Exploration of data space through trans-dimensional sampling: A case study of 4D seismics

Nicola Piana Agostinetti[1], Maria Kotsi[2], and Alison Malcolm[3]

[1]University of Milano-Bicocca
[2]PanGeo Subsea Inc
[3]Memorial University of Newfoundland

November 28, 2022

Abstract

We present a novel methodology for exploring 4D seismic data in the context of monitoring subsurface resources. Data-space exploration is a key activity in scientific research, but it has long been overlooked in favour of model-space investigations. Our methodology performs a data-space exploration that aims to define structures in the covariance matrix of the observational errors. It is based on Bayesian inferences, where the posterior probability distribution is reconstructed through trans-dimensional (trans-D) Markov chain Monte Carlo sampling. The trans-D approach applied to data-structures (termed "partitions') of the covariance matrix allows the number of partitions to freely vary in a fixed range during the McMC sampling. Due to the trans-D approach, our methodology retrieves data-structures that are fully data-driven and not imposed by the user.
We applied our methodology to 4D seismic data, generally used to extract information about the variations in the subsurface. In our study, we make use of real data that we collected in the laboratory, which allows us to simulate different acquisition geometries and different reservoir conditions. Our approach is able to define and discriminate different sources of noise in 4D seismic data, enabling a data-driven evaluation of the quality (so-called "repeatability') of the 4D seismic survey. We find that: (1) trans-D sampling can be effective in defining data-driven data-space structures; (2) our methodology can be used to discriminate between different families of data-structures created from different noise sources. Coupling our methodology to standard model-space investigations, we can validate physical hypothesis on the monitored geo-resources.

Exploration of data space through trans-dimensional sampling: A case study of 4D seismics

Nicola Piana Agostinetti

ZED Depth Exploration Data GmbH, Vienna, Austria

Maria Kotsi

PanGeo Subsea Inc

Earth Sciences Department, Memorial University of Newfoundland, St John's, NL, Canada

Alison Malcolm

Earth Sciences Department, Memorial University of Newfoundland, St John's, NL, Canada

---

ZED Depth Exploration Data GmbH, Vienna, Austria

PanGeo Subsea Inc

Earth Sciences Department, Memorial University of Newfoundland, St John's, NL, Canada

Abstract.

We present a novel methodology for exploring 4D seismic data in the context of monitoring subsurface resources. Data-space exploration is a key activity in scientific research, but it has long been overlooked in favour of model-space investigations. Our methodology performs a data-space exploration that aims to define structures in the covariance matrix of the observational errors. It is based on Bayesian inferences, where the posterior probability distribution is reconstructed through trans-dimensional (trans-D) Markov chain Monte Carlo sampling. The trans-D approach applied to data-structures (termed ”partitions”) of the covariance matrix allows the number of partitions to freely vary in a fixed range during the McMC sampling. Due to the trans-D approach, our methodology retrieves data-structures that are fully data-driven and not imposed by the user.

We applied our methodology to 4D seismic data, generally used to extract information about the variations in the subsurface. In our study, we make use of real data that we collected in the laboratory, which allows us to simulate different acquisition geometries and different reservoir conditions. Our approach is able to define and discriminate different sources of noise in 4D seismic data, enabling a data-driven evaluation of the quality (so-called “repeatability”) of the 4D seismic survey. We find that: (1) trans-D sampling can be effective in defining data-driven data-space structures; (2) our methodology can be used to discriminate between different families of data-structures created from different noise

sources. Coupling our methodology to standard model-space investigations, we can validate physical hypothesis on the monitored geo-resources.

1. Introduction

In their investigations of the Earth system, geo-scientists have to deal with two complementary spaces: data space and model space. The *model space* is generally defined as the space of the investigated parameters. For a given parameterization of the system, each point of the model space defines a possible model of the system, represented by a combination of values of the model parameters. To make inferences on the model parameters, we need to take measurements of relevant geo-observables. The *data space* contains all the possible combinations of such observations [*Tarantola*, 2005] and the measured data points form a local subset of the data space with its own structure. While there is a vast literature about methodologies for investigating the model space [*Sambridge and Mosegaard*, 2002, e.g.], few attempts have been made at a systematic exploration of the data space. Exploration of the data space is an ordinary activity for geo-scientists, and includes, for example, data preparation, quality controls (QC)s for data selection and estimation of data errors. Some of those activities, for example the data selection, could have a strong impact on the data space, modifying, for example, the data structure. Generally, such activities rely on the *expert-opinion* of the geoscientists and are carried out ahead of the main geophysical investigations that are related to the model space.

There are two main reasons for considering a systematic exploration of the data space. First, the ever growing amount of geo-data available to geo-scientists needs to be tackled with more automated workflows; expert opinion is generally a time-consuming process. Second, more interestingly, expert opinion, as a human activity, implies the separation of data into categories (i.e. a discrete number of outputs) rather than a more general continuous evaluation of probability. For example, in data selection activities, the expert can select and, then exclude, part of the data based on their experience, using a two category model (in/out, good/bad). Conversely, a more automated workflow, developed in a statistical framework, can associate a probability value to each data point, avoiding the need to remove any of them from the analysis.

In recent years, some studies have reported cases of systematic exploration of the data space, even if such analyses take often a marginal role in the scientific studies themselves. In particular, there are some examples [*Bodin et al.*, 2012a; *Dettmer and Dosso*, 2012; *Xiang et al.*, 2018] where Bayesian inference is applied to a geophysical inverse problem for defining both physical parameters (i.e. investigating the model space) and the errors associated to the data (i.e. exploring the data space), the so called *Hierarchical Bayes* approach [*Malinverno and Briggs*, 2004] . In Hierarchical Bayes algorithms, the uncertainties related to the data are assumed to be poorly known and need to be estimated during the process. This approach usually assumes a fixed number of parameters which represent the unknown part of the data space. In most applications of the Hierarchical Bayes approach, the absolute value of the data errors is considered an unknown in the problem that needs to be inferred [*Bodin et al.*, 2012a]. Sometimes, in cases where the structure of the data errors is known (i.e. we know which data points are measured with more precision with respect to other points), a scaling factor of the data error is used as the unknown [*Piana Agostinetti and Malinverno*, 2018]. In more complex cases, the Hierarchical Bayes approach is adopted to somehow define a function of the data uncertainties, so called “data structures” or “states” hereinafter, which include: estimating an auto regressive model of the data errors [i.e. a form of error correlation, *Dettmer and Dosso*, 2012], and estimating an increasing linear model for the data errors as a function of the geometrical distance between measurement points [e.g. *Galetti et al.*, 2016]. In all of these cases however, the number of parameters representing the data structure is fixed a-priori (usually one or two parameters, rarely more than three). By contrast, *Steininger et al.* [2013] and *Xiang et al.* [2018], extend Hierarchical Bayes approach to make inferences on the data space by considering data structures that are represented by a variable number of parameters. *Xiang et al.* [2018] make use of a transdimensional (trans-D) sampler [*Sambridge et al.*, 2006, 2013] for sampling models belonging to two different states: in one state, one unknown defines an autoregressive model of the first order for the data errors, i.e. assume uncorrelated errors, while in a second state, two unknowns are used to define an autoregressive model of the second order, i.e. exponential correlation between data uncertainties. Using this ability to jump from one state to the other, the algorithm is able to indicate the "predominant" auto-regressive model associated to the data errors. As far as we know, *Steininger et al.* [2013] and *Xiang et al.* [2018] are the first applications of a trans-D algorithm in Geophysics, for sampling different states representing different error models, even if they are limited to a transition between states represented by one and two parameters.

