Bollettino di Geofisica Teorica ed Applicata Vol. 62, n. 1, pp. 73-88; March 2021
DOI 10.4430/bgta0328

A data-driven transdimensional approach to include lateral constraints on 2D target-oriented AVA inversion

A. Salusti[1,2] and M. Aleardi[1]

[1] *Earth Sciences Dept., University of Firenze, Italy*
[2] *Earth Sciences Dept., University of Pisa, Italy*

(Received: 7 November 2019; accepted: 27 April 2020)

ABSTRACT Seismic inversion aims to infer subsurface properties from processed seismic data; since these are often ill-conditioned procedures, numerous strategies can be investigated. To date, currently adopted procedures assume an *a priori* structural knowledge of the investigated area and impose such constraints on the recovered solution. To overcome this shortcoming, we apply a transdimensional reversible jump-Markov chain Monte Carlo (Rj-McMC) algorithm to solve the interval-oriented amplitude versus angle (AVA) inversion on 2D synthetic seismic data. This approach considers the model parameterisation as an unknown, together with the elastic properties of the investigated area. The algorithm samples models discretised in Voronoi cells characterised by similar AVA responses. The elastic values associated with each Voronoi cell are obtained taking the average of the elastic properties of the Common Dip Points (CDPs) falling within it. This data-driven approach, therefore, needs no external assumption over the investigated area and ensures an automatically inferred strategy to include lateral variability of data inside the inversion kernel. We compare results obtained with a standard Bayesian approach for different signal-to-noise ratios (SNR), showing how the increase of random noise contaminating the data strongly affects the linear approach, while the Rj-McMC generates model predictions in accordance with the true model, producing more reliable results.

Key words: AVA inversion, lateral constraint, transdimensional Markov chain, Monte Carlo method.

1. Introduction

The inclusion of lateral constraints in the inversion framework is the most popular strategy devoted to attenuating the ill-conditioning of the seismic inversion. Several advanced regularisation strategies exist, such as the inclusion of geostatistical constraints in the form of isotropic model correlation functions (Buland and Omre, 2003), multipoint statistics (Cordua *et al*., 2012; Hansen *et al*., 2012), stratigraphic constraints (Tetyukhina, 2011), or constraints derived by the seismic bandwidth (Haas and Dubrule, 1994). The main limitation of all these approaches is that they rely on an *a priori* structural knowledge of the investigated area and force the recovered model to obey such *a priori* constraints. These are essentially model-driven regularisation strategies that could provide biased model parameter estimations in case of erroneous *a priori* assumptions. More advanced approaches train sparse dictionary learning algorithms to impose sparse representations

of subsurface model parameters (She *et al*., 2018) or make use of anisotropic Markov random fields (Aleardi *et al*., 2018; Guo *et al*., 2018). The goal of these approaches is to locally adapt the structural constraint to the local structural characteristics of the subsurface model that can be iteratively inferred from the local characteristics (i.e. variability) of the observed data.

In line with these data-driven approaches, we present a numerical transdimensional reversible jump-Markov chain Monte Carlo (Rj-McMC) algorithm for target-oriented amplitude versus angle (AVA) inversion. Numerical reservoir models play a key role in the exploration industry since they describe the spatial variability of reservoir elastic properties. To this end, interesting works have focused on establishing inversion procedures for complex-media reservoir characterisation that reliably estimate rock properties from seismic data (Ba *et al*., 2017; Chen and Zhang, 2017).

In our case target-oriented means that only the AVA responses of the target layer are inverted and these AVA responses are extracted for each considered Common Dip Points (CDP) position. The results are 2D maps representing the lateral variability of the elastic contrasts along the considered reflecting interface. The AVA inversion is a severely ill-conditioned problem, highly affected by noise contamination, in which it is crucial to adopt a reliable regularisation strategy to retrieve reliable and stable results. Furthermore, the extraction of reflection coefficients from seismic data is an extremely delicate process that needs a robust procedure (Grion *et al*., 2002). In a transdimensional inversion, the number of model parameters (that codes the optimal subsurface model parameterisation) is considered unknown and is estimated using a probabilistic sampling. In our case, the inverted 2D horizon is divided into Voronoi cells, whose number and shape are automatically determined by the Rj-McMC sampling. The algorithm autonomously partitions the considered 2D horizon on the basis of the spatial variability of data, producing subsurface 2D models discretised in Voronoi polygons, each one enclosing CDP positions with similar AVA responses. This also means that the CDPs falling within the same cell also share similar elastic properties and for this reason the same elastic property values are assigned to these CDPs. These values are computed by averaging the model properties pertaining to the CDPs falling within each cell. Similarly, the observed data for each polygon is computed by averaging the AVA responses of the CDPs falling within each Voronoi cell. On one hand, this strategy constitutes a data-driven approach to include lateral constraints into the AVA inversion because these constraints are automatically inferred from the lateral variability of the data and not arbitrarily infused into the inversion framework. On the other hand, the averaging of the AVA responses pertaining to CDPs falling within the same cell inherently increases the signal to noise ratio (SNR) of the observed data. These two aspects proved of crucial importance for stabilising the inversion, even in the case of severely noise-contaminated data.

For the lack of field data, we test the implemented Rj-McMC algorithm by performing synthetic inversions with different SNR. The proposed method is benchmarked against a more standard Bayesian AVA inversion without lateral constraints.

2. Bayesian formulation

The Bayesian approach originates from the probability theory of the English mathematician Thomas Bayes (1702-1761). There are two main differences between the classical deterministic and Bayesian approach: the solution, which is described by a probability distribution, and the

inclusion of *a priori* information inside the inversion kernel. This *a priori* information is derived from seismic independent data (as well as log information, geological information, and so on) and is always expressed by a probability distribution, which respects the model parameters.

Seismic data are, thus, combined with *a priori* information through Bayes' theorem, to obtain the posterior distribution for model parameters. The theorem can be formulated as:

$$p(\boldsymbol{m}|\boldsymbol{d}) = \frac{p(\boldsymbol{d}|\boldsymbol{m})p(\boldsymbol{m})}{p(\boldsymbol{d})} \quad (1)$$

where $p(\boldsymbol{m})$ denotes the *a priori* probability: the probability density function (PDF) over $\boldsymbol{m}$. The $p(\boldsymbol{d}|\boldsymbol{m})$ is the likelihood function and expresses a conditional probability and quantifies the fit among observed and predicted data from model $\boldsymbol{m}$. The denominator of Eq. 1 is considered a constant scale factor, that normalises the conditional distribution $p(\boldsymbol{m}|\boldsymbol{d})$ (Aster *et al*., 2005) and that can be ignored when only the shape of the posterior PDF is of interest. The $p(\boldsymbol{m}|\boldsymbol{d})$ represents the *a posteriori* PDF describing the PDF of $\boldsymbol{m}$ once the observations $\boldsymbol{d}$ are obtained.

