probabilistic sampling and multimodal optimization. Geophysical Journal International, 196(1), 357-374.

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

Stolt, R. H., and Weglein, A. B. (1985). Migration and inversion of seismic data. Geophysics, 50(12), 2458-2472.

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

ter Braak, C. J., and Vrugt, J. A. (2008). Differential evolution Markov chain with snooker updater and fewer chains. Statistics and Computing, 18(4), 435-446.

Tierney, L. and Mira, A., (1999). Some adaptive Monte Carlo methods for Bayesian inference, Stat. Med., 18(1718), 2507–2515.

Zhu, X., and McMechan, G. A. (2012). Elastic inversion of near-and postcritical reflections using phase variation with angle. Geophysics, 77(4), R149-R159.