In this study, we move a step forward in the development of algorithms for data space exploration. We make use of a trans-D sampler for exploring different "states" (represented by a different number of variables), where each state reproduces a partition of the data space (i.e. a data structure). The number of states to be explored is no longer strictly limited [e.g. two states, like in *Xiang et al.*, 2018], and the number of variables representing each state can vary between a user-defined minimum and maximum. The algorithm is developed in a Bayesian framework, used to define the posterior probability of the data structures. Data space structures are expressed in terms of partitions of the covariance matrix of the errors, which allow us to define regions of the data space where measured data are in agreement with a given working hypothesis. The algorithm is applied to the data analysis workflow used for time-lapse seismics (also called *4D seismics*), a technology used primarily by oil&gas companies for monitoring their reservoirs. The 4D seismic data consist of time-repeated active seismic surveys that need to be investigated for detecting noise/distortions and focusing the subsequent geophysical inversion on the portion of active seismic data where temporal changes have occurred. The algorithm is applied on laboratory data that mimic active seismic surveys and the results are discussed in light of the potential of the algorithm for statistically separating signals with different origins.

1.1. 4D seismics: key-concepts and present-day challenges

The term *4D seismics* indicates the data workflow adopted by oil&gas companies for monitoring their reservoirs through the repetition, after a few years, of active seismic surveys. The 4D seismic workflow consists of three main phases: acquisition, processing and interpretation. 4D seismics is generally performed for off-shore reservoirs, but the first successes were obtained on-shore [e.g. *Porter-Hirsche and Hirsche*, 1998; *Davis et al.*, 2003]. This technology is also used for monitoring CO_2 underground storage sites [*Lumley*, 2010; *Cheng et al.*, 2010; *Yang et al.*, 2014; *Roach et al.*, 2015]. Briefly, a first active seismic survey, the so-called *baseline survey*, is performed just before starting production to image the untouched resources. After some time and while the reservoir is under production, the active seismic survey is repeated, the so-called *monitor survey*. If the seismic acquisition and data processing are exactly the same as those used for the baseline survey, the differences between the images can be uniquely attribute to changes in the physical properties of the reservoir due to its exploitation. Through the analysis of such differences, scientists can make informed decisions about the next phases of exploitation of the reservoir.

An important question is: how can we get relevant information from 4D seismics? Production related effects on images obtained from the monitor survey can be obscured by distortions induced by the lack of repeatability of the data acquisition and processing. This is one of the main technical barriers for getting the correct information from 4D seismics [*Koster et al.*, 2000]. The concept of *repeatability* between two or more seismic surveys indicates the degree to which the data-sets can be considered to be generated from the same operational and computational workflows. Measures of repeatability between two seismic surveys generally include Normalized Root Mean Square (NRMS) and trace correlation [also called *predictability Kragh and Christie*, 2002]. Increasing and evaluating the repeatability of 4D seismics have been the focus of a number of studies in the last decades [*Landro*, 1999; *Houck*, 2007; *Pevzner et al.*, 2011], with the main efforts going into increasing acquisition quality, i.e. hardware solutions. Statistical approaches to 4D data analysis have been limited to the interpretation phase [e.g. applying Machine Learning algorithms to porosity inversion *Dramsch*, 2019].

1.2. Methodological framework: Bayesian inference, Markov chain Monte Carlo and trans-dimensional algorithms

Various geophysical inverse problems have been solved following a probabilistic Bayesian framework [*Tarantola*, 2005, 2006]. Bayes' theorem

$$p(\mathbf{m} \mid \mathbf{d}) = \frac{p(\mathbf{m})p(\mathbf{d} \mid \mathbf{m})}{p(\mathbf{d})} \tag{1}$$

connects (probabilistic) prior information $p(\mathbf{m})$ about some subsurface properties (m) and data measured (d), generally at the surface, to extract new information about such properties (the so-called *posterior probability distribution* $p(\mathbf{m} \mid \mathbf{d})$ or PPD), through an (assumed) known error statistics [the Likelihood $p(\mathbf{d} \mid \mathbf{m})$, or $L(\mathbf{m})$ hereinafter, *Bayes*, 1763]. Thus, in contrast with other approaches, the solution of geophysical inverse problems is given in the form of a probability distribution over the investigated parameters, and not as a single value for each parameter (i.e. a single model). In simple cases, Bayes' theorem can give an analytic solution to geophysical inverse problems [*Tarantola*, 1987]. However, numerical methods have been widely used in more complex cases. In particular, Markov chain Monte Carlo (McMC) sampling has been found to be well suited for sampling a chain of Earth models with a probability proportional to the PPD and, thus, to make inferences on relevant parameters based on such sampled models [*Sambridge and Mosegaard*, 2002]. Here, we follow the approach presented in *Mosegaard and Tarantola* [1995] and we develop a sampler of the prior probability distribution which can be "switched" to sample models with a probability that follows the PPD. After collecting a relevant number of models from the PPD, we compute numerical estimators of the investigated parameters directly from the sampled models. For example, the mean value of the parameter m, can be estimated as

$$\hat{m} = \frac{1}{N_s} \sum_{j}^{N_s} m^j, \tag{2}$$

where N_s is the number of samples computed during the McMC sampling and m^j is the value of parameter m for the j-th model sampled. Following the approach in *Mosegaard and Tarantola* [1995], we define the probability of accepting a new model along the Markov chain as:

$$\alpha = \min[1, L(\mathbf{m}_{cand})/L(\mathbf{m}_{cur})], \tag{3}$$

where $\mathbf{m}_{cand}$, the candidate model, and $\mathbf{m}_{cur}$, the current model, are two consecutive Earth models along the Markov chain and $L(\mathbf{m})$ is the likelihood of the model given the observed data. In other words, the candidate is always accepted if $L(\mathbf{m}_{cand}) \geq L(\mathbf{m}_{cur})$. If $L(\mathbf{m}_{cand}) < L(\mathbf{m}_{cur})$, the random walk moves to the candidate model with probability equal to $L(\mathbf{m}_{cand})/L(\mathbf{m}_{cur})$. The last point, $L(\mathbf{m}_{cand}) < L(\mathbf{m}_{cur})$, guarantees that the McMC sampler will not get stuck in a local maximum of the likelihood function, because models which worsen the fit to the data may still be accepted.

Two fundamental points in Bayesian inferences are the initial states of knowledge about the investigated parameters, the so-called *priors*, which can take a closed analytical form, or be represented by a set of rules (e.g. one parameter has to be smaller than a second parameter, like in P- and S- waves velocities in rocks). More interestingly, the statistics of the data uncertainties should be known at a certain level. Such statistics is used to compute the likelihood value of an Earth model. Simplified statistics can be adopted (e.g. a diagonal covariance matrix in Gaussian distributed errors) but has been proven to give un-realistic results in some cases [*Birnie et al.*, 2020]. Both of these assumptions have to hold to make inferences on physical parameters and, given Equation 1, the solution to the geophysical inverse problem may change under different assumptions.