3. Rj-McMC algorithm

Among the numerical approaches for non-linear inverse problems, Markov chains Monte Carlo (McMC) methods are the most commonly used, gaining growing attention over the last decades. Their genesis dates back to the mid-1940s when the method was formalised by Nicholas Constantine Metropolis and Stanislaw Marcin Ulan (Metropolis and Ulan, 1949); nonetheless, the use of random numbers to solve problems of probabilistic nature is quoted in a paper by Lord Kelvin back in 1901. Actually, the development of McMC techniques and the statistical theory at its base constitute a large and active research field (Flournoy and Tsutakawa, 1991) and in recent years the different typologies of problems where it has been applied have grown exponentially (Moosegard and Tarantola, 1995; Sambridge and Moosegard, 2002).

McMC algorithms are iterative procedures where a model is generated through a random walk on the model space. This random path satisfies the Markov property of "absence of memory": each generated model depends only on the previous model in the chain. The new model is, then, accepted or rejected based on some acceptance rules. Iteration by iteration the chain collects models spread all over the model space and reaches the stationary regime. The ensemble of all collected models during the random walk defines the posterior PDF. These methods, therefore, convert the inverse problem into a sampling problem in which the sampling density is proportional to the posterior, so that the sampled models can be used to approximate the statistical properties of the posterior PDF.

McMC algorithms require considerable computational effort with respect to analytical inversion approaches and for this reason, a specific McMC recipe, tailored to the problem in hand, is usually needed to efficiently sample the PDF. Like all the global optimisation methods, the choice of the starting model is not important, since we consider a user-defined number of iterations, called burn-in, to let the algorithm move toward the stationary regime; models collected before burn-in will be discarded from the computing of the posterior PDFs as they are not representative of the true posterior distribution. An accurate choice for the burn-in period is, therefore, essential to prevent biased PDF.

Another very important issue related to all McMC methods is their inability to sample multi-modal complex distributions. In a Monte Carlo sampling, the algorithm typically moves towards regions of high probability and, once reached, it will likely exploit it, at the expense of a strongly decreased model space exploration. Several methods have been deployed and applied to overcome this problem: Sen and Biswas (2017) hybridised the Rj-McMC with a faster Hamiltonian Monte Carlo algorithm, Lan proposed modifying the Riemannian geometric properties of the target distribution to create wormholes connecting high probability density regions (Lan *et al*., 2014). For this work, a parallel tempering approach has been adopted: several multiple chains are run at different temperatures to speed up the search on parameter space. This method has been widely used in computational statistics (Geyer, 1991). In exploration geophysics, it has been gaining attention (Dosso *et al*., 2012; Sambridge, 2013), since the results have demonstrated that the improved parallel-tempered McMC methods effectively accelerate the algorithm convergence speed.

4. Method

4.1. Forward model

Zoeppritz equations describe the partitioning of wave seismic energy at an interface in terms of P-wave velocity, S-wave velocity, and density of the upper and lower layers. In order to reduce the ambiguity within the AVA method (Drufuca and Mazzotti, 1995) and further constrain the inversion, the number of unknowns can be reduced. In this work, we have adopted the Ursenbach and Stewart (2008) equations as the forward model to describe the system in terms of relative contrasts in P-impedance and S-impedance (R_I and R_J):

$$Rpp(\Theta) = \left(1 + \sin^2\Theta \frac{4\gamma^2 + cos^2\Theta - 1}{5}\right)\frac{R_I}{cos^2\Theta} - 8\gamma \sin^2\Theta R_J \tag{2}$$

where Rpp is the P-wave reflectivity, Θ is the angle of incidence across the interface, $\gamma = \frac{V_{s_1} + V_{s_2}}{V_{p_1} + V_{p_2}}$ with the subscripts 1 and 2, respectively, referring to the upper and lower layer,

$$R_I = \frac{\Delta\, Ip}{2\,\tilde{Ip}};\;\; R_J = \frac{\Delta\, Is}{2\,\tilde{Is}} \tag{3}$$

where Ip and Is are respectively the P-wave and S-wave impedance; the operator Δ indicates the contrast and the symbol ˜ indicates the mean across the interface.

4.2. Space discretisation

In this work, the model space is discretised using the Voronoi polygons over equally spaced grid points representing the CDP positions. A number of discrete set of points, the centre of Voronoi cells, partitions the plane and the cells are composed by those points whose distance from the centre is smaller. By doing so polygons do not overlap each other (all points of the grid fall in just one polygon) and the parameterisation is totally defined by the position of the nuclei. Our models are, therefore, represented by all the CDPs over the grid, to each of which we associate

a P-impedance and S-impedance value. Fig. 1 shows an example of the Voronoi diagram, which forms an irregular set of cells that partitions the plane.

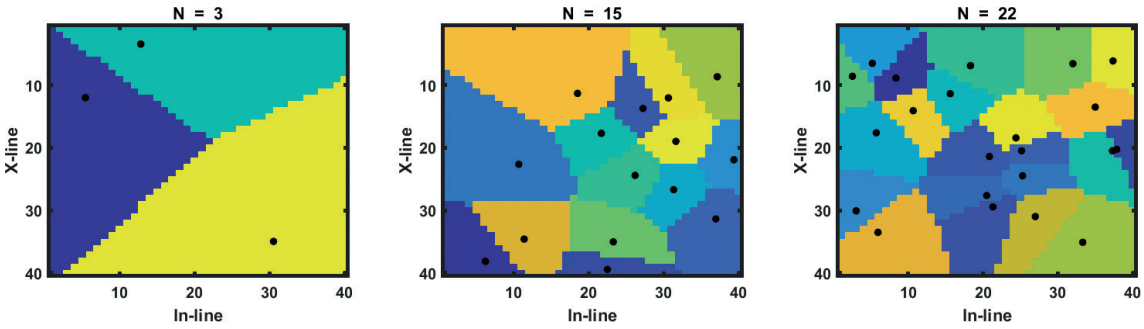

Fig. 1 - Example of model space partition in 3, 15 and 22 randomly distributed polygons; black dots indicate the nuclei centres.

4.3. Metropolis-Hastings

The Metropolis-Hastings rule (Metropolis *et al.*, 1953; Hastings, 1970) determines the probability to accept or reject a new model $\boldsymbol{m'}$ from previous model $\boldsymbol{m}$. For a given model space parametrisation with different number of unknowns γ, the probability can be expressed as:

$$\alpha = p\left(\boldsymbol{m}',\gamma' \middle| \boldsymbol{m},\gamma\right) = \min\left[1, |\mathbf{J}| \times prior\ ratio \times likelihood\ ratio \times proposal\ ratio\right] = \min\left[1, |\mathbf{J}| \times \frac{p\left(\boldsymbol{m}' \middle| \gamma'\right)}{p\left(\boldsymbol{m} \middle| \gamma\right)} \times \frac{p\left(\boldsymbol{d} \middle| \boldsymbol{m}',\gamma'\right)}{p\left(\boldsymbol{d} \middle| \boldsymbol{m},\gamma\right)} \times \frac{p\left(\boldsymbol{m},\gamma \middle| \boldsymbol{m}',\gamma'\right)}{p\left(\boldsymbol{m}',\gamma' \middle| \boldsymbol{m},\gamma\right)}\right] \quad (4)$$

where $\mathbf{J}$ represents the Jacobian matrix of the model transformation, here needed to take into account the possible dimensional scale change from $\boldsymbol{m}$ to $\boldsymbol{m'}$. It can be shown that for the present formulation the determinant of the Jacobian matrix is equal to 1 and can, therefore, be safely ignored (Bodin and Sambridge, 2009). The $p(\boldsymbol{d}|\boldsymbol{m'},\gamma')$ term describes the likelihood function, $p(\boldsymbol{m'}|\gamma')$ represents the *a priori* distribution and $p(\boldsymbol{m},\gamma|\boldsymbol{m'},\gamma')$ is the proposal distribution that defines the probability of generating a new model as a random perturbation of the previous model.

