

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/268810044>

Multidimensional scaling for the evaluation of a geostatistical seismic elastic inversion methodology

Article in *Geophysics* · January 2014

DOI: 10.1190/geo2013-0037.1

CITATION

1

READS

84

6 authors, including:



Pedro Correia

Geovariances

26 PUBLICATIONS 6 CITATIONS

SEE PROFILE



Ruben F. M. Nunes

Technical University of Lisbon

18 PUBLICATIONS 67 CITATIONS

SEE PROFILE



A. Soares

Technical University of Lisbon

145 PUBLICATIONS 1,273 CITATIONS

SEE PROFILE



Guenther Schwedersky neto

Petróleo Brasileiro S.A.

26 PUBLICATIONS 14 CITATIONS

SEE PROFILE

Multidimensional scaling for the evaluation of a geostatistical seismic elastic inversion methodology

Leonardo Azevedo¹, Ruben Nunes², Pedro Correia², Amílcar Soares²,
Luis Guerreiro³, and Guenther Schwedersky Neto⁴

ABSTRACT

Due to the nature of seismic inversion problems, there are multiple possible solutions that can equally fit the observed seismic data while diverging from the real subsurface model. Consequently, it is important to assess how inverse-impedance models are converging toward the real subsurface model. For this purpose, we evaluated a new methodology to combine the multidimensional scaling (MDS) technique with an iterative geostatistical elastic seismic inversion algorithm. The geostatistical inversion algorithm inverted partial angle stacks directly for acoustic and elastic impedance (AI and EI) models. It was based on a genetic algorithm in which the model perturbation at each iteration was performed recurring to stochastic sequential simulation. To assess the reliability and convergence of the inverted

models at each step, the simulated models can be projected in a metric space computed by MDS. This projection allowed distinguishing similar from variable models and assessing the convergence of inverted models toward the real impedance ones. The geostatistical inversion results of a synthetic data set, in which the real AI and EI models are known, were plotted in this metric space along with the known impedance models. We applied the same principle to a real data set using a cross-validation technique. These examples revealed that the MDS is a valuable tool to evaluate the convergence of the inverse methodology and the impedance model variability among each iteration of the inversion process. Particularly for the geostatistical inversion algorithm we evaluated, it retrieves reliable impedance models while still producing a set of simulated models with considerable variability.

INTRODUCTION

Geophysical inverse problems aim to infer properties of the subsurface geology, such as acoustic and elastic impedance (AI and EI) models, from a set of indirect geophysical measurements. Particularly for seismic inversion problems, the retrieved inverse solution is nonlinear, ill conditioned, and nonunique due to the intrinsic limitations of the geophysical method: the limited bandwidth and resolution of the seismic data, noise, measurement errors, and physical assumptions about the involved forward models (Tarantola, 2005). The best-fit AI and EI models, retrieved at the end of a seismic inverse process, are just one possibility among several earth models that satisfy the observed seismic data. In other words,

if the match between the real and the synthetic seismic data, created with the best AI/EI pair at the end of an elastic inversion process, is poor, one can conclude that the correspondence between the real AI/EI models with the inverted ones is also poor. However, the opposite may not be true: A good match between observed and inverted synthetic seismic data does not ensure that the real and inverted AI and EI models are converging to each other; the inverse solution is converging toward a local minimum instead of toward a global minimum (Tarantola, 2005). Therefore, there is a variable degree of uncertainty associated with any inverted subsurface AI and EI models. This uncertainty needs to be assessed, not neglected, during the interpretation of these inverse models (Bosch et al., 2010; Tompkins et al., 2011).

Manuscript received by the Editor 31 January 2013; revised manuscript received 21 September 2013; published online 27 November 2013.

¹Universidade Técnica de Lisboa, Center for Modeling Petroleum Reservoirs, Cerena/DECivil, Instituto Superior Técnico, Lisbon, Portugal and Universidade de Aveiro, CESAM & Departamento de Geociências, Campus Universitário de Santiago, Aveiro, Portugal. E-mail: leonardo.azevedo@ua.pt.

²Universidade Técnica de Lisboa, Center for Modeling Petroleum Reservoirs, Cerena/DECivil, Instituto Superior Técnico, Lisbon, Portugal. E-mail: nunessrfm@gmail.com; pedrocorreia1984@gmail.com; asoares@ist.utl.pt.

³Partex Oil & Gas, Rua Ivone Silva, Lisboa, Portugal. E-mail: lguerreiro@partex-oilgas.com.

⁴Petrobras Research Center, Rio de Janeiro, Brazil. E-mail: guenther@petrobras.com.br.

© 2013 Society of Exploration Geophysicists. All rights reserved.

Seismic inverse problems may be developed following two main different approaches: a deterministic, or optimization, one, and by posing the inverse problem in a probabilistic basis. Within the deterministic framework, the sparse-spike and model-based methodologies are the most widespread inversion techniques among the geophysical community (Russel, 1988; Bosch et al., 2010). The uncertainty assessment of deterministic solutions is limited because it is represented by a linearization problem around least-squares inverse solutions. Under this linear assumption, deterministic solutions lack the wide exploration of the uncertainty space. On the other hand, probabilistic approaches ensure the propagation of the uncertainty from the prior probability distributions, estimated from experimental data (e.g., well-log data) to the distributions of the model parameters space (Grana et al., 2012). In probabilistic solutions, the uncertainty space is considerably more explored when compared with deterministic ones because the inverse solution is a probability density function on the model parameters space. These methodologies overcome the linearity assumptions of the deterministic solutions. However, the uncertainty assessment depends greatly on the parameterization of the inverse problem, e.g., assumptions about the prior distributions and the spatial continuity pattern (Scales and Tenorio, 2001; Tarantola, 2005).

Geostatistical seismic inversion techniques may be divided in two main methodologies: Bayesian linearized and based on sequential simulation algorithms (Doyen, 2007). Bayesian linearized solutions allow the definition of the inverse problem in a well-defined mathematical formulation by assuming a Gaussian distribution of the prior probability distributions of the variables to invert and the linearization of the forward model operator. These assumptions enable an analytical solution for the inverted posterior distributions, estimated as being Gaussian (Loertzer and Berkhout, 1992; Buland and Omre, 2003; Tarantola, 2005). However, most times the prior probability distributions of the properties to invert are not close to Gaussian. In fact, in real data sets, they are most often multimodal. The Gaussian assumption was later overcome by Grana and Della Rosa (2010), who developed the mathematical formalism for the Bayesian linearized inversion assuming Gaussian mixture models to estimate the prior probability distributions. In their approach, the posterior probability distribution is analytically expressed as Gaussian mixture models as well. Gaussian linearized methods allow mathematically tractable solutions but lack the real exploration of the uncertainty space because the posterior probability distributions are assumed to be Gaussian, or Gaussian mixture.