An efficient design of the McMC sampler is fundamental for achieving robust results (in terms of number of samples extracted from the PPD) in a limited amount of time. Several different *recipes* have been designed in the past for proposing a *candidate model*, i.e. a new point in the model space, as a perturbation of the *current model*, i.e the last visited point in the model space [*Bodin et al.*, 2012b]. In fact, if the sampling is too limited to the neighbourhood of the current model, McMC will converge too slowly toward the global maximum of the likelihood function. Conversely, too strong a perturbation of the current model will likely lead to poorly fitting candidate models, most of which will be rejected. In recent years, one ingredient that has been added to many implementations of the McMC sampler is the possibility of sampling a candidate model which has a different number of variables than the current model [*Malinverno*, 2002; *Sambridge et al.*, 2006]. In practise, we relax the hard constraint of a fixed number of variables in the models, allowing it to vary between fixed minimum and maximum values. This new generation of McMC samplers are collectively called trans-dimensional samplers [e.g. *Sambridge et al.*, 2013] and are based on the pioneering works of *Geyer and Møller* [1994] and *Green* [1995]. For trans-dimensional samplers, Equation 3 holds under specific assumptions on the model space transformation and its Jacobian matrix [see Appendix B in *Piana Agostinetti and Malinverno*, 2010, for details].

2. Data

We consider a simple time-lapse scenario that consists of an overburden layer and a reservoir. To better mimic a real world application, we use a scaling factor of 10000 such that a frequency of 200 kHz represents a frequency of 20 Hz, and a dimension of 1 mm represents 10 m. To build this experiment in the lab we take two Plexiglas blocks with dimensions $310 \times 154 \times 77$ mm, and attach them together (Figure 1). The first Plexiglas block represents the overburden layer with elastic properties of $V_p = 2780$ m/s, $V_s = 1480$ m/s, and $\rho = 1.19$ g/cm^3. This overburden layer remains unchanged between the two surveys. To build the reservoir layer we remove a rectangular cube from the second block, allowing us to insert different fluids into our 'reservoir'.

For the baseline survey, we keep the second block empty, representing a gas-filled reservoir. In this case, the elastic properties of the air are $V_p = 332$ m/s, $V_s =$ N/A, and $\rho \sim 0$ g/cm^3. For the monitor survey, we fill the block with water, miming a scenario where the gas in the reservoir has been replaced with brine. The elastic properties of the water are $V_p = 1500$ m/s, $V_s =$ N/A, $\rho \sim 1$ g/cm^3. Figure 1 shows the experimental setup for the data acquisition. For the source we use a P-wave transducer with a single-cycle sine wavelet at 200 kHz, generated through the function generator (top left corner of Figure 1). This P-wave transducer has a diameter of

10 mm. For the receivers, we use a laser vibrometer that measures the particle velocity along the direction of the laser beam (perpendicular to the surface), and sends it to the oscilloscope to be saved. The laser measures the signal at 160 points along the tape, giving us a total of 160 receivers with a sampling distance of 0.5 mm. The nearest offset in this case is 10 mm. Figure 1 top right corner shows the signal reading at the nearest offset for the baseline case. Throughout the data acquisition the P-wave transducer is glued to the Plexiglas box, and the laser is attached to a stage that stably moves it along the tape. This allows for a controlled and repeatable time-lapse experiment. Summarising, the experimental set-up allows us to record 160 "wiggles" for each of the two different reservoir-states, composing two "shot-gathers". For the first 100 wiggles in each shot-gather, clear arrivals from the surface and the reservoir can be separated. These shot-gathers compose a homogeneous, discrete (x, t)-space, where x is the wiggle offset, and t is the recording time (Figure 2). In general, we use the first shot-gather from the first reservoir-state experiment as the "baseline survey" (Figure 2a). We combine the wiggles for the two experiments to simulate different monitoring scenarios. For example, in Figure 2b, we mimic: (a) the misplacement of some sensors (wiggles between 15 and 25), replacing the correct baseline wiggles with wiggles from the baseline survey but with a four-wiggles shift; and (b) the presence of changes in the reservoir (wiggles 60 to 90), replacing wiggles from the baseline with wiggles from the second reservoir-state experiment. Point-wise measurements of the squared difference between baseline and monitor surveys can be larger for misplacement sensors than for reservoir alteration (Figure 2c), making the discrimination between the two effects quite challenging.

To test our methodology, we used one in five wiggles for the first 100 wiggles, thus, we collect 20 "traces" for each survey, $N_w = 20$. Downsampling the number of wiggles allows us to have enough data for simulating the misplacement of the receiver in the monitor survey. In the following , we continue to call “wiggles” the recording for a single detector position as a function of time in each shot-gather, and we call “traces” the wiggles selected to compose the baseline and monitor surveys. Each trace is composed of $N_S = 1251$ samples. Thus, our (x, t)-space is composed of $N_w \cdot N_s = 25020$ data-points.

2.1. Error statistics

To rigorously compare the monitor and baseline survey we need to know how the errors are statistically distributed in the two data-sets, i.e. the error covariance matrix. Computing the rank of such a large $(N_w \cdot N_s) \times (N_w \cdot N_s)$ matrix could be intractable. To avoid this, we estimate the covariance matrix from the data themselves with the following assumptions. First, we do not consider inter-trace correlation, so our model of the covariance matrix is block-diagonal, one block for each trace. Note that this assumption means that near-by traces are not correlated, which could be un-realistic under some scenarios, e.g. weather conditions, acquisition systems and so on. Second, we assume the same error statistics for the baseline and monitor surveys. Again, this assumption could be partially false for, e.g., surveys acquired with a large (10s of years) time-gap. However, under our assumptions, we can estimate a tractable error covariance matrix $\mathbf{C}^*_{\mathbf{e,ij}}$ which can be decomposed following the approach developed in *Malinverno and Briggs* [2004], with an adequate correlation model [*Kolb and Lekić*, 2014].

Given the nature of our data, i.e. band-limited waveforms, our covariance matrices are semi-positive definite Toeplitz matrices and they can be decomposed as:

$$\mathbf{C}^*_{\mathbf{e,ij}} = \mathbf{SRS} \tag{4}$$

where:

$$\mathbf{S} = \begin{pmatrix} \sigma_{1,1} & 0 & 0 & \dots & 0 \\ 0 & \sigma_{2,1} & 0 & \dots & 0 \\ 0 & 0 & \sigma_{3,1} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \sigma_{N_s,N_w} \end{pmatrix} \tag{5}$$

represents the diagonal matrix containing the standard deviation of each data point b_{ij} in the baseline [*Malinverno and Briggs*, 2004].

With the assumption of independent traces, the correlation matrix $\mathbf{R}$ can be represented as a block-diagonal matrix with N_w blocks, each of dimension: $N_s \times N_s$. The block $\mathbf{R}_j$ represents the error correlation within the j-th trace and can be estimated from the data [*Piana Agostinetti and Malinverno*, 2018; *Piana Agostinetti and Martini*, 2019]. However, such data-derived correlation matrices $\mathbf{R}_j$ are often not positive definite and need to be approximated, e.g., with the singular value decomposition, to use them for estimating the covariance matrix and computing the likelihood $L(\mathbf{m}_{cand})$. In this study, we make use of a correlation model that results in positive definite matrices and guarantees stable matrix inversion [*Kolb and Lekić*, 2014]. Thus, our blocks $\mathbf{R}_j$ assume the form:

$$\mathbf{R}_j = R_{ik,j} = e^{-\lambda_j |t_i - t_k|} \cos(\lambda_j \omega_j |t_i - t_k|) \tag{6}$$

where t_k and t_i are the time of the b_{kj} and b_{ij} samples, respectively, while λ_j and ω_j are estimated from the data in the j-th trace. In Figure 3, we illustrate the computation of σ_{ij}, λ_j and ω_j. In Figure 3a, we show how we estimate the standard deviation of each point in each trace. For the j-th trace (red), we consider all traces between $j-5$ and $j+5$ and we compute a stack of these traces (Figure 3b). From the stack, we compute a residual for each trace considered (Figure 3c) and the residuals are autocorrelated. The autocorrelation functions are stacked to obtain an average autocorrelation (orange line in Figure 3d). This function is used to estimate λ_j and ω_j (green line in Figure 3d), through a 2-parameter grid search. Our model for the autocorrelation function fits the empirical function well before $10\mu s$ and somewhat over-estimates sample correlation at longer periods, thus it should be considered a conservative model.