4.3.1. Likelihood function

The likelihood describes how well a given model $\boldsymbol{m}$ reproduces the observed data. In this work, the likelihood function is quantified by a least squares misfit function:

$$\Phi\left(\boldsymbol{m},\gamma\right) = \left\| \mathbf{C}_\mathbf{d}^{-1} \left(\mathbf{G}(\boldsymbol{m}) - \boldsymbol{d}_{obs} \right) \right\|_2^2 \quad (5)$$

where $\mathbf{G}(\boldsymbol{m})$ is the forward modelling operator (Eq. 2), $\boldsymbol{d}_{obs}$ represents the observed AVA and $\mathbf{C}_\mathbf{d}$ is the covariance matrix of data noise. Since seismic data are usually pre-processed with several correction methods to remove coherent noise, we can assume that noise left inside seismic data is white: Gaussian distributed and uncorrelated along the incidence angles. As known, minimising the least squares misfit function is equivalent to maximising the probability for a Gaussian-shaped likelihood function:

$$p(\boldsymbol{d}\,|\,\boldsymbol{m},\gamma) = \frac{1}{\sqrt{(2\pi)^L\,|\,\mathbf{C_d}\,|}}\exp\left\{-\frac{\Phi(\boldsymbol{m},\gamma)}{2}\right\} \tag{6}$$

with L given by the product of incidence angles with the number of CDPs inside each cell.

4.3.2. The prior

Every information that can be expressed as a probability distribution can be used as prior knowledge on the model (Gouveia and Scales, 1998). Here, we have used bounded and independent uniform distributions between fixed ranges. Following Bodin and Sambridge (2009), the prior distribution is therefore expressed as:

$$p(\boldsymbol{m},\gamma) = p(\mathbf{e}\,|\,\gamma)p(\mathbf{z}\,|\,\gamma)p(\gamma) \tag{7}$$

where $p(\boldsymbol{e}|\gamma)$ represents the prior on the elastic properties, $p(\boldsymbol{z}|\gamma)$ is the prior on the layer position and $p(\gamma)$ refers to the number of layers:

$$p(\gamma) = \begin{cases} \frac{1}{\Delta\gamma}, & \text{if } \gamma_{min} \le \gamma \le \gamma_{max} \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

$$p(\mathbf{z}\,|\,\gamma) = \frac{\gamma!(N-\gamma)!}{N!} \tag{9}$$

$$p(e\,|\,\gamma) = \begin{cases} \prod_{i=1}^{\gamma} p(e_i\,|\,\gamma) = \prod_{i=1}^{\gamma} \frac{1}{\Delta e} = \frac{1}{(\Delta e)^{\gamma}}, & \text{if } \quad e_{min} \le e \le e_{max} \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

where γ_{min} and γ_{max} are the minimum and maximum number of nuclei that the algorithm is allowed to generate, N are all the possible locations of the nucleus centres, e_{min} and e_{max} define the range of elastic values the algorithm associates to each Voronoi nucleus.

For a rigorous mathematical derivation of the *a priori*, likelihood and proposal formulation used in this work, the reader may refer to Bodin and Sambridge (2009).

5. Proposed inversion procedure

For each chain the inversion starts from a single Voronoi cell representing a laterally invariant R_I and R_J value drawn from the prior distribution. The values of γ_{min} and γ_{max} defining the *a priori* range of the Voronoi cells number are, respectively, equal to 1 and three-quarters of maximum number of CDPs within the inverted area (chosen for computational reason). The model parameters include the number of Voronoi cells, the position of their nuclei, and the elastic properties associated with each polygon. To all the N CDPs positions falling within the same cell are assigned the same elastic property values that are computed as the average of their R_I and R_J values. Similarly, the average AVA response of the N CDPs constitutes the observed data for the

considered cell. Then, the algorithm evolves by sampling the model space that is by sampling the R_I and R_J values, the number of cells, and the positions of their nuclei. The sampling is driven by the acceptance probability given in Eq. 4, in which models with better fit with the observed data and with a parsimonious parameterisation (lower number of cells) are more likely to be accepted. During the iterations the algorithm tends to gather within the same Voronoi cell adjacent CDPs with similar AVA responses to which are assigned the elastic properties that produce a good fit with the observed data. On each iteration, the algorithm applies a perturbation to the current model chosen with equal probability from the following list:

1 - death move*:* delete one polygon from the Voronoi tessellation and rearrange the shape of the remaining polygons based on the positions of their nuclei (Fig. 2b);
2 - cell move: randomly choose one Voronoi cell and change the position of the corresponding centre without modifying the associated R_I and R_J values. This perturbation will produce a slight rearrangement of the Voronoi tessellation over the considered 2D horizon (Fig. 2c);
3 - elastic move: randomly choose one Voronoi cell and perturb the R_I and R_J values for all the CDP positions enclosed in the selected cell. This perturbation follows a Gaussian proposal centred on the current R_I and R_J values (Fig. 2d);
4 - birth move: create a new polygon within the Voronoi tessellation and assign the elastic properties to the newly created polygon by drawing a random realisation from the prior R_I and R_J distributions. Note that only the neighbouring cells of the new-born cell have their geometry changed during this step (Fig. 2e).

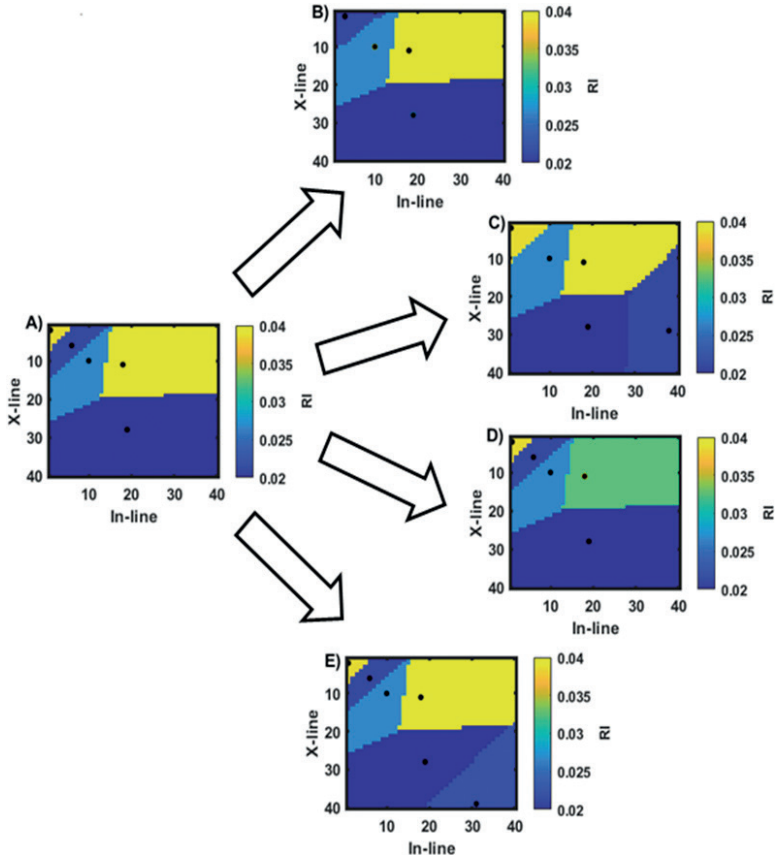

Fig. 2 - Proposed perturbation strategies adopted: a) starting model with 5 nuclei, b) perturbed model after death step, c) perturbed model after cell move, d) perturbed model after elastic property change, e) perturbed model after birth step.