Geostatistical inverse methodologies based on stochastic simulation algorithms overcome the need for Gaussian assumption about the prior probability distributions or the linearization of the forward operator. Genetic algorithms (e.g., Mallick, 1995, 1999; Boschetti et al., 1996; Soares et al., 2007) and simulated annealing (Sen and Stoffa, 1991; Ma, 2002) fall within this class of inverse methodologies. Stochastic solutions are achieved by sampling the parameter model space directly from the prior probability distribution. Because the assumptions within this framework are less strict when compared with the Bayesian linearized assumptions, the uncertainty space is considered to be more extensively sampled. On the other hand, one loses the mathematical tractability of the inverse solution, increasing the computational costs involved (Tarantola, 2005; Bosch et al., 2010). Iterative stochastic seismic inversion methodologies (Bortoli et al., 1992; Haas and Dubrule, 1994; Soares et al.,

2007; González et al., 2008; Grana et al., 2012; Nunes et al., 2012) solve the seismic inverse problem recurring to stochastic sequential simulations and cosimulations algorithms (Deutsch and Journel, 1998) as the perturbation technique of the elastic models at each iteration. Briefly, these inversion algorithms can be described by the following sequence of steps: First, a set of equivalent earth models, the model parameter space, is simulated at each iteration using a predefined stochastic sequential simulation algorithm based on prior information, usually from available well-log data, available secondary information (e.g., normally information about a background trend), and a spatial continuity pattern (e.g., variogram, training image). Two different approaches can be considered concerning the simulation of the impedance models: trace-by-trace methods (Bortoli et al., 1992; Haas and Dubrule, 1994), where just one trace is simulated at the time, and global inversion methods (Soares et al., 2007), where the entire model is simulated and cosimulated at each iteration. Then, for each simulated model, and for every location on the simulation grid, synthetic seismic traces are calculated by convolving an estimated wavelet with the reflection coefficients computed from the impedance models. The resulting synthetic trace is, afterward, compared with the corresponding recorded seismic trace in terms of correlation coefficient. In trace-by-trace algorithms, the best impedance trace is retained as conditioning data and the simulation procedure continues until all the traces are simulated. In the global inversion methodologies, the AI and EI traces that ensure the highest correlation coefficient between the real and the synthetic seismic from the entire simulation ensemble are used as seed for the generation of impedance models for the next iteration. One family of inverse algorithms is based on crossover genetic algorithms. The iterative procedure stops when the global correlation coefficient between the entire synthetic and real seismic volumes is above a certain threshold.

In geostatistical seismic inverse solutions, it is of great importance to assess how the inverted models are representative of the plausible model parameter space: if this space is being explored widely, or if on the contrary all the inverted models are located around a narrow region of this space. It is essential to ensure that the inverted models are converging toward the real ones while the inverted seismic converge toward the real one. For real data, one does not know what the real earth's impedance model is; therefore, the convergence of the method may be assessed for example at the location of existing wells not used to constraint the inversion procedure.

Multidimensional scaling (MDS) (Cox and Cox, 1994) is a mathematical tool that enables the representation of the uncertainty between different earth models based on the concept of distances (Suzuki and Caers, 2008; Scheidt and Caers, 2009). By computing the dissimilarity between the inverted and real earth models, which refers to the notion of distance between models, one is able to evaluate how good the uncertainty of the parameter model space is being sampled (Caers, 2011). Similar models in terms of the spatial distribution of its internal properties will be plotted closer, or in a cluster, in the MDS space, while on the other hand, models with larger differences are plotted far away in the same metric space.

In this paper, we introduce a novel approach to combine the MDS for assessing the performance of geostatistical seismic inversion methodologies, in this case, the global elastic inversion (GEI) (Nunes et al., 2012) in two complementary ways: in terms of how well the retrieved inverse acoustic and elastic models converge

toward the reality, while the mismatch between the real and the inverted seismic data is being minimized, and how well the inverted models explore the parameter model space. The proposed approach was tested in two different examples: a 3D synthetic seismic data set from where the real AI and EI models are known for the entire study area and a real seismic data set from where the real AI and EI models are sparsely known at the well location. In the first example, the MDS was used to project in the metric space the ensemble of simulated acoustic and elastic models, derived from the geostatistical iterative seismic inversion procedure and the real AI and EI models. This allowed the comparison between the spatial location, at the MDS referential, of inverted and real impedance models. By plotting all the models in the MDS space, we could access how the simulated models of AI and EI evolve while the inverse method converges to the desired solution — the match between synthetic seismograms and the real seismic. In addition, we assessed how the parameter model space is being explored. If the sampling of the plausible parameter model space, the space of all possible AI and EI models given a prior distribution, is significant or if the simulated models are all located around a particular location, for example, around a local minimum. In the second example, we inverted a real seismic data set composed by two partial angle stacks and five wells. The geostatistical inversion procedure ran five times independently. On each run, a different well was removed from the conditioning data and the impedance models were compared at the well location not used in that specific inversion run. Beside the apparent spatial clusterization of the available wells, this cross-validation technique ensures that the global prior probability distributions for acoustic and elastic properties as estimated from the well-log is fairly reproduced on each of the five runs. In addition, by removing only one well at the time the local conditional distributions built during the stochastic simulation of the impedance models is not significantly changed, allowing the comparison of the results between runs. The results of these two case studies show a rapid convergence of the simulated earth models toward the real solution at the end of six iterations. The convergence is achieved without compromising the exploration of the model parameter space, particularly for all the simulated pair of acoustic and elastic models created during iteration one.

METHODOLOGY

Global elastic inversion

The GEI algorithm is an iterative geostatistical seismic inversion methodology that allows the inversion of partial angle stacks directly for AI and EI models. It is based on two main ideas: the use of stochastic sequential simulation at each iteration as the perturbation technique of the inverted elastic models and the use of a genetic algorithm as a global optimizer to converge the simulated models toward the solution. The local correlation coefficients between each individual seismic trace from the real and the corresponding inverted partial angle stack are used as the affinity criterion to create the next generation of models. The objective function, which determines when the iterative process stops, is defined by the global correlation coefficient between the real and the inverted seismic data for all angle stacks simultaneously.

As part of the geostatistical inversion methodology, we choose the direct sequential simulation (DSS) (Soares, 2001) and the co-DSS with joint probability distributions (Horta and Soares,

2010) as the simulation algorithms to generate AI and EI models, respectively. The great advantage of DSS, when compared with other traditional sequential simulation algorithms such as sequential Gaussian simulation (Deutsch and Journel, 1998), is that it allows the use of a prior distribution directly estimated from the well-log data without the need of a Gaussian transform. The use of co-DSS with joint probability distributions ensures the reproduction of the joint probability distribution between the simulated AI and EI models as estimated from the well-log data. In stochastic sequential simulation, the reservoir grid is simulated following a predefined random path that visits all the nodes of the simulation grid. Following the predefined random path, and at each node of the simulation grid, the local mean and variance are estimated by a simple kriging estimate. Then, a value is drawn from the global distribution function (DSS and co-DSS) (Soares, 2001) or from the local conditional probability distribution (co-DSS with joint probability distributions) (Horta and Soares, 2010). Different sequential simulation runs, commonly designated as realizations, produce distinct models because the random path changes on each run and consequently, the conditioning data at every grid node change as well. Each realization reproduces the experimental data (e.g., well-log data) at its location, the prior probability distribution of the simulated property and a spatial continuity pattern, estimated by a variogram in the two-point geostatistics framework.