3. Exploration of the data-space through trans-dimensional sampling: methodology

Exploring the data space of 4D seismics implies the separation of multiple sources for the "4D signal" (i.e. the signal arising when monitor and baseline surveys differ). Here we consider a simplified case using three signal sources: ambient random noise (*noise*, hereinafter), sensor misplacement (*perturbation*) and physical changes in the reservoir (*target signal*). With perfect survey repetition (no sensor misplacement) and no change in the reservoir, the unique source of 4D signal is the noise. Assuming an empirically estimated noise model, we can define our working hypothesis: in the case of a unique source of 4D signal from the noise, the fit of the monitor survey with respect to the baseline survey should close to the number of data-points $N_w \times N_s$, where the fit is statistically represented by:

$$\phi^* = \left(\mathbf{e_{ij}^T} \, (2 \times \mathbf{C^*_{e,ij}})^{-1} \, \mathbf{e_{ij}} \right), \tag{7}$$

which is used to compute the likelihood of the monitoring to the baseline survey:

$$L^* = \prod_{i=1}^{N_w} \frac{1}{[(2\pi)^{N_s} |2 \times \mathbf{C^*_{e,ij}}|]^{1/2}} \exp\left(-\frac{1}{2}\phi \right), \tag{8}$$

and we assume Gaussian distributed noise with the error model defined in Section 2.1. Here, the covariance matrix $\mathbf{C^*_{e,ij}}$ is directly estimated from the data through their autocorrelation and their standard deviation. It is interesting to note that the likelihood computation is what we need to advance our McMC sampling, following Equation 3.

When there signals in the 4D data caused by different sources, we can adopt a Hierarchical Bayes approach to define a different configuration for the covariance matrix so that the new covariance matrix will again closely fit our error model and the working hypothesis defined by Equation 8. As detailed in *Bodin et al.* [2012a], modifications to the covariance matrix obtained through a Hierarchical Bayes algorithm not only represent improved estimates of the data uncertainties, but also include any additional source of uncertainty arising from, e.g., un-realistic modelling or, as in our case, incorrect assumptions. In fact, the likelihood function above does represent the differences in the two surveys in case of noise only (our assumption), and the covariance matrix needs to be modified appropriately when this hypothesis is violated. In the case of sensor mis-placement (i.e. when errors occur in the geometry of the monitor survey), the modification of the covariance matrix should be the same for all the points belonging to the misplaced traces. Conversely, when changes in the reservoir occur, the covariance matrix needs to be modified only for those seismic phases generated at the top of the reservoir for some consecutive traces (in our simplified data, from the top and the bottom in field measurements). Summarising, we will try to define a different structure for the covariance matrix so that the modified covariance matrix will approximate our error model.

3.1. Partition of the error covariance matrix

Here we define a new structure of the covariance matrix as an unambiguous correspondence between a partition of the data and a partition of the covariance matrix, so that separating regions of the data space separates distortions in the covariance. Given the properties of the covariance matrix and assigning a relevant weight to each sampled point (x,t), we can create a modified covariance matrix such as

$$\mathbf{C_{e,ij}}(\mathbf{m}) = \mathbf{W}(\mathbf{m}) \times 2 \times \mathbf{C^{*}_{e,ij}} \times \mathbf{W}(\mathbf{m}) \tag{9}$$

where

$$\mathbf{W_{ij}}(\mathbf{m}) = 10^{w_{ij}(\mathbf{m})}, \tag{10}$$

and w_{ij} is a weight associated to sample point (x,t), derived by the model sampled during the McMC process. Note that our assumptions on the original covariance matrix (block-diagonal matrix generated from a modelled correlation function) are not necessary for generating $\mathbf{C_{e,ij}}$. Thus, the following discussion can be generalized to any covariance matrix. The goal now is to generate sensitive weights for all points, to be able to separate the portion of the monitor survey where the signal follows the likelihood in Equation 8, from the signal where other distortions are present. Given the nature of the distortions considered here, we can assume that, in the case of the misplacement of a single sensor, all the weights associated to the corresponding trace have to be modified by the same amount. This means that, for a given j, the weights w_{ij} would be the same for one entire block along the diagonal of the covariance matrix, associated to the misplaced trace. Conversely, in case of a change in the reservoir, all weights associated to the same seismic phase need to be homogeneously modified. Thus, w_{ij} would be the same for the same time interval across different traces (assuming an almost flat interface generating phases arriving almost at the same time at the receivers, as in Figure 2a at about 70μs). This second kind of distortion strongly impacts the covariance matrix, equivalently modifying many blocks along its diagonal. Having homogeneous weights for different portions of the covariance matrix, we can create a partition of the covariance matrix based on the corresponding partition of the (x, y)-space associated to the relevant distortion. Giving the nature of our algorithm, i.e. a new way for elaborating partitions of the data, it could be categorized as a member of the family of clustering algorithms, where the number of cluster is not pre-specified by the user or chosen during or after the data analysis, but it is self-defined by the data themselves [e.g. *Mechelen et al.*, 2018].

3.1.1. Model parameterization

We model our partition of the covariance matrix as rectangular partitions of the data-space (Figure 4). Our model is represented by a variable number of rectangular patches (so-called *cells*) that cover the data-space, where each patch has an associated constant weight. In detail, our model $\mathbf{m}$ is composed of a scalar n and five n-vectors, $\mathbf{m} = (n, \mathbf{c_n}, \mathbf{r_n}, \mathbf{t_n}, \mathbf{s_n}, \pi_n)$, where n is the number of cells, $\mathbf{c_n}$ the vector of position of cell centres along the x-axis, $\mathbf{r_n}$ the vector of cell radii along the x-axis, $\mathbf{t_n}$ the vector of the time-position of the cell centres along the time axis, $\mathbf{s_n}$ the vector of the time-width of the cells, and π_n the vector of the cell weights. Keeping the model definition in mind, we can assume that the relevant weight for each point in the data space is the sum of the weights of the cells that extend to cover that particular point:

$$w_{ij}(\mathbf{m}) = 0 \qquad \text{if} \qquad x_{ij} \notin C_m \forall m = 1, ..., n \tag{11}$$

$$w_{ij}(\mathbf{m}) = \sum_{m=1}^{n} \pi_m \qquad \text{if} \qquad x_{ij} \in C_m \tag{12}$$

where C_m represents the time-space extension of the cell associated to the m-th nucleus, i.e.:

$$x_{ij} \in C_m \Leftrightarrow \begin{cases} c_m - 1/2 \cdot r_m < x_i < c_m + 1/2 \cdot r_m, \\ t_m - 1/2 \cdot s_m < x_j < t_m + 1/2 \cdot s_m \end{cases} \tag{13}$$

Having defined the weight for each data point as a function of the partitioning model of the data space, we now have most of the elements for sampling the model space according to our McMC strategy. In fact, the weights define the likelihood of the model from Equation 8 substituting $\mathbf{C_{e,ij}}$ for $\mathbf{C}^*_{\mathbf{e,ij}}$, i.e.:

$$L(\mathbf{m}) = p(\mathbf{d} \mid \mathbf{m}) = \prod_{i=1}^{N_w} \frac{1}{[(2\pi)^{N_s}|\mathbf{C_{e,ij}}|]^{1/2}} \exp\left(-\frac{1}{2}\phi\right), \tag{14}$$

where:

$$\phi = \left(\mathbf{e}_{\mathbf{ij}}^{\mathbf{T}} \mathbf{C_{e,ij}}^{-\mathbf{1}} \mathbf{e_{ij}}\right). \tag{15}$$

The novelty of our approach resides in the fact that, differently from standard McMC schemes, here the dependence of the likelihood function on the model is solely expressed in the covariance matrix and not in the residuals **e** [e.g. *Malinverno*, 2002].