Fig. 3 illustrates an example of the evolution of the R_I models sampled by the algorithm. These models represent the R_I values estimated over the 2D stratigraphic horizon. We note that the algorithm starts from very different R_I values from the reference model and from a Voronoi tessellation with a number of cells, which are not enough to successfully predict the observed data. As the iterations proceed, the sampled R_I values get closer to the reference model, whereas the Voronoi tessellation successfully partitions the considered 2D horizon by gathering within the same polygon CDPs with similar AVA response. In other words, the algorithm successfully recognizes adjacent CDP positions with similar AVA responses and similar relative contrasts in the P- and S-impedance values.

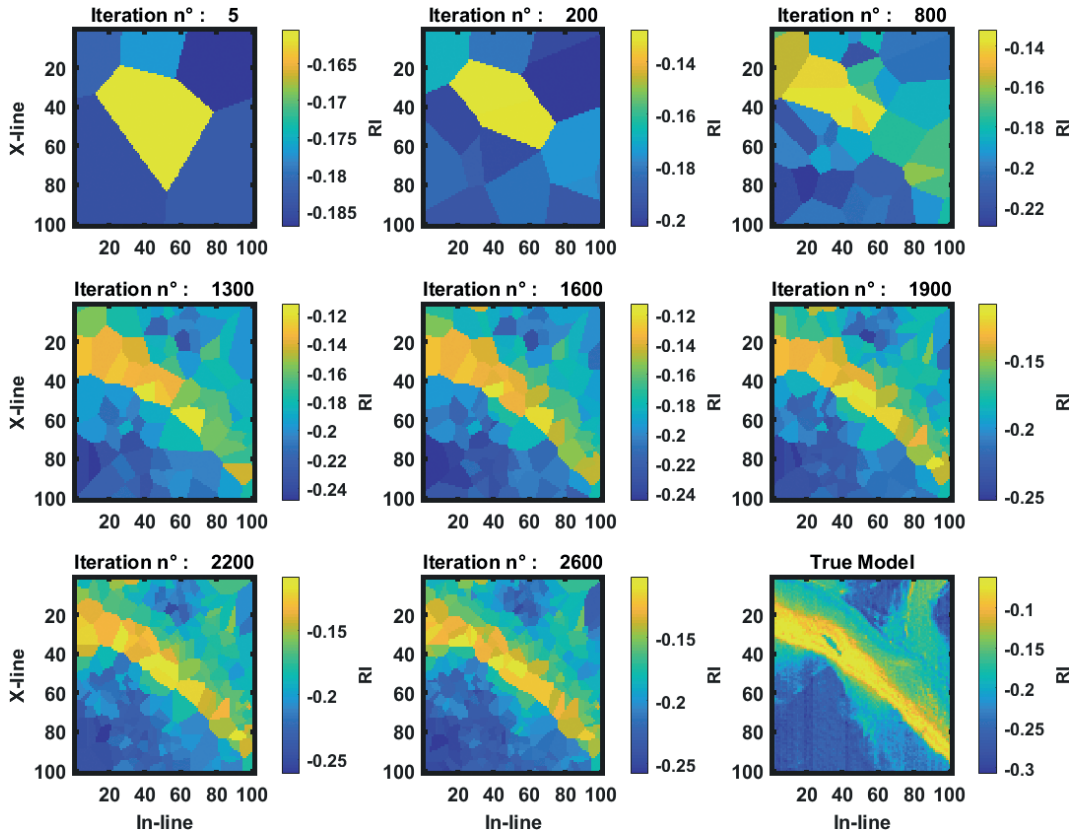

Fig. 3 - Example of evolution of the sampled R_I models from iteration 5 (top left) to iteration 2600 (bottom centre) over a regular spaced grid of 100×100 CDPs. The reference model is shown in the bottom right corner.

Defining a reliable method to assess the convergence of a transdimensional McMC inversion is still a matter of debate (Brooks *et al*., 2003). Indeed, in transdimensional optimisations the standard measures of convergence such as the potential scale reduction factor (Gelman *et al*., 2013) cannot be applied. In this work, the convergence of the algorithm has been addressed by comparing the PDFs estimated by each chain in the first and second halves of the sampled models after the burn-in. If these PDFs are in good agreement no further sampling is needed.

6. Parallel tempering

In order to increase the model space exploration and to ensure good properties mixing between the chains, we applied a parallel tempering simulation method. Each chain is initialised with a temperature and, after a fixed number of iterations we let the temperature swap between 2 chains, chosen with equal probability among all running chains. The likelihood for a given chain is:

$$p(\boldsymbol{d} \mid \boldsymbol{m}, \gamma) \propto \exp\left\{ -\frac{\Phi(\boldsymbol{m}, \gamma)}{2\,T} \right\} \tag{11}$$

according to Sambridge (2013), the probability p to swap two running chains temperatures, respectively identified with models $\mathbf{m}_1$ and $\mathbf{m}_2$ and temperatures T_1 and T_2, is equal to:

$$p = \min\left[1, \frac{\exp\left(-\frac{\Phi(\mathbf{m}_2)}{T_1} - \frac{\Phi(\mathbf{m}_1)}{T_2} \right)}{\exp\left(-\frac{\Phi(\mathbf{m}_1)}{T_1} - \frac{\Phi(\mathbf{m}_2)}{T_2} \right)} \right] \tag{12}$$

the temperature acts as a normalising factor, modifying the shape of the likelihood function (Fig. 4). Higher temperatures will produce a flatter shape, enabling the chain to freely explore the model space; lower temperatures will force the chains to exploit their current model regions. We proceed with collecting models from chains at temperature = 1, since all other models are representative of biased posterior PDFs (Dosso *et al.*, 2014).

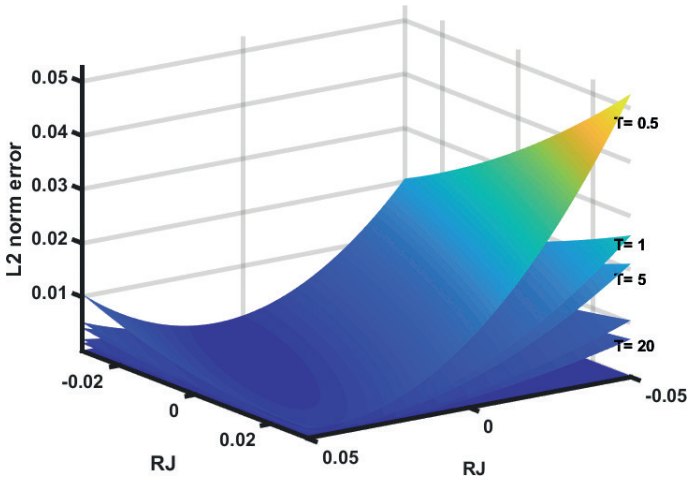

Fig. 4 - Effect of temperature values on the L2 norm error between true and observed data.