The GEI is performed at the seismic scale, with a typical vertical cell size between 1 and 4 ms. Prior to the inversion, the impedance data from well-logs, with high resolution, need to be upscaled into the simulation grid scale, with lower resolution. The upscaling method should be selected cautiously because it can dramatically change the estimation of the prior probability distribution. A discussion about upscaling methodologies is outside the scope of this work. In our examples, we chose an upscaling method that ensures the reproduction of the first- and second-order statistics (mean and variance) of the original well-log data. In the proposed methodology, the spatial continuity pattern is imposed by a variogram model. The horizontal and vertical ranges of the variogram are commonly modeled from the available well-log data. However, many times the number of available wells does not allow modeling reliable horizontal ranges. In these situations, it is a common practice to estimate the variogram horizontal extent directly from the seismic reflection data, which will often result in an overestimation of the range value.

The GEI workflow can be summarized in the following steps:

- 1) Simulation of a set of N_s AI models for the entire simulation grid using DSS, given the available AI well-log data. Each individual simulated model reproduces the observed data at the well location, the estimated prior distribution (Figure 1), and the spatial continuity patterns as revealed by the variogram model (Figure 2).
- 2) Cosimulation of a set of N_s EI models, given the available EI data from well logs and the previous simulated AI models (used as a collocated secondary variable in the simulation of EI) recurring to the co-DSS with the joint-distributions algorithm (Horta and Soares, 2010). The secondary variable is used to ensure the reproduction of the joint probability distribution of AI versus EI (Figure 3). All the simulated EI models reproduce the observed data of EI at the well location, the spatial pattern as revealed by the variogram model, and the joint probability distribution between AI and EI as estimated from the hard data.

- 3) Calculation of N_s synthetic partial angle stacks from the N_s simulated duplet of (AI, EI). The reflection coefficients are computed following Fatti's approximation (Fatti et al., 1994) and are then convolved with an estimated wavelet for that specific partial angle stack. The underlying assumptions on the wavelet estimation are outside the scope of this work.
- 4) In a trace-by-trace basis, calculate the correlation coefficient between each synthetic trace from the various N_s synthetic angle stacks with the correspondent trace from the observed real partial angle stack. This step results in the calculation of N_s local correlation cubes per angle.
- 5) The elastic traces that ensure the highest local correlation coefficient between observed and synthetic seismic data are selected to build the "best" AI and EI models. The corresponding

correlation coefficient is also stored in a best local correlation volume. These best AI and EI models are then used as secondary variables for the cosimulation process of the next iteration along with the collocated local correlation coefficient cube. In this way, the elastic traces associated with low correlation coefficients will be able to change significantly in the next iteration while the ones associated with higher correlation values will only suffer small-scale variations. This procedure guarantees the local and global convergence of the inverted seismic toward the real one.

- 6) Generation of a new set of AI and EI models using the best impedance model, along with the local correlation coefficients by direct sequential cosimulation. Return to step 3 until the global correlation coefficient between the real and synthetic seismic simultaneously for all the angle stacks is above a certain threshold.

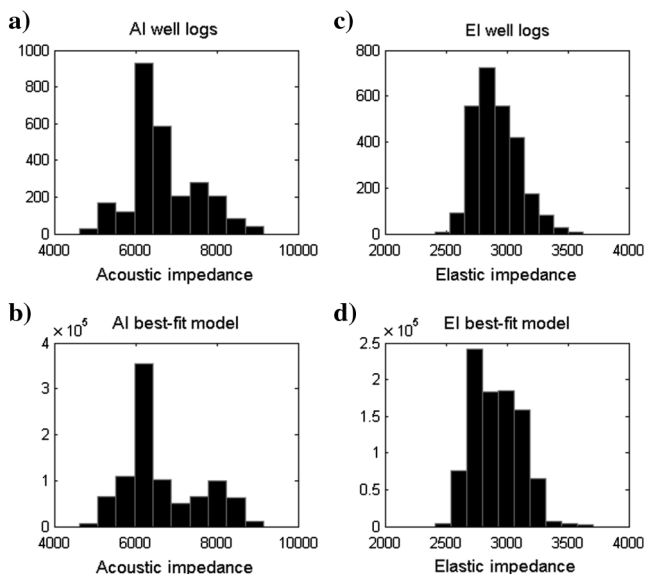


Figure 1. AI and EI histograms calculated from (a and c) available well-log data and (b and d) the inverse best-fit acoustic and elastic models. The inverted models reproduce the original distribution as estimated from the experimental data.

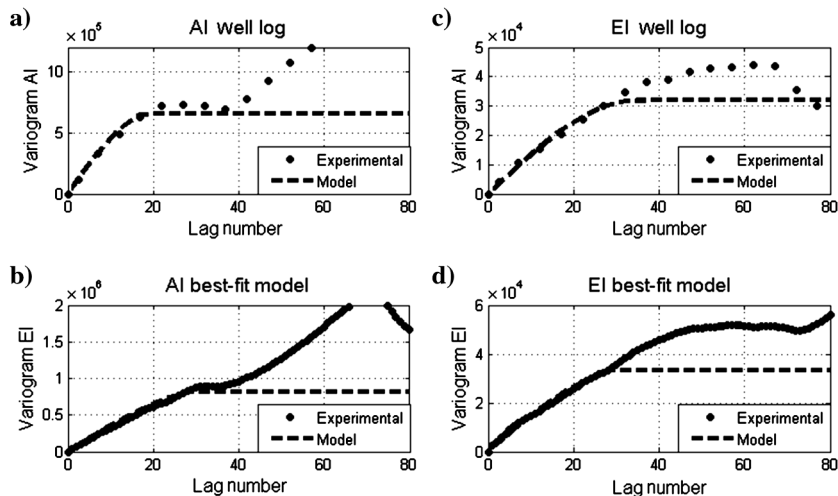


Figure 2. Comparison between the vertical variograms estimated from the available (a and c) well-log data and the (b and d) best-fit inverse model of AI and EI.

MDS

MDS is a multivariate statistical technique based on the concept of distances, which is able to reveal in few dimensions patterns between a set of multidimensional models (Caers, 2011). Within the proposed methodology, the MDS was applied to assess how the simulated acoustic and elastic models produced by the GEI evolve from iteration to iteration.

In a metric space, such as that created by the MDS, the relative position between several simulated earth models is directly related to how similar these models are in terms of their internal configuration. Briefly, the MDS (the detailed mathematical description about the method is presented in Appendix A) is a statistical methodology for uncertainty assessment in stochastic frameworks. It converts a dissimilarity matrix (\mathbf{D}) into points, which can then be plotted in a Cartesian space, the MDS space (Cox and Cox, 1994; Borg and Groenen, 1997; Caers, 2011). First, matrix \mathbf{D} is converted into matrix \mathbf{A} by a scalar product. Then, \mathbf{A} is decomposed by eigenvector decomposition where only the first d principal components, or eigenvectors, are retained (Scheidt and Caers, 2009; Caers, 2011). In the MDS space, the distances between projected points, where each point represents an earth model (e.g., AI models),

are proportional to the similitude, in terms of its internal properties, between different models. In our examples, this allowed the assessment of how the uncertainty of the solution model space is being explored by the ensemble of all the simulated models derived from the GEI. If the models that comprise a simulation ensemble are all very similar among themselves they will be projected in a cluster in the new metric space produced by the MDS (Figure 4). On the other hand, if the models are very different among themselves they will be projected with greater distances among themselves and consequently exploring a larger region of the uncertainty model space (Figure 4). because the MDS space only cares about the relative distances between earth models, the coordinates of each point, that correspond to a unique AI/EI model, itself is not important. In fact, the MDS plot may suffer rotation, translation and reflection, without any loss of information (Caers, 2011).