Our choice of rectangular cells is optimal for the case of vertical and horizontal anomalies, because the trans-D sampler can easily mimic this kind of distortions with a limited number of cells. However, all models sampled from the PPD will have vertical and horizontal boundaries, thus generating a somewhat “blocky” PPD. For more complex, i.e. dipping, anomalies, more general functions such as “anisotropic Gaussian kernels [*Belhadj et al.*, 2018] can be adopted.

3.2. Priors

To make Bayesian inferences about the data partitions we define appropriate prior probability distributions on the model parameters. We make use of uniform probability distributions between minimum and maximum values for all investigated parameters. Minimum and maximum values are reported in Table 2. Uniform priors have several advantages from a computational point of view, and keep the number of pieces of prior information to a minimum (two values per parameter). We do not impose any constraints on the radius and time-window parameters for cell centres approaching the boundary of the (x,t) space, i.e. some cells could span outside the (x,t) space (this is the reason why some cells seem to have their centres not exactly in the middle of the cells in Figure 4). While this assumption can introduce some combinations of parameters with very limited impact on the likelihood function (e.g. when c_m is close to one or close to N_W and r_m is small), the /it parsimonious behaviour of our trans-D approach guarantees that useless cells are removed from the model at some point, thus avoiding keeping too many cells.

3.3. Candidate selection

We now need to define how to progress in our McMC sampling, i.e. how to propose a new candidate model to be compared to the current one, the so called *recipe*. Defining an efficient recipe, in terms of convergence to the global maximum of the likelihood function and ability to explore a (potentially) multi-modal distribution, is fundamental for keeping the required computational resources reasonable.

Our recipe comprises seven moves, each of which represents a different way of perturbing the current model. During the definition of the candidate model only one of the moves is performed. Moves are selected with different probability. In detail, we define the following moves:

1. perturb the time-position t_n of a randomly picked cell nucleus (this move has a probability of 0.15 to be selected);
2. perturb the space-position c_n of a randomly picked cell nucleus (0.15)
3. perturb the time-extension s_n of a randomly picked cell nucleus (0.15);
4. perturb the space-extension r_n of a randomly picked cell nucleus (0.15);
5. perturb the weight π_n of a randomly picked cell (0.2);
6. birth of a new cell: one cell is ad dded to the model (0.1);
7. death of a cell: one cell is removed from the model (0.1).

Perturbation of the parameters in moves [1]-[5] are made according to the scheme in Appendix A in *Piana Agostinetti and Malinverno* [2010]. Following this scheme, the normal proposal distributions for sampling the uniform priors have the following variances ${\sigma_i}^2$: ${\sigma_1}^2 = {\sigma_3}^2 = 8 \times 10^{-3}$ for moves [1] and [3]; ${\sigma_2}^2 = {\sigma_4}^2 = 0.0025$ for moves [2] and [4]; ${\sigma_5}^2 = 10^{-6}$ for move [5]. Moves [6] and [7] are called trans-dimensional moves because they imply the changing of the number of variables associated to the candidate model with respect to the current model. Such moves are defined as in Appendix B in *Piana Agostinetti and Malinverno* [2010], so that the determinant of their Jacobian matrix is equal to 1. We follow the approach developed in *Mosegaard and Tarantola* [1995] for moves [6] and [7]. Thus, we make use of a sampler that walks across the prior distributions (the so-called *sampling from the priors* approach), and we accept or reject the candidate model with the probability in Equation 3. It is worth noticing that *sampling from the priors* can be quite inefficient if the data contain a lot of information about the investigated parameters, and thus the PPD likely differs from the prior probability distribution. On the contrary, if there is limited information contained in the data, *sampling from the priors* is a convenient sampling strategy, as it removes the need to define a proposal distribution [as in, e.g., *Bodin et al.*, 2012a].

4. Results

4.1. Simple cases: Misplaced sensors or changes in the physical properties of the rocks

In this section we consider three simple tests. As a first illustration of the algorithm, we construct a monitor survey which mimics the mis-placement of some sensors (Figure 5). The baseline survey is composed of twenty traces (Wiggle numbers: 5, 10, 15, ..., 100) from the first experimental set-up (Plexiglas/air). For the monitor survey, we use the same traces as in the baseline survey, and substitute five traces (Wiggle numbers: 50, 55, ..., 70) with shifted traces (Wiggle numbers: 54, 59, 64,..., 74, all positions have been shifted by the same amount) from the same Plexiglas's/air experimental set-up. In this way, the amplitude of the arrivals do not have relevant changes, but we introduce a temporal shift. It is worth noticing that the number of traces used, the number of shifted traces, and the shift amplitude have been selected to keep a reasonable number of traces in the inversion (20 wiggles out of 100 available) while having enough space to introduce a significant shift in the traces (four wiggles). The results are obtained by running 5 parallel McMC samplings. Each chain is composed of 2×10^6 models, half of which are discarded as part of the burn-in phase [*Somogyvari and Reich*, 2019]. For each chain, we used 20 CPUs on a Linux cluster for about 17 hours. The full computation time was about 5x350 core-hours. Computation time is almost constant across all tests presented in this study, due to the same number of traces and the limited number of rectangular cells used by the trans-D sampler.

In Figure 5, we show the most relevant information extracted from the PPD, together with the monitor and baseline surveys. The misplaced traces in the monitor survey are marked (yellow box in Figure 5b). For each point in the discrete (x, t)-space, we compute the 1D marginal PPD of w_{ij} and plot its mean posterior value (Figure 5c) and standard deviation (*std*, Figure 5d). As a rule of thumb, high values of the mean posterior w_{ij} indicate regions where the baseline and monitor surveys differ the most. Low and high values of the std differentiate well- and less- constrained regions, respectively. Our results illustrate how the algorithm works in this simple case. Due to the kind of distortion used, i.e. misplaced sensors, we should attribute almost the same weight to the entire set of misplaced traces. The algorithm accomplishes this task using a limited number of rectangular cells (about 20 cells, see Figure S1), confined in the vertical area of misplaced traces. The std also displays the same pattern with low values indicating a robust result. Due to the realistic nature of our test (traces obtained in laboratory and not synthetic traces), the results are not "perfect" and there are some anomalies (higher std for surface arrivals and a vertical stripe in the std plot within the misplaced traces) due to complexity in the experimental set-up (hardware noise).

The performance of the algorithm (Figure S1) highlights some key-aspects of the sampling. First, we are not overfitting the data because the number of cells in the sampled models is limited, and thus so is the number of inverted parameters. The acceptance probability for trans-D moves is very low, so we need long chain (> 1 million of models) to guarantee the necessary exploration of the data-space. However, after 1 million models, the number of cells used is almost stable between 15 and 30, but not constant, i.e. chains are still sampling models with variable number of dimensions but within a limited range of values.