Note that too high temperatures will make the likelihood a non-informative parameter (i.e. at $T = \infty$, the likelihood will tend to 0), doing so the inversion will just follow the *a priori* information (Eq. 1). The temperature limits are set in order to ensure an acceptance ratio around 0.2-0.4, which, according to Sambridge and Moseggard (2002), is the optimal rate for McMC methods.

The entire process described above is summarised with a workflow reported in Fig. 5.

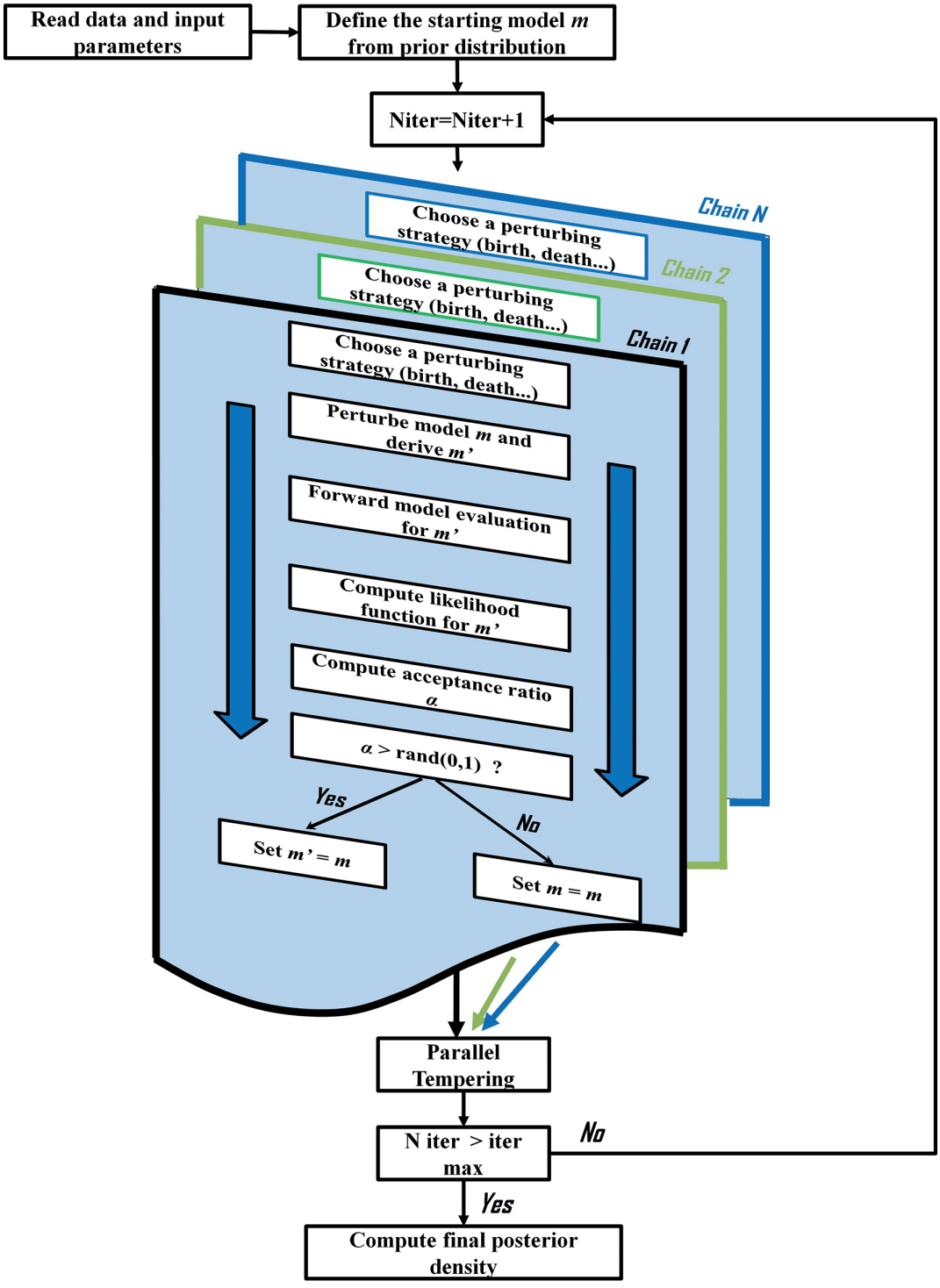

Fig. 5 - Flowchart of the implemented Rj-McMC inversion.

7. Synthetic inversion tests and results

The previously described reversible jump McMC is compared with a classical Bayesian linear approach. This is a standard, fast marching method with good reliability where the model $\boldsymbol{m}$ is obtained as:

$$\boldsymbol{m} = \left(\mathbf{C_d}^{-1}\left(\mathbf{G}(\boldsymbol{m})^{'}\,\mathbf{G}(\boldsymbol{m})\right) + \mathbf{C_m}^{-1}\right)^{-1} \quad \boldsymbol{G}'\left(\boldsymbol{d}_{obs} - \left(\mathbf{G}(\boldsymbol{m}_{prior})\right)\right)\mathbf{C_d}^{-1} \tag{13}$$

with $\mathbf{C_d}$ covariance matrix of data noise, $\mathbf{C_m}$ model covariance matrix, $\mathbf{G}$ represents the forward model operator (Eq. 2), $\boldsymbol{d}_{obs}$ is the observed seismic AVA and $\boldsymbol{m}_{prior}$ indicates the *a priori* model.

The synthetic observed AVA data are computed from a reference elastic model constituted by a regularly spaced grid of 320 by 480 CDPs (leftmost part of Figs. 7, 8, and 9). This reference model represents a stratigraphic section of a deltaic depositional area characterised by twisted river channel systems where yellow regions are representative of predominantly sandy channels, nested in blue-coloured background shale portions with significant lateral contrasts in the R_I and R_J values.

The Ursenbach and Stewart (2008) equation is applied to the reference model to compute the AVA response for each CDP gather position within an angle range between 0 and 45 degrees and a 40-Hz Ricker source wavelet with a sampling interval of 0.002 s. We considered three scenarios with different SNR and with Gaussian-distributed noise affecting the data: in the first, second, and last case the noise standard deviation is respectively equal to 0.001, 0.050, and 0.070. The SNR is defined as the difference between the power in decibels of a signal and the power in decibels of the background noise:

$$NR_{dB} = \; P_{signal,\,dB} - P_{noise,\,dB} \;\; . \tag{14}$$

In each scenario we used 20 chains, each one running for 10^5 iterations. Fig. 6 indicates that after the first 3000 iterations, approximately, the McMC algorithm reaches a stationary regime after which the algorithm begins to sample the posterior distribution. The V_P/V_S ratio oscillates from 1.9 to 2.2.