Particularly, for our examples we begin by calculating the matrix \mathbf{D} between each AI/EI model that belong to the simulation ensemble created during the GEI (all realizations per iteration and the best AI/EI model computed at the end of each iteration) using the Euclidean distance:

$$d_{i,j} = \sqrt{\sum_{i,j=1}^N (x_i - x_j)^2}, \quad (1)$$

where \mathbf{x}_i and \mathbf{x}_j are a pair of earth model vectors with dimension N and the simulation ensemble (\mathbf{X}) is of size $L \times N$, with $\mathbf{X} = [\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_L]^T$. Then, we applied the classical MDS methodology to the matrix \mathbf{D} (Cox and Cox, 1994; Borg and Groenen, 1997; Caers, 2011).

EXAMPLES

For the examples shown in this section, we propose the combination of the MDS as a statistical tool to evaluate how the GEI algorithm explores the parameter model space from iteration to iteration and how the simulation ensemble is evolving from iteration to iteration toward the objective. This methodology was applied to two different scenarios: a 3D synthetic seismic data set, from where the real AI and EI models are known, allowing the comparison between the location of the inverse models created during the iterative process and the location of the real impedance models in the MDS space and to a 3D real seismic data set, where several GEI were run using a cross-validation methodology. We ran five different GEI, as many as the number of available wells, and on each geostatistical inversion run, a different well was removed from the conditioning data. In all five runs, the assumptions about the spatial continuity pattern were kept the same as estimated from the entire set of available wells. With this cross-validation scheme, we ensure that for all the independent inversion runs, the spatial continuity pattern remains fairly constant and the prior probability distributions are only slightly changed. This allows a comparison between the results of the five runs because the assumptions about the priors are similar. At the end of each inversion process, we then plotted the observed AI and EI models from the well-log data along with all the simulated models of AI and EI, which were computed as part of the GEI at the same well location. In this way, we were able to evaluate the convergence of the GEI at each well location individually. We took for granted all the intrinsic uncertainties related with the well-log data acquisition and processing and assumed to be the real impedance values of the subsurface geology.

Synthetic example

In this example, we used a synthetic 3D seismic reflection data with four partial angle-stacks, mean angles of 5°, 15°, 25°, and 35°, from where the real AI and EI models are known. This synthetic data set was built based on a real deep-water turbidite field and has a total size of $101 \times 101 \times 90$ blocks in the i -, j -, and k -directions, respectively. A log set of 31 wells with AI and EI well logs and the

wavelet used for the calculation of the synthetic seismic data, designated from now on as real seismic data, were also available.

We inverted the synthetic seismic data set using the GEI methodology. The geostatistical inversion converged after six iterations, simulating 32 pairs of AI/EI per iteration. It resulted in a total of 384 impedance models (192 AI and 192 EI modes). The global correlation coefficient for all the partial angle stacks between the observed and the inverted seismic reflection data computed from the best pair of AI and EI models is greater than 0.75. The retrieved best-fit models are visually similar with the real ones, and they honor the marginal and joint distributions of AI and EI (Figures 1 and 3), estimated from the original well-log data, the spatial continuity pattern as revealed by a variogram model (Figure 2), and the data from the well set used as conditioning data (Figure 5). Larger differences between the inverted and the real elastic models are located in areas far from the well locations. These areas will

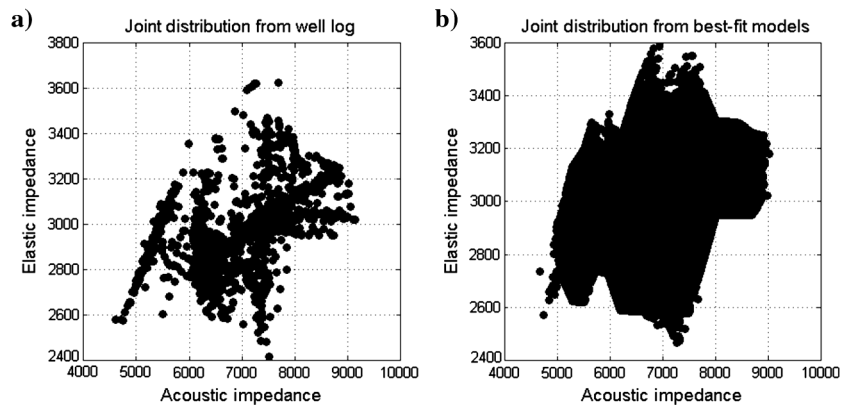


Figure 3. Joint distribution between AI versus EI estimated from (a) the available well-log and (b) the best-fit inverse models.

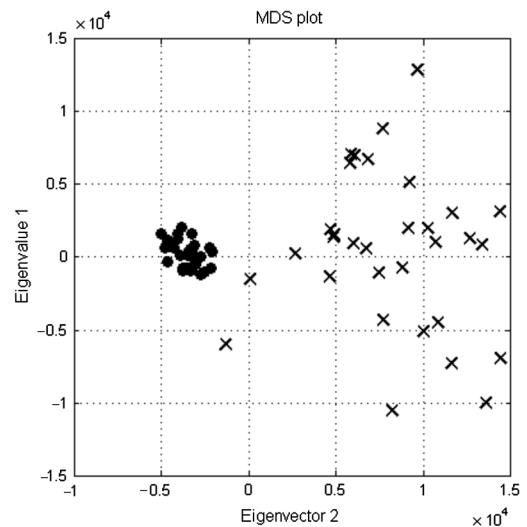


Figure 4. MDS plot showing the difference of relative position between similar models (filled circles) and considerably different models (crosses). The distribution of the models in the uncertainty space allows an estimation of how this space is being sampled; the models represented by the crosses explore far more the uncertainty space when compared with the models represented by the circles.

preserve higher variability among the models simulated within the same iteration, representing areas of higher uncertainty because they are less constrained by the experimental data. Then, we calculated the Euclidean distances between each of the 192 acoustic/elastic models, the best impedance model created at the end of each iteration, and the real AI and EI models, respectively. The resulting MDS plot for an individual impedance domain contains a total of 199 points, each representing one of the models referred above.