Our second test is designed to complement the previous one and considers a monitor survey where only changes in the reservoir state are present (Figure 6). In this case, we make use of the same baseline as in the previous test, but in the monitor survey we substitute five traces (Wiggle number: 50 to 70) with the traces recorded at the same position but for the Plexiglas/water experimental set-up. Both posterior mean and std of w_{ij} share the same structure, with a vertical block and a pinched horizontal structure. The main difference in the results, with respect to the previous test, is the presence of a dark (large weights) spot in the location of the change in the reservoir-state, i.e. limited to the arrivals from the top of the reservoir and not including the surface waves (Figure 6c). Also, while the results contain a vertical stripe in the mean posterior w_{ij} in the region of the reservoir changes, as in Figure 5c, the std along the same stripe is very large. Horizontally, the rectangular cells seem to be able to move slightly and the dark region in the mean posterior w_{ij} (defining the reservoir changes) propagates across some traces, suggesting a higher vertical than horizontal resolution.

The third test considers the presence of both reservoir-changes and receiver misplacement in two separated regions of the (x,t)-space (Figure 7). In this case, while the baseline is kept the same as in previous tests, the monitor survey is composed as follows: for the misplaced sensors, three traces (wiggle numbers 15, 20 and 25) are replaced with wiggles from the same experimental set-up but with a 4 wiggle shift (so replaced with wiggle numbers: 19, 24 and 29); for the reservoir-changes, we substitute seven traces from 60 to 90, with the wiggles recorded in the same position but with the second experimental set-up. Note that the number of traces representing the two anomalies is different from the previous tests, to keep them separated and to be able to split it into two regions (see next section).

The results clearly show that, in the case of not-interacting anomalies, the two kinds of distortions can be separately identified (Figure 7c). Both anomalies can be seen in the mean posterior of w_{ij} with the same characteristics as in the previous tests. In the analysis of the std there is a clear difference, with respect to the previous tests, in the bright spot defining the reservoir-change, but also in the value (lower here) of the vertical stripe defining the misplaced sensors. However, such changes could be attributed to the different numbers of traces composing the anomalies (Figure 7d), indicating that the std is more sensitive to the lateral extension of the anomaly than to the mean posterior value.

4.2. Complex case: simultaneous retrieval of misplaced sensor and changes in the physical properties of the rocks

The most interesting case represents the co-existence of both misplaced receivers and reservoir-changes in the same region of (x,t)-space. To test this, the baseline is kept the same as in previous tests. The monitor survey is composed of the baseline traces with substitutions in three different and contiguous regions. In the first region, called “A”, six traces are substituted by shifted wiggles from the same experimental set-up (i.e. mimic misplacement receivers only: wiggles numbers 30, 35, ..., 55 are replaced with 34, 39, ..., 59). Also in the second region “B” we have misplaced traces (three traces, wiggles numbers 60,65 and 70 replaced with 64, 69 and 74) but from the second experimental set-up, to simultaneously reproduce both misplaced receivers and reservoir-changes. Finally in the third region ”C”, we consider reservoir changes only. Four traces (wiggles numbers 75 to 90) are replaced with the wiggles recorded in the same position, but from the second experimental set-up. The minimum region dimension is three traces, but the “misplaced sensors” anomaly covers nine traces, while the “reservoir-changes” anomaly covers seven traces (Figures 8 and 9).

As expected, the outcomes from a complex case are more challenging to describe. The mean posterior of w_{ij} still clearly defines the reservoir changes as a dark (large values) elongated region that covers exactly the expected traces (Figure 8 and Figure 9b). However, recognizing the boundaries between regions “A” and “B”, and “B” and “C” is not easy in the mean posterior. In fact the value of the mean posterior of w_{ij} does not change significantly through regions “A” to “C” away the reservoir-changes zone, with fluctuation given by experimental noise and lateral smearing of the reservoir-changes anomaly. It is hard to recognise which traces have only been shifted (from the region between traces number 1 to 5 where the two surveys share the same wiggles) or which traces are both shifted and have a reservoir-change. Knowing the monitor survey composition, we can see that more traces than the ones composing region “C” have been locally perturbed, from the occurrence of the high-weights at localised times (dark region), but we cannot really discriminate which of the traces that also have the reservoir-change signature have been displaced.

The results for the posterior std of the w_{ij} furnish some additional insights into the separation of the three regions. In fact, comparing both mean and std shows that the posterior std is generally uniform, but very large in the region where we only have reservoir changes (as seen also in Figure 6). The posterior std is lower and more variable for the region where we have misplaced traces (both with and without simultaneous reservoir changes). In practise, only the simultaneous analysis of both mean and std posterior for w_{ij} can somewhat unequivocally define the three regions.

Finally, the posterior std is very low in the core of the reservoir-changes anomaly, as found in the previous test (compare to Figure 8d), likely caused by the large lateral extent of the anomaly (quite large, seven traces (one third of the total)). Moreover, we observe that the area of the std where we only have misplaced sensors is not uniform as expected, due to the interaction with the reservoir anomaly (anomaly lateral smearing). However, the std is large where the two anomalies interact.

5. Discussion

We propose a new methodology for exploring 4D seismic data and detecting potential noise sources other than random ambient noise, and relevant signals from the alteration of a reservoir. The algorithm has been proven to correctly perform in isolating simple case scenarios (one noise source or one reservoir change, or both present in two different portions of the 4D seismic data). In such cases, our algorithm identifies the different anomalies and their position, and it is able to characterise them in terms of both the amplitude of the posterior weights and their standard deviation. In particular, anomalous signals related to a misplacement of the sensors is identified as a broad portion of the monitoring survey where the posterior weights are uniformly increased by a limited amount, and their standard deviation is uniform too. Conversely, in the portion of the monitoring survey where the anomaly is related to a reservoir change, the posterior weights are extremely high in a localised 2D patch. Their standard deviation also displays a peculiar pattern, with very low values in the inner portion of the anomaly and very high values along its border. We suggest that the rapid change in the standard deviation is the key-element that can define the shape of the anomaly related to reservoir changes.

In more complex cases, i.e. where both noise sources and reservoir signals coexist, the interpretation of the results is more challenging. Dis-aggregating co-existing changes/mis-positioning is not easy (Figure 9), but we observe that reservoir changes are always the most striking and isolated feature. Also in this case, the analysis of the standard deviation of the weights is a critical point for making inferences. In fact, even here the sharp change in the standard deviation defines the border of the anomaly given by reservoir changes. Moreover, the standard deviation also helps to define the area where the mis-placed sensors are present (these regions have a lower standard deviation compared to area where only reservoir changes are present). It is worth noting that the estimation of the standard deviation of the weights is a brand new outcome of our algorithm, given by our statistical approach to data-space exploration.

Our results display to some extent the boundaries of our rectangular patches (i.e they seems to have a block-structure). Such blockiness indicates the resolution limits of our model to some extent, and are related to our choice of rectangular partitions. In trans-D algorithms, the effects of the parameterization on the retrieved results is an on-going research field [e.g. *Gao and Lekic*, 2018]. Here, we suggest that other choices of partition shape could be more efficient on biggerscale data, such as the *anisotropic kernels,* proposed in *Belhadj et al.* [2018], which could more easily reproduce the true shape of anomalies in field measurements.