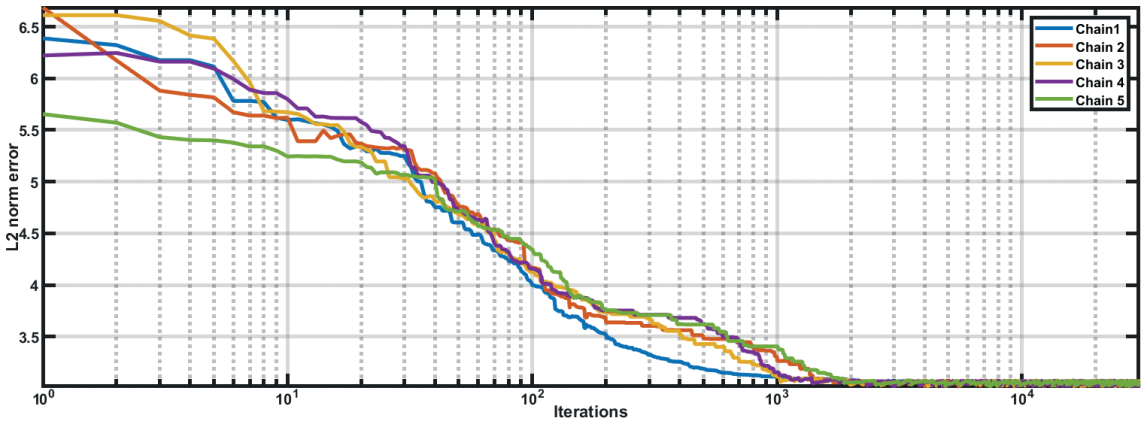

Fig. 6 - Example of evolution of the L2 norm misfit for 5 chains out of 20.

In order to significantly reduce the computational cost of the inversion procedure on the simulated 2D model for the Rj-McMC, we split the inversion area in smaller areas of 40×40 CDPs. The standard Bayesian has instead been applied separately on each CDP.

Fig. 7 shows that for a SNR of 70 dB both approaches used provide final map solutions in good agreement with the reference model and where the formation boundaries can be mapped with high accuracy. Differently, as the noise increases, (Figs. 8 and 9) the standard Bayesian algorithm without lateral constraints provides an estimated model characterised by significant scattering that is produced by the noise propagation from the data to the model space. On the contrary, the implemented Rj-McMC algorithm efficiently attenuates the ill-conditioning of the inversion procedure. On one hand, the averaging of the AVA response within the same Voronoi cell significantly increases the SNR of the observed data. On the other, the averaging of the elastic properties estimated for the CDPs position falling within the same cell, inherently introduces lateral constraints into the inversion framework, which can particularly be highlighted where the lateral contrasts are stronger (as in *xline* 150-250, *inline* 150-300 or *xline* 50-150, *inline* 350-450 of Figs. 7, 8, and 9). Both these characteristics of the Rj-McMC algorithm ensure a more stable inversion procedure and more reliable results.

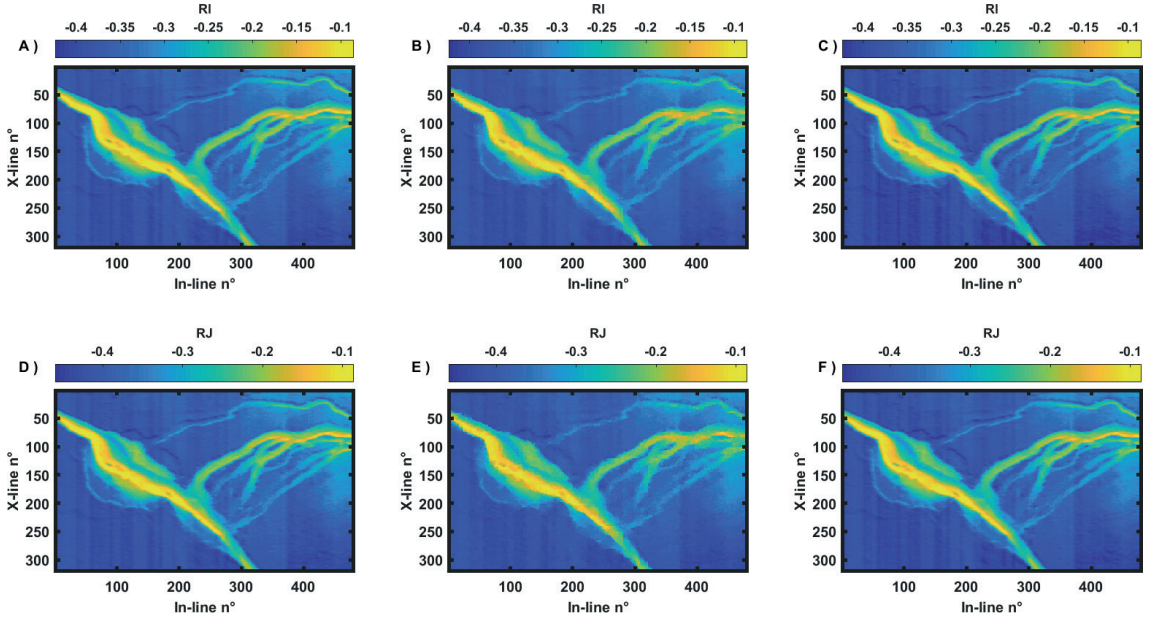

Fig. 7 - Comparison between the reference model (A and D), the solution provided by the Rj-McMC (B and E) and by the standard Bayesian approach (C and F) for noise standard deviation is equal to 0.001. R_I and R_J represent the relative contrasts in the *Ip* and *Is* values at the reflecting interface, respectively.

In order to give a quantitative indicator to compare the results obtained, we calculated the correlation coefficient (*RR*) between the synthetic data and the Rj-McMC and the standard Bayesian approach, as reported in Table 1.