By plotting all the acoustic and elastic models in the MDS space (Figure 6), it is easily recognizable that the parameter model space is considerably well explored by all the models computed during the inversion procedure. This space was created by retaining the first three eigenvalues that explain about 75% variance of the original model space. By plotting dimension two versus dimension three, it is easily recognizable that the inversion algorithm is converging toward the real AI and EI models from iteration to iteration. The same behavior is observed in the space defined by the other dimensions that belong to the reduced space. Figure 6 shows that, as the iteration number increases, the gravity center corresponding to the cloud of a particular iteration moves toward the real impedance model while the match between synthetic and real seismic is also being improved. However, in the MDS space, there is the possibility that a particular simulated model, mainly from the first iterations, achieves a shorter distance to the real models when compared with the real impedance model. Nevertheless, due to the nonunique solution nature of the seismic inverse problem, these models cannot ensure a good correlation between the corresponding synthetic and the real seismic. Figure 6 also shows a slower convergence of the elastic domain when compared with the acoustic one. The convergence of EI to a local minimum can be a possible explanation for the slowest convergence of EI. However, it is worth noting that the proposed methodology assures a joint convergence of AI and EI instead of assuring

the individual best model of AI and EI. In fact, the convergence of the GEI in terms of similitude between the real and the inverted synthetic seismic is evaluated by minimizing the objective function, the global correlation coefficient between the inverted synthetic seismic and the real seismic (see the Methodology section).

All the simulated models of iteration 1 explore considerably well the MDS space. Once an iteration is finished, the best impedance model from that iteration approximates, in terms of distance, and therefore in terms of similitude, toward the real impedance model. The next set of simulated models will preferably occupy the space around the best model from the previous iteration. This effect is directly related with the GEI methodology because the best impedance volume, computed at the end of the current iteration, is used as a secondary variable in the geostatistical cosimulation process of the next iteration.

The MDS is an efficient technique to evaluate the GEI in terms of its convergence in the model parameters space. It allows the assessment of how the inverted AI and EI models evolve from iteration to

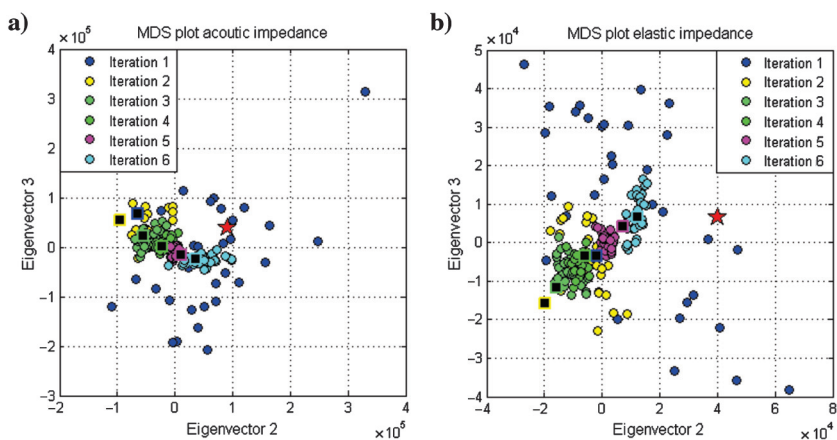


Figure 6. MDS plots for the synthetic example: (a) AI models and (b) EI models. True models are represented by the red star, and the best models, computed at the end of each iteration, are represented by the squares filled in black.

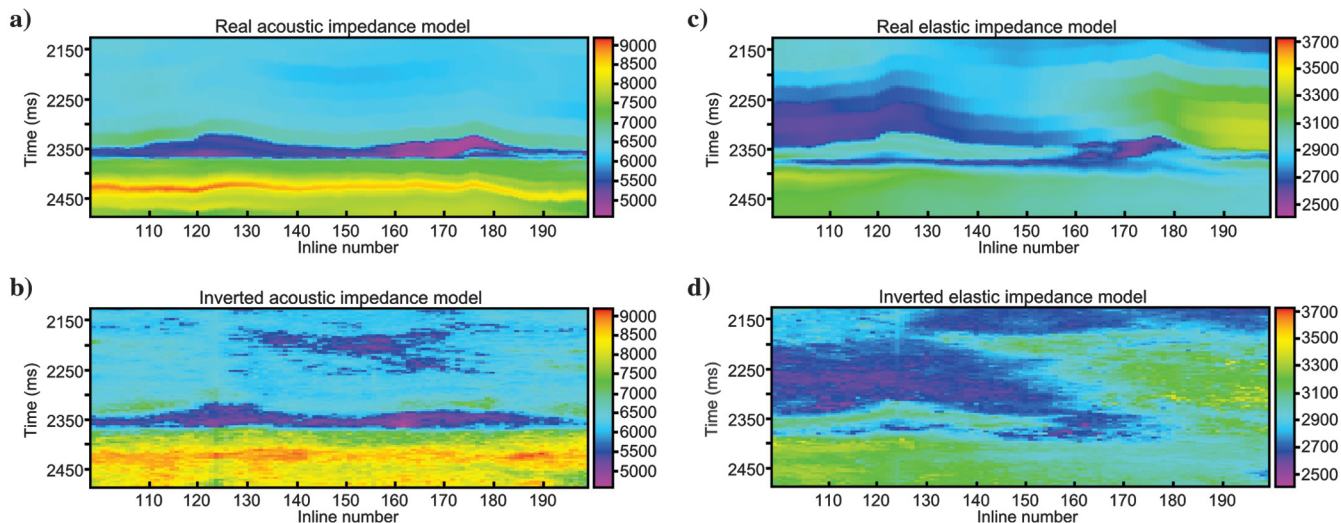


Figure 5. Comparison between a vertical section of the (a and c) real elastic model and a (b and d) vertical section extracted from the inverse best-fit model of (a and b) AI and (c and d) EI.

iteration comparing to the location of the inverted model against the real ones.

Real data example

In a field seismic data set, one does not know the real AIs and EIs that produced the acquired seismic data. In fact, this reality is only sparsely known at the well locations. Consequently, the retrieved impedance models from any seismic inversion process cannot be directly compared. For this reason and to assess the GEI convergence with real seismic reflection data sets, we propose a cross-validation technique to an onshore data set. The available data are composed of two partial angle stacks, corresponding to mean angles of 15° and 27.5° , over an area of $336 \times 212 \times 49$ cells, in the i -, j -, and k -directions, respectively, and five wells with AI and EI well logs. A total of five GEI were run. Each run comprised six iterations and 32 simulations per iteration. For all the GEI runs, a different well, from the available five, was removed and not used as conditioning data. This cross-validation techniques allowed the comparison of the inverted models at the well location against the observed AI and EI models from the well-log data. Beside the apparent spatial clusterization of the well logs, we chose this cross-validation technique because we want to guarantee that the global parameters such as prior probability distributions and the spatial variogram model remain fairly constant during the five independent runs. If we retain at once from the conditioning data wells 3, 4, and 5, apparently located in a spatial cluster, the assumptions about priors and the spatial continuity pattern would be considerably different and therefore the results of the inversion would change dramatically. In addition, wells 3, 4, and 5 are far enough when compared with the horizontal variogram ranges used for the inversion. The best five pairs of AI an EI model produced synthetic seismic data sets with global correlation coefficient values between the real and the inverted synthetic angle stacks of about 0.80. We then calculated the Euclidean distances between each of the 192 acoustic/elastic models, the best impedance models created at the end of each iteration, and the real AI and EI models, respectively. Therefore, the MDS plot, for an individual impedance domain, contains a

total of 199 points, each representing one of the models referred above at the location of the well removed from the conditioning data for that particular seismic inversion.