Our approach to 4D seismic data analysis could be used to support more complex data workflows adopted in energy industries. In Figure 10, we compare the results of our complex case, with a standard analytic indicator (NRMS) commonly used in data-workflow for 4D seismics. Comparing Figure 10a and 10b, it seems that mis-positioning is the most impactful issue in terms of likelihood between baseline and monitoring surveys, but it is easily separated from reservoir changes, which have the strongest W_{ij} in our case. As seen in Figure 10c, NRMS is clearly higher in the area of sensor misplacement. Such an anomaly masks the signal coming from the "altered conditions in the reservoir". In fact such a signal can be seen as a small amplitude anomaly (i.e. around 40% at trace 16-19, still higher NRMS with respect to trace 1-5 where no anomaly is present at all), but it is totally obscured between traces 11 and 15, where the dominant effect is the sensor misplacement. Our approach could be used as a support to standard data-workflow and could save time during subsequent physical modelling of the reservoir (an extremely time-consuming task). Because it makes no preliminary assumption on the reservoir geometry, our approach does not risk bringing an initial bias into the results and thus could furnish more reliable information on the state of the geo-resources.

6. Conclusions

In this study, we presented a new methodology for the exploration of the data-space. We followed a trans-D sampling approach to recreate and validate data-structures in the form of partitions of the covariance matrix. We applied the new methodology to 4D seismic data acquired for monitoring the sub-surface. Our results indicate that:

1. the trans-D approach can be applied to data-space exploration for defining unknown data-structures and separating data-volumes that are coherent with a-priori physical hypotheses;

2. the analysis of the full PPD of the data-structures can be used for classifying different sources of 4D signal, like repeatability noise and 4D signal from the geo-resources;

3. In comparison with standard measures of repeatability like NRMS, our approach is less biased by the presence of different sources if 4D signal in the same data-volume and can be used to efficiently separate such sources.

In the future, we will further develop our methodology to include different shapes and orientation of the partitions [i.e. not rectangular patches, also called *anisotropic kernels, as in Belhadj et al.*, 2018] for increasing the efficiency of the McMC sampling; and to consider 3D partitions and the comparison of two entire 3D volumes.

7. Acknowledgements

NPA would like to thank Daniele Melini at INGV for assistance with the linux cluster. NPA publications are printed with the financial support of the Austrian Science Fund (FWF), project number: M2218-N29. We are also grateful for support at Memorial provided by Chevron and with grants from the Natural Sciences and Engineering Research Council of Canada Industrial Research Chair Program and InnovateNL (IRCPJ 491051-14). We would also like to thank Kamal Moravej for his help and instructions at the lab in order to collect the data used in this study. Raw data (i.e. waveform used as baseline and monitoring surveys) can be requested to MK via email: mk7251@mun.ca. The Generic Mapping Tools software was used for plotting the figures of this manuscript [*Wessel and Smith*, 1998].

References

Bayes, T. (1763), An essay towards solving a problem in the doctrine of chances, *Philos. Trans. R Soc. London,*, *53*, 370–418.

Belhadj, J., T. Romary, A. Gesret, M. Noble, and B. Figliuzzi (2018), New parameterizations for bayesian seismic tomography, *Inverse Problems*, *34*(6), 065,007, doi:10.1088/1361-6420/aabce7.

Birnie, C., K. Chambers, D. Angus, and A. L. Stork (2020), On the importance of benchmarking algorithms under realistic noise conditions, *Geophysical Journal International*, *221*(1), 504–520, doi:10.1093/gji/ggaa025.

Bodin, T., M. Sambridge, N. Rawlinson, and P. Arroucau (2012a), Transdimensional tomography with unknown data noise, *Geophys. J. Int*, doi: 10.1111/j.1365-246X.2012.05414.x.

Bodin, T., M. Sambridge, H. Tkalcic, P. Arroucau, K. Gallagher, and N. Rawlinson (2012b), Transdimensional inversion of receiver functions and surface wave dispersion, *J. Geophys. Res.*, *117*(B02301), doi:10.1029/2011JB008560.

Cheng, A., L. Huang, and J. Rutledge (2010), Time-lapse VSP data processing for monitoring co2 injection, *The Leading Edge*, *29*(2).

Davis, T. L., M. J. Terell, R. D. Benson, R. Cardona, R. R. Kendall, and R. Winarsky (2003), Multicomponent seismic characterization and monitoring of the CO2 flood at Weyburn field, Saskatchewan, *The Leading Edge*, *22*(7), 606–700.

Dettmer, J., and S. E. Dosso (2012), Trans-dimensional matched-field geoacoustic inversion with hierarchical error models and interacting Markov chains, *J. Acoust. Soc. Am.*, *132*(4), 2239–2250.

Dramsch, J. S. (2019), Machine learning in 4d seismic data analysis: Deep neural networks in geophysics, Ph.D. thesis, Technical University of Denmark, Lyngby, Denmark.

Galetti, E., A. Curtis, B. Baptie, D. Jenkins, and H. Nicolson (2016), Transdimensional Love-wave tomography of the British Isles and shear-velocity structure of the East Irish Sea Basin from ambient-noise interferometry, *Geophysical Journal International*, *208*(1), 36–58, doi:10.1093/gji/ggw286.

Gao, C., and V. Lekic (2018), Consequences of parametrization choices in surface wave inversion: insights from transdimensional Bayesian methods, *Geophysical Journal International*, *215*(2), 1037–1063, doi:10.1093/gji/ggy310.

Geyer, C. J., and J. Møller (1994), Simulation procedures and likelihood inference for spatial point processes, *Scand. J. Stats*, *21*, 359–373.

Green, P. J. (1995), Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, *Biometrika*, *82*(4), 711–732.

Houck, R. T. (2007), Time-lapse seismic repeatability – How much is enough?, *The Leading Edge*, *26*(7), –.

Kolb, J. M., and V. Lekić (2014), Receiver function deconvolution using transdimensional hierarchical Bayesian inference, *Geophysical Journal International*, *197*(3), 1719–1735, doi:10.1093/gji/ggu079.

Koster, K., P. Gabriels, M. Hartung, J. Verbeek, G. Deinum, and R. Staples (2000), Time-lapse seismic surveys in the North Sea and their business impact, *The Leading Edge*, *19*(3), 286–293, doi:10.1190/1.1438594.

Kragh, E., and P. Christie (2002), Seismic repeatability, normalized RMS, and predictability, *The Leading Edge*, *21*(7), 640–647, doi:10.1190/1.1497316.

Landro, M. (1999), Repeatability issues of 3D VSP data, *Geophysics*, *64*(6), –.

Lumley, D. (2010), 4D seismic monitoring of CO_2 sequestration, *The Leading Edge*, *29*(2), https://doi.org/10.1190/1.3304817.

Malinverno, A. (2002), Parsimonious Bayesian Markov chain Monte Carlo inversion in a non-linear geophysical problem, *Geophys. J. Int.*, *151*(3), 675–688.

Malinverno, A., and V. A. Briggs (2004), Expanded uncertainty quantification in inverse problems: Hierarchical Bayes and empirical Bayes, *Geophysics*, *69*(4), 1005–1016, doi:10.1190/1.1778243.

Mechelen, I. V., A.-L. Boulesteix, R. Dangl, N. Dean, I. Guyon, C. Hennig, F. Leisch, and D. Steinley (2018), Benchmarking in cluster analysis: A white paper.

Mosegaard, K., and A. Tarantola (1995), Monte Carlo sampling of solutions to inverse problems, *J. Geophys. Res.*, *100*(B7), 12,431–12,447.