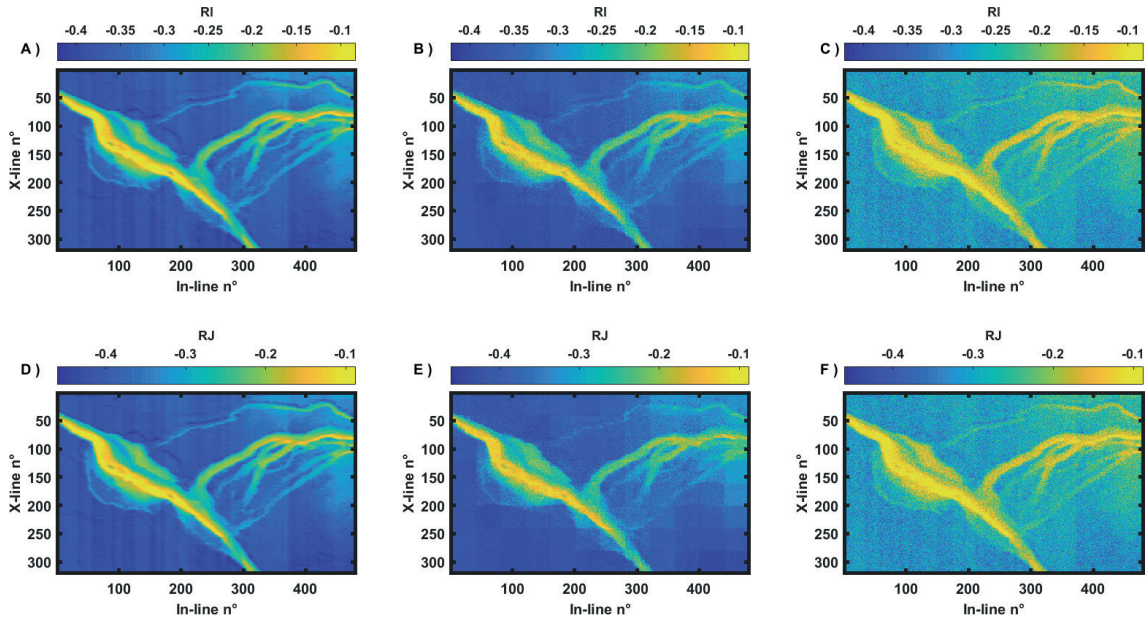

Fig. 8 - Comparison between the reference model (A and D), the solution provided by the Rj-McMC (B and E) and by the standard Bayesian approach (C and F) for noise standard deviation is equal to 0.050. R_I and R_J represent the relative contrasts in the *Ip* and *Is* values at the reflecting interface, respectively.

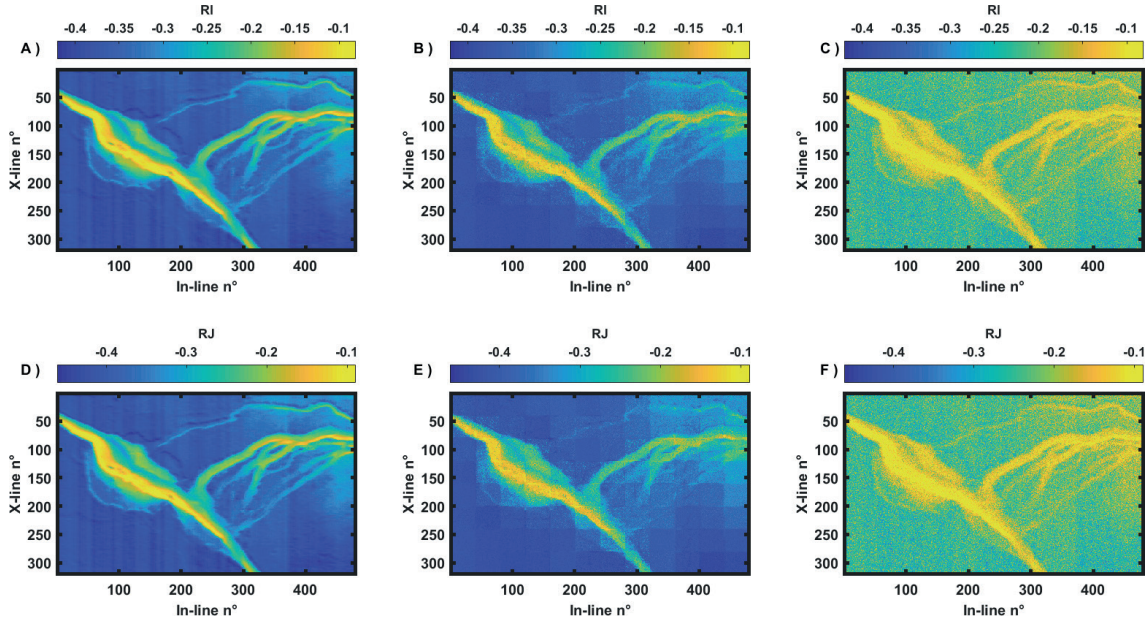

Fig. 9 - Comparison between the reference model (A and D), the solution provided by the Rj-McMC (B and E) and by the standard Bayesian approach (C and F) for noise standard deviation is equal to 0.070. R_I and R_J represent the relative contrasts in the *Ip* and *Is* values at the reflecting interface, respectively.

Table 1 - Correlation coefficients derived from the posterior PDFs estimated by the Rj-McMC and the standard Bayesian approach.

RR for noise standard deviation of 0.001	R_I	R_J
Rj-McMC	0.984	0.977
standard Bayesian	0.999	0.999
RR for noise standard deviation of 0.070	R_I	R_J
Rj-McMC	0.949	0.949
standard Bayesian	0.777	0.776

Obtained results confirm what was previously qualitatively highlighted, namely that for high SNR both approaches show a very good match with the synthetic data. The slightly lower match observed with the Rj-McMC for both the R_I and R_J parameters can be explained considering that in these synthetic cases, the γ_{max} value has been fixed to three-quarters of the maximum number of CDPs within the inverted area. By doing so, the algorithm generated models where naturally one or more CDPs shared the same elastic properties. This yielded a natural lower correlation with the synthetic data with respect to the standard Bayesian approach, where single CDPs were inverted. In the lower SNR case, where the ill-conditioning is severe, the differences between the outcomes of the two approaches can clearly be appreciated, in which the better match obtained with the Rj-McMC can be quantitatively calculated at about 17% more with respect to the standard Bayesian method.

The implemented inversion algorithm runs in approximately 12 hours on two compute nodes equipped with two deca-core Intel(R) Core(TM) i7-8550U @2.00 GHz (8 Gb RAM). Clearly, the computational cost is primarily related to the number of chains used, the maximum number of Voronoi cells the algorithm is allowed to create, the maximum number of iterations and the seismic data size (i.e. the considered range of angles of incidence over which the AVA is calculated). We observed that doubling the numbers of unknowns (i.e. bigger inversion areas or increased incidence angles) lead to an exponential increase of the computational cost, for which a further optimised code should be applied.

8. Conclusions

We implemented a Rj-McMC algorithm for target-oriented AVA inversion on 2D seismic data. The aim was to estimate the elastic variables and their lateral variability over a reference model characterised by strong lateral contrasts. Results have been benchmarked against those obtained with a standard Bayesian linear approach for different SNR. Due to the lack of observed seismic data, the proposed algorithm has been applied here only to synthetic seismic data. Since the convergence of the McMC algorithms is generally very slow, a parallel tempering approach was used to ensure optimal exploitation and exploration capabilities and to preserve the mutual correlation between the elastic properties. Our synthetic tests showed that for very high SNR the standard Bayesian approach fails to reconstruct the actual subsurface structures and produces a final prediction totally covered by noise. Differently, the proposed Rj-McMC ensures much more stable results, in which the lateral discontinuities are recovered accurately. The superior solution

provided by the proposed Rj-McMC method is ensured by its transdimensional framework that inherently adapts the parameterisation of the proposed models to the lateral variability of the data. Indeed, the implemented algorithms do not require any statistical test to choose the adequate model parameterisation and do not include any regularisation operator to force the model to obey some external constraints. Instead, the Rj-McMC automatically adjusts the underlying model parameterisation to produce solutions with an appropriate level of complexity to fit the data to statistically meaningful levels. Clearly, for a correct application of the proposed algorithm, a careful check of all the AVA inversion requirements must be made (improved data quality and controlled amplitude data processing).