Globally, we may say that at the well locations for AI and EI models, the inversion method is converging toward the real AI and EI well-log data (Figure 7). For all the examples, we retained only the first three eigenvalues. This reduced space explains around 70% of the variance of the original AI model space and around 65% of the variance of the original EI model space. Figure 7 shows the dimensions where the convergence of the models is clearly observed. Remember that each point in the MDS plot represents an impedance model located at the well location not used as conditioning data of the GEI. It is possible to observe that the plausible parameter model space is considerably well explored from iteration to iteration. However, the convergence, shorter distances between inverted and the real impedance models at the well location, is not equal for all the available wells. Consistently for the acoustic and the elastic domain, well 1 and well 2 approximate worst toward the solution when compared with the other three wells. This effect is summarized in Figure 8. For all five GEI runs, we plotted the Euclidean distance, calculated in the 3D reduced space, between the best model at the end of each iteration and the observed AI model, normalized by the greatest distance of all the best AI models. It is clear that wells 1 and 2 converge far less than wells 3, 4, and 5 for the elastic and the acoustic domains. Nevertheless, the results are consistent with the ones obtained for the synthetic case: While the synthetic seismic converges in terms of correlation coefficient toward the real, the associated AI and EI models converge toward the real ones.

DISCUSSION

From the interpretation of the MDS plots for both the examples shown here (Figures 6 and 7), it is clear that the inverted models for AI converge better and faster toward the real AI model when compared with the inverted EI models. This directly refers to the degree of uncertainty associated with each inverted property and to the

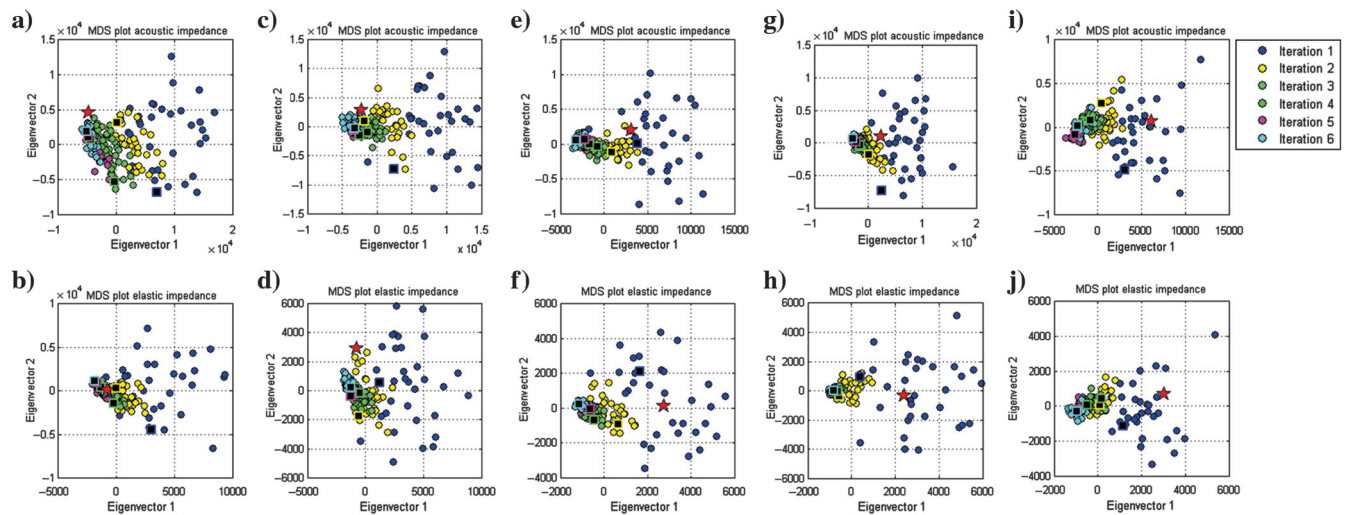


Figure 7. From left to right: MDS plots for all models produced by the GEI without using (a and b) well 1, (c and d) well 2, (e and f) well 3, (g and h) well 4, and (i and j) well 5 as conditioning data: (a, c, e, g, and i) AI models and (b, d, f, h, and j) EI models. True models are represented by the red star, and the best models, computed at the end of each iteration, are represented by the squares filled in black.

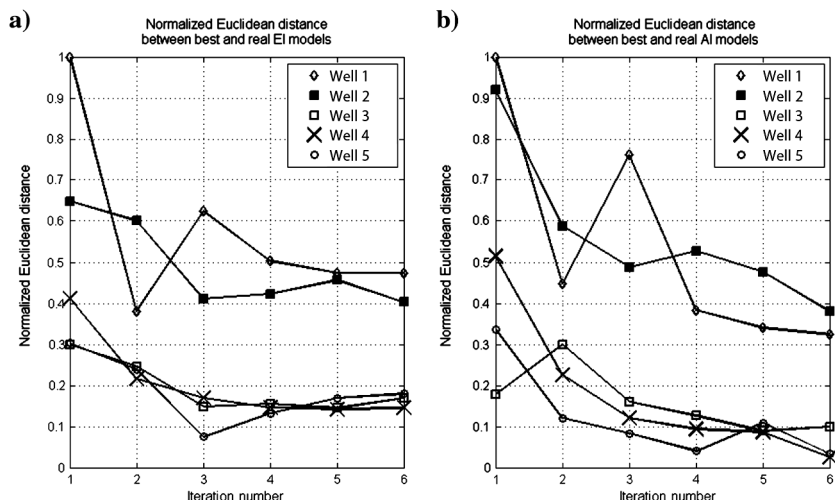


Figure 8. (a) Euclidean distance between the best AI models computed at the end of each iteration and the observed AI model. (b) Euclidean distance between the best EI models computed at the end of each iteration and the observed EI model.

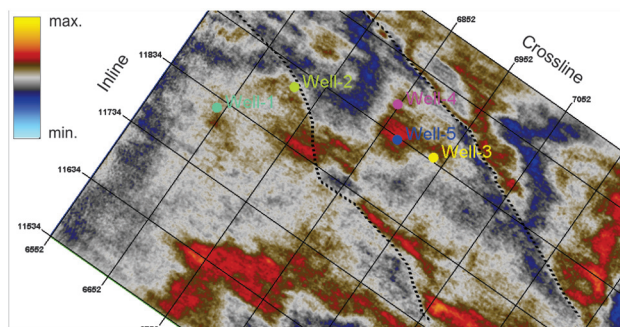


Figure 9. Horizontal time slice from the real seismic volume extracted at the reservoir level and well locations. Notice that wells 4-6 are located on a different fault block when compared with the location of wells 1 and 2. Dashed black lines represent both fault directions.

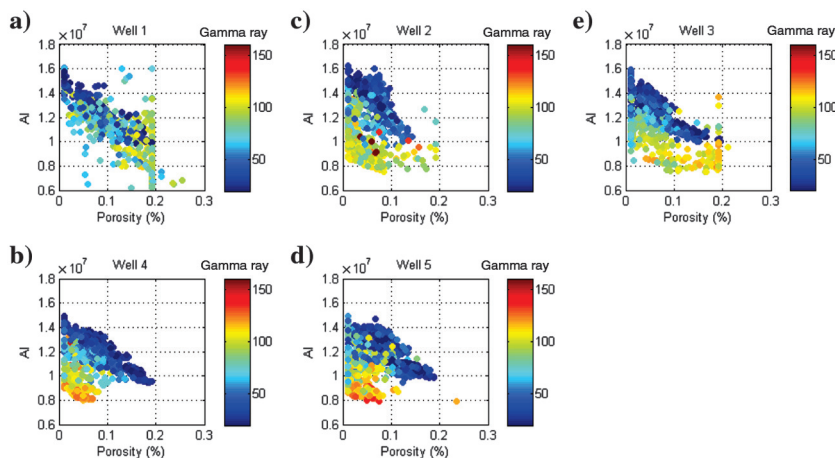


Figure 10. Well-log data for all the available wells: AI versus porosity, color-coded by gamma ray. It is possible to infer that from high to low impedance values, the clay content increases (assuming a linear relationship between gamma ray and clay content).

approximation used in the calculation of the reflection coefficient (Fatti et al., 1994). In fact, the normal-incidence component of the seismic reflection data, which directly relates with the AI, is much easier to retrieve than the angle-dependent one. For this reason, the inverted EI models for the synthetic case (Figure 6b) explore considerably more the uncertainty space when compared with the AI models (Figure 6a). This effect can also be visible for the real case study but is harder to distinguish.

When comparing the distribution configuration of the MDS space for the real data example with the one from the synthetic example, it is clear that we are exploring the uncertainty space in a more comprehensive way from iteration to iteration: The AI and EI models now appear more scattered in the MDS plot. This effect is directly related with the noisy nature of field seismic data when compared with the noise-free synthetic example. From iteration to iteration, the geostatistical inversion algorithm struggles to find impedance models that produce synthetic seismic accepted by the objective function. This effect is particularly noticed at the end of iteration 1: The resulting best AI and EI models produce synthetic seismic that have lower local correlation coefficient values, when compared with the synthetic example. Therefore, the simulation ensemble of iteration 1 has impedance models that are considerably different among themselves. This is a natural result of using the secondary variable in the cosimulation associated with lower local correlation coefficients.

Also of interest are the differences in the distances between the best-fit inverse models for all five wells in the real data example. For wells 1 and 2, the Euclidean distance between the best AI/EI model and the real one is always larger than for the rest of the wells. This difference may be directly related with the geologic environment of each well: Despite that all the wells were drilled in the same hydrocarbon field, they are located in different fault blocks — wells 1 and 2 are located in the footwall block, and wells 3, 4, and 5 are located in the hanging wall block (Figure 9). The differences in the geologic setting can, for example, also be addressed in the AI versus porosity domain (Figure 10). Wells 3, 4, and 5 drill three distinct facies in terms of its clay content. Figure 10 shows an increase of clay content from high to low AI values, assuming a linear correlation between the values of gamma ray and the clay content. Well 1 penetrates only a sandy facies, characterized by low gamma ray values and high AI/EI values, and it lacks the shaly facies characterized by lower AI/EI values. Well 2, which is spatially located between well 1 and the rest of the wells (Figure 9), samples the sandy and the shaly facies but misses the transitional facies between the higher and lower AI/EI values. Unfortunately, and due to the lack of the available well-log data, it was not possible to calibrate a rock physics model (e.g., stiff-sand model) to this case study. We believe that a calibrated rock physics model would give new insights to the

approximation used in the calculation of the reflection coefficient (Fatti et al., 1994). In fact, the normal-incidence component of the seismic reflection data, which directly relates with the AI, is much easier to retrieve than the angle-dependent one. For this reason, the inverted EI models for the synthetic case (Figure 6b) explore considerably more the uncertainty space when compared with the AI models (Figure 6a). This effect can also be visible for the real case study but is harder to distinguish.

discussion about the different convergence rates for the five available wells.

CONCLUSIONS

The MDS technique was successfully applied to assess how the plausible parameter model space is being explored by a geostatistical seismic inversion algorithm and how well the retrieved inverse-impedance models are converging toward the real impedance models.

MDS was computed over the entire ensemble of simulated AI and EI models resulting from the GEI applied to a synthetic data set. The results showed that the retrieved best-fit AI and EI models are reliable because they both converged toward the global solution, the real AI and EI models.

To assess the performance of the GEI with real seismic data sets, we applied a cross-validation technique. We ran distinct geostatistical inversions, and on each a different well, from the available five, was removed from the conditioning set of data. We then evaluated the inverted AI and EI models at the well location for the well not used during that specific inversion run. Once again, the results of the MDS plot showed a good convergence of the inverted earth models toward the solution without reducing the exploration of the uncertainty space. However, from the retrieved results, we may say that the geologic context may influence the convergence performance of the geostatistical inversion method.

We believe that the MDS technique may also be used to benchmark the convergence of different seismic inversion algorithms (e.g., deterministic, linearized Bayesian). Using synthetic data sets and plotting in the metric space the best-fit models, retrieved from different inverse methodologies, against the real impedance models we could access how well the convergence of the retrieved elastic models is.

ACKNOWLEDGMENTS

The authors would like to acknowledge CERENA/CMRP for supporting this work, Petrobras for the permission to use this synthetic data set, and Partex Oil & Gas for the permission to publish the results from the real case study. We also acknowledge H. Djikpesse and the two anonymous reviewers for their helpful comments and suggestions.

APPENDIX A

MULTIDIMENSIONAL SCALING

Here, we recall the mathematical expressions for the MDS following [Borg and Groenen \(1997\)](#) and [Caers \(2011\)](#).

Consider an ensemble of earth models, composed by a number L of individual earth models, represented by a vector \mathbf{x}_i of dimension N . Traditionally, $L \ll N$,

$$\mathbf{X} = [\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_L]^T. \quad (\text{A-1})$$

A metric space \mathbf{M} , which by definition is only equipped with distances, can be constructed computing a distance (e.g., a Euclidean distance) between any \mathbf{x}_i earth model constructing the matrix \mathbf{D} , the dissimilarity matrix. Matrix \mathbf{D} may be centered such that the

origin of the map is zero. This transformation may be done in a scalar (equation A-2) or in a matrix (equation A-3) way:

$$b_{ij} = -\frac{1}{2} \left(d_{ij}^2 - \frac{1}{L} \sum_{k=1}^L d_{ik}^2 - \frac{1}{L} \sum_{l=1}^L d_{il}^2 - \frac{1}{L^2} \sum_{k=1}^L d_{ij}^2 \sum_{l=1}^L d_{kl}^2 \right), \quad (\text{A-2})$$

$$a_{ij} = -\frac{1}{2} d_{ij}^2. \quad (\text{A-3})$$

Then, we can center the matrix by

$$\mathbf{B} = \mathbf{H} \mathbf{D} \mathbf{H}, \quad (\text{A-4})$$

where \mathbf{H} can be expressed by

$$\mathbf{H} = \mathbf{I} - \frac{1}{L} \mathbf{1} \mathbf{1}^T, \quad (\text{A-5})$$

where $\mathbf{1}$ is a row of ones with size L and \mathbf{I} is the identity matrix of size L . \mathbf{B} can be rewritten in such a form, defining a $L \times L$ matrix:

$$\mathbf{B} = (\mathbf{H} \mathbf{X})(\mathbf{H} \mathbf{X})^T. \quad (\text{A-6})$$