REFERENCES

Aleardi M., Ciabarri F. and Gukov T.; 2018: *Reservoir characterization through target-oriented AVA-petrophysical inversions with spatial constraints*. Pure Appl. Geophys., **176**, 901-924, doi: 10.1007/s00024-018-2009-4.

Aster R.C., Borchers B. and Thurber C.H.; 2018: *Parameter estimation and inverse problems. 3rd ed*. Elsevier Academic Press, Cambridge, MA, U.S.A., 404 pp., ISBN: 9780128046517.

Ba J., Xu W., Fu L.Y., Carcione J.M. and Zhang L.; 2017: *Rock anelasticity due to patchy saturation and fabric heterogeneity: a double double-porosity model of wave propagation*. J. Geophys. Res., Solid Earth, **122**, 1949-1976, doi: 10.1002/2016JB013882.

Bodin T. and Sambridge M.; 2009: *Seismic tomography with the reversible jump algorithm*. Geophys. J. Int., **178**, 1411-1436, doi: 10.1111/j.1365-246X.2009.04226.x.

Brooks S.P., Giudici P. and Roberts G.O.; 2003: *Efficient construction of reversible jump Markov chain Monte Carlo proposal distributions*. J. R. Stat. Soc.: Series B (Stat. Method.), **65**, 3-39, doi: 10.1111/1467-9868.03711.

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

Chen H. and Zhang G.; 2017: *Estimation of dry fracture weakness, porosity, and fluid modulus using observable seismic reflection data in a gas-bearing reservoir*. Surv. Geophys., **38**, 651-678, doi: 10.1007/s10712-017-9410-x.

Cordua K.S., Hansen T.M. and Mosegaard K.; 2012: *Monte Carlo full-waveform inversion of crosshole GPR data using multiple-point geostatistical a priori information*. Geophys., **77**, H19-H31, doi: 10.1190/geo2011-0170.1.

Dosso S.E., Holland C.W. and Sambridge M.; 2012: *Parallel tempering for strongly nonlinear geoacoustic inversion*. J. Acoust. Soc. Am., **132**, 3030-3040, doi: 10.1121/1.4757639.

Dosso S.E., Dettmer J., Steininger G. and Holland C.W.; 2014: *Efficient trans-dimensional Bayesian inversion for geoacoustic profile estimation*. Inverse Prob., **30**, 114018, doi: 10.1088/0266-5611/30/11/114018.

Drufuca G. and Mazzotti A.; 1995: *Ambiguities in AVO inversion of reflections from a gas-sand*. Geophys., **60**, 134-141, doi: 10.1190/1.1443740.

Flournoy N. and Tsutakawa R.K. (eds); 1991: *Statistical multiple integration*. Proceedings 1989 Joint Summer Research Conference Held at Humboldt University, American Mathematical Society, Providence, RI, USA, vol. 115.

Gelman A., Carlin J.B., Stern H.S., Dunson D.B., Vehtari A. and Rubin D.B.; 2013: *Bayesian data analysis, 3rd ed*. Chapman and Hall-CRC Press, Texts in Statistical Science, Boca Raton, FL, USA, 676 pp., doi: 10.1111/j.1467-985X.2014.12096_1.x.

Geyer C.J.; 1991: *Markov chain Monte Carlo maximum likelihood*. In: Proc., Keramidas E.M. (ed.), 23rd Symposium on the Interface, Computing Science and Statistics, Interface Foundation of North America, Seattle, WA, USA, pp. 156-163.

Gouveia W.P. and Scales J.A.; 1998: *Bayesian seismic waveform inversion: parameter estimation and uncertainty analysis*. J. Geophys. Res., Solid Earth, **103**, 2759-2779, doi: 10.1029/97JB02933.

Grion S., Mazzotti A. and Spagnolini U.; 2002: *Joint estimation of AVO and kinematic parameters*. Geophys. Prospect., **46**, 405-422, doi: 10.1046/j.1365-2478.1998.970332.x.

Guo Q., Zhang H., Han F. and Shang Z.; 2018: *Prestack seismic inversion based on anisotropic Markov random field*. IEEE Trans. Geosci. Remote Sens., **56**, 1069-1079, doi: 10.1109/TGRS.2017.2758800.

Haas A. and Dubrule O.; 1994: *Geostatistical inversion - a sequential method of stochastic reservoir modelling constrained by seismic data*. First Break, **12**, 561-569, doi: 10.3997/1365-2397.1994034.

Hansen T.M., Cordua K.S. and Mosegaard K.; 2012: *Inverse problems with non-trivial priors: efficient solution through sequential Gibbs sampling*. Comput. Geosci., **16**, 593-611, doi: 10.1007/s10596-011-9271-1.

Hastings W.K.; 1970: *Monte Carlo sampling methods using Markov chains and their applications*. Biometrika, **57**, 97-109, doi: 10.1093/biomet/57.1.97.

Lan S., Streets J. and Shahbaba B.; 2014: *Wormhole Hamiltonian Monte Carlo*. In: Proc., 28th AAAI Conference on Artificial Intelligence, Québec City, QC, Canada, pp. 1953-1959.

Metropolis N. and Ulam S.; 1949: *The Monte Carlo method*. J. Am. Stat. Assoc., **44**, 335-341, doi: 10.1080/01621459.1949.10483310.

Metropolis N., Rosenbluth A.W., Rosenbluth M.N., Teller A.H. and Teller E.; 1953: *Equation of state calculations by fast computing machines*. J. Chem. Phys., **21**, 1087-1092., doi: 10.1063/1.1699114.

Mosegaard K. and Tarantola A.; 1995: *Monte Carlo sampling of solutions to inverse problems*. J. Geophys. Res., Solid Earth, **100**, 12431-12447, doi: 10.1029/94JB03097.

Sambridge M.; 2013: *A parallel tempering algorithm for probabilistic sampling and multimodal optimization*. Geophys. J. Int., **196**, 357-374, doi: 10.1093/gji/ggt342.

Sambridge M. and Mosegaard K.; 2002: *Monte Carlo methods in geophysical inverse problems*. Rev. Geophys., **40**, 3\1-3\29, doi: 10.1029/2000RG000089.

Sen M.K. and Biswas R.; 2017: *Transdimensional seismic inversion using the reversible jump Hamiltonian Monte Carlo algorithm*. Geophys., **82**, R119-R134, doi: 10.1190/geo2016-0010.1.

She B., Wang Y., Liang J., Liu Z., Song C. and Hu G.; 2018: *A data-driven amplitude variation with offset inversion method via learned dictionaries and sparse representation*. Geophys., **83**, R725-R748, doi: 10.1190/geo2017-0615.1.

Tetyukhina D.; 2011: *High-resolution reservoir characterization by seismic inversion with geological constraints*. Geophys., **76**, Z2, doi: 10.1190/1.3554713.

Ursenbach C.P. and Stewart R.R.; 2008: *Two-term AVO inversion: equivalences and new methods*. Geophys., **73**, C31-C38, doi: 10.1190/1.2978388.

Corresponding author: Alessandro Salusti
Dipartimento di Scienze della Terra, Università di Pisa
Via Santa Maria 53, 56126 Pisa, Italy
Phone: +39 050 2215752; e-mail: alessandro.salusti@unifi.it