For most of the cases $L \ll N$, \mathbf{B} is positive definite and can be decomposed by eigenvalue decomposition as

$$\mathbf{B} = v_i \lambda_i v_i^T \quad (\text{A-7})$$

where v_i are the eigenvectors and λ_i are the eigenvalues. Choosing the first d eigenvectors, which correspond to the d higher eigenvalues, \mathbf{X}_d contains the location of the models in the MDS space for the retained dimension:

$$\mathbf{X}_d = [\mathbf{x}_{1,d} \mathbf{x}_{2,d} \dots \mathbf{x}_{L,d}]^T. \quad (\text{A-8})$$

REFERENCES

- Borg, I., and G. Groenen, 1997, *Modern multidimensional scaling: Theory and applications*: Springer.
- Bortoli, L. J., F. Alabert, A. Haas, and A. G. Journel, 1992, Constraining stochastic images to seismic data: *Geostatistics Tróia*, **1**, 325–337, doi: [10.1007/978-94-011-1739-5_27](https://doi.org/10.1007/978-94-011-1739-5_27).
- Bosch, M., T. Mukerji, and E. F. Gonzalez, 2010, Seismic inversion for reservoir properties combining statistical rock physics and geostatistics: A review: *Geophysics*, **75**, no. 5, A165–A176, doi: [10.1190/1.3478209](https://doi.org/10.1190/1.3478209).
- Boschetti, F., M. C. Dentith, and R. D. List, 1996, Inversion of seismic refraction data using genetic algorithms: *Geophysics*, **61**, 1715–1727, doi: [10.1190/1.1444089](https://doi.org/10.1190/1.1444089).
- Buland, A., and H. More, 2003, Bayesian linearized AVO inversion: *Geophysics*, **68**, 185–198, doi: [10.1190/1.1543206](https://doi.org/10.1190/1.1543206).
- Caers, J., 2011, *Modeling uncertainty in earth sciences*: Wiley-Blackwell.
- Cox, T. F., and M. A. A. Cox, 1994, *Multidimensional scaling*: Chapman & Hall.
- Deutsch, C., and A. G. Journel, 1998, *GSLIB: Geostatistical software library and users' guide*: Oxford University Press.
- Doyen, P. M., 2007, *Seismic reservoir characterization: EAGE*.
- Fatti, J. L., G. C. Smith, P. J. Vail, P. J. Strauss, and P. R. Levitt, 1994, Detection of gas in sandstone reservoirs using AVO analysis: A 3-D seismic case history using the geostack technique: *Geophysics*, **59**, 1362–1376, doi: [10.1190/1.1443695](https://doi.org/10.1190/1.1443695).
- González, E. F., T. Mukerji, and G. Mavko, 2008, Seismic inversion combining rock physics and multiple-point geostatistics: *Geophysics*, **73**, no. 1, R11–R21, doi: [10.1190/1.2803748](https://doi.org/10.1190/1.2803748).

- Grana, D., and E. Della Rossa, 2010, Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion: *Geophysics*, **75**, no. 3, O21–O37, doi: [10.1190/1.3386676](https://doi.org/10.1190/1.3386676).
- Grana, D., T. Mukerji, J. Dvorkin, and G. Mavko, 2012, Stochastic inversion of facies from seismic data based on sequential simulations and probability perturbation method: *Geophysics*, **77**, no. 4, M53–M72, doi: [10.1190/geo2011-0417.1](https://doi.org/10.1190/geo2011-0417.1).
- Haas, A., and O. Dubrule, 1994, Geostatistical inversion — A sequential method of stochastic reservoir modeling constrained by seismic data: *First Break*, **12**, 561–569.
- Horta, A., and A. Soares, 2010, Direct sequential co-simulation with joint probability distributions: *Mathematical Geosciences*, **42**, 269–292, doi: [10.1007/s11004-010-9265-x](https://doi.org/10.1007/s11004-010-9265-x).
- Loertzer, G. J., and A. J. Berkhout, 1992, An integrated approach to lithologic inversion: Part I, Theory: *Geophysics*, **57**, 233–244, doi: [10.1190/1.1443236](https://doi.org/10.1190/1.1443236).
- Ma, Xin-Quan, 2002, Simultaneous inversion of prestack seismic data for rock properties using simulated annealing: *Geophysics*, **67**, 1877–1885, doi: [10.1190/1.1527087](https://doi.org/10.1190/1.1527087).
- Mallick, S., 1995, Model based inversion of amplitude variations with offset data using a genetic algorithm: *Geophysics*, **60**, 939–954, doi: [10.1190/1.1443860](https://doi.org/10.1190/1.1443860).
- Mallick, S., 1999, Some practical aspects of prestack waveform inversion using a genetic algorithm: An example from the East Texas woodbine gas sand: *Geophysics*, **64**, 326–336, doi: [10.1190/1.1444538](https://doi.org/10.1190/1.1444538).
- Nunes, R., A. Soares, G. Schwedersky, L. Dillon, L. Guerreiro, H. Caetano, C. Maciel, and F. Leon, 2012, Geostatistical inversion of prestack seismic data: Presented at Ninth International Geostatistics Congress.
- Russel, B. H., 1988, Introduction to seismic inversion methods: SEG.
- Scales, J. A., and L. Tenorio, 2001, Prior information and uncertainty in inverse problems: *Geophysics*, **66**, 389–397, doi: [10.1190/1.1444930](https://doi.org/10.1190/1.1444930).
- Scheidt, C., and J. Caers, 2009, Representing spatial uncertainty using distances and kernels: *Mathematical Geosciences*, **41**, 397–419, doi: [10.1007/s11004-008-9186-0](https://doi.org/10.1007/s11004-008-9186-0).
- Sen, M. K., and P. L. Stoffa, 1991, Nonlinear one dimensional seismic waveform inversion using simulated annealing: *Geophysics*, **56**, 1624–1638, doi: [10.1190/1.1442973](https://doi.org/10.1190/1.1442973).
- Soares, A., 2001, Direct sequential simulation and cosimulation: *Mathematical Geology*, **33**, 911–926, doi: [10.1023/A:1012246006212](https://doi.org/10.1023/A:1012246006212).
- Soares, A., J. D. Diet, and L. Guerreiro, 2007, Stochastic inversion with a global perturbation method: EAGE Petroleum Geostatistics, 10–14.
- Suzuki, S., and J. Caers, 2008, A distance-based prior model parameterization for constraining solutions of spatial inverse problems: *Mathematical Geosciences*, **40**, 445–469, doi: [10.1007/s11004-008-9154-8](https://doi.org/10.1007/s11004-008-9154-8).
- Tarantola, A., 2005, Inverse problem theory: SIAM.
- Tompkins, M. J., J. L. Fernández Martínez, D. L. Alumbaugh, and T. Mukerji, 2011, Scalable uncertainty estimation for nonlinear inverse problems using parameter reduction, constraint mapping, and geometric sampling: Marine controlled-source: *Geophysics*, **76**, no. 4, F263–F281, doi: [10.1190/1.3581355](https://doi.org/10.1190/1.3581355).