Pevzner, R., V. Shulakova, A. Kepic, and M. Urosevic (2011), Repeatability analysis of land time-lapse seismic data: CO2 CRC Otway pilot project case study, *Geophysical Prospecting*, *59*(1), –.

Piana Agostinetti, N., and A. Malinverno (2010), Receiver Function inversion by trans-dimensional Monte Carlo sampling, *Geophys. J. Int*, *181*, doi:10.1111/j.1365-246X.2010.04530.x.

Piana Agostinetti, N., and A. Malinverno (2018), Assessing uncertainties in high-resolution, multi-frequency receiver function inversion: a comparison with borehole data, *Geophysics*, *83*(3), KS11–KS22, doi: 10.1190/geo2017-0350.1.

Piana Agostinetti, N., and F. Martini (2019), Sedimentary basins investigation using teleseismic p-wave time delays, *Geophysical Prospecting*, *67*(6), 1676–1685, doi:10.1111/1365-

2478.12747.

Porter-Hirsche, J., and K. Hirsche (1998), *Repeatability study of land data acquisition and processing for time lapse seismic*, pp. 9–11, doi:10.1190/1.1820663.

Roach, L. A. N., D. J. White, and B. Roberts (2015), Assessment of 4d seismic repeatability and co2 detection limits using sparse permanent land array at the aquistore co2 storage site, *Geophysics*, *80*(2).

Sambridge, M., and K. Mosegaard (2002), Monte Carlo methods in geophysical inverse problems, *Rev. Geophys.*, *40*(3), doi:10.1029/2000RG000,089.

Sambridge, M., K. Gallagher, A. Jackson, and P. Rickwood (2006), Trans-dimensional inverse problems, model comparison and the evidence, *Geophys. J. Int.*, *167*(2), 528–542, doi:10.1111/j.1365-246X.2006.03155.x.

Sambridge, M., T. Bodin, K. Gallagher, and H. Tkalcic (2013), Transdimensional inference in the geosciences, *Phil. Trans. R. Soc. A*, *371*, 20110547.

Somogyvari, M., and S. Reich (2019), Convergence tests for transdimensional markov chains in geoscience imaging, *Math Geosci*, https://doi.org/10.1007/s11004-019-09811-x.

Steininger, G., J. Dettmer, J. Dosso, and S. Holland (2013), Transdimensional joint inversion of seabed scattering and reflection data, *J. acoust. Soc. Am.*, *133*, 1347–1357.

Tarantola, A. (1987), *Inverse problem theory: methods for data fitting and model parameter estimation*, Elsevier Science publishing Co.

Tarantola, A. (2005), *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM.

Tarantola, A. (2006), Popper, Bayes and the inverse problem, *nature physics*, *2*.

Wessel, P., and W. H. F. Smith (1998), New, improved version of the generic mapping tools released, *EOS Trans. AGU*, *79*, 579.

Xiang, E., R. Guo, S. E. Dosso, J. Liu, H. Dong, and Z. Ren (2018), Efficient hierarchical trans-dimensional Bayesian inversion of magnetotelluric data, *Geophysical Journal International*, *213*(3), 1751–1767, doi:10.1093/gji/ggy071.

Yang, D., A. Malcolm, M. Fehler, and L. Huang (2014), Time-lapse walkaway vertical seismic profile monitoring for CO_2 injection at the SACROC enhanced oil recovery field: A case study, *Geophysics*, *79*(2).

Variables	Description
N_w	number of traces in the survey
N_s	number of samples per trace
i, k	indices for samples
j	index for a trace
$\mathbf{x}_{ij}$	space (x_i) and time (x_j) position of the i-th point for the j-th trace
b_{ij}	amplitude of baseline survey at the i-th point for the j-th trace
m_{ij}	amplitude of monitor survey at the i-th point for the j-th trace
$\mathbf{e_{ij}} = (b_{ij} - m_{ij})$	sample-wise difference between baseline and monitor surveys (at the i-th point for the j-th trace)
Terms	**Description**
	Data
shot-gather	original data from the laboratory, one for each experimental set-up
wiggle	one recording (in time) at a fixed position within one shot-gather
survey	input data for the algorithm: new shot-gather composed of selected wiggles
trace	one recording of the survey
4D signal	differences in the monitoring and baseline surveys
	Sources of 4D signal
target signal	changes in reservoir properties
noise	ambient random noise
perturbation	sensor misplacement.

Table 1. Description of variables and terminology.

Model parameter	Minimum	Maximum
Number of cells, n	1	200
Cell centre along x-axis, c_n	1	20
Cell radius, r_n	1	10
Cell centre along t-axis, t_n	1	1251
Cell time-window, s_n	1	625
Weight, π_n	0.0	1.0

Table 2. Uniform prior distributions of model parameters in the **m** vector.

Figure 1. Experimental setup and photos of the equipment. (a) function generator showing the parameters of the source pulse (b) oscilloscope showing an example of a recorded wiggle. The red spot on the model is the location of the laser receiver, which is moved vertically in controlled increments to generate wiggles at different locations, which are combined into the final shot record.

Figure 2. Example of seismic surveys: (a) Baseline survey using all wiggles generated with air/Plexiglas interface. (b) Monitor survey. Same wiggles as in (a), but: wiggles from 15 to 25 have been replaced with the wiggles from 19 to 29, same interface (simulating misplaced receivers); wiggles from 60 to 89 have been replaced with wiggles recorded in the same position but with a different interface (water/Plexiglas, simulating a change in the physical properties of the reservoir). (c) Squared differences of the two survey, computed for each sample separately. Notably the largest values are associated with “misplaced receivers”. See Section 4.1 for the details of this experiment.

Figure 3. Example of data analysis for reconstructing the Covariance matrix of the error associated to trace 155. (a) Zoom of the traces close to trace 155. The yellow box indicates the traces used for estimating the standard deviation and the correlation model needed to compose the Covariance matrix. (b) Stack and standard deviation for the traces in the yellow box in (a). The orange line and the dashed orange lines represent the stack and the standard deviation, respectively. Grey lines report the traces in the yellow box in (a). (c) Residuals between the stack and each single trace in the yellow box in (a). (d) Auto-correlation of the residuals in (c). The orange line shows the average of all autocorrelation curves (grey lines). The green line displays the best-fitting curve, modelled using the function in Eq. 6 [*Kolb and Lekić*, 2014].

ADDITIONAL FIGURES

Figure 4. Example of a model. The rectangles represent the cell, coloured according to their weights. Where cells overlap, weights are summed. Each data point (dots) has an associated weight. Data points outside all cells are associated to a weight $w_{ij} = 0.0$. Yellow circles represent cell nuclei. To make the figure readable, only one of every 15 data-point is plotted.

Figure 5. Results for a simple case: misplacement of receivers. (a) Baseline survey. The grey area denotes where the signal is absent. (b) Monitor survey. See Section 4.1 for details on how the monitor survey is created. (c) Mean posterior weight w_{ij} associated to each data point (i-th sample on the j-th trace). (d) Posterior standard deviation of w_{ij}. The yellow box indicates the wiggles that changed between the Baseline and Monitor surveys.

Figure 6. Results for a simple case: changes in the physical properties of the reservoir. See Figure 5 for details.

Figure 7. Results for a complex case: misplacement of receivers and changes in the physical properties of the reservoir, separated. See Figure 5 for details.

Figure 8. Results for a complex case: misplacement of receivers and changes in the physical properties of the reservoir, overlapping. See Figure 5 for details. Yellow boxes indicate changes between monitoring and baseline surveys in Figure 5 have been removed for improving readability.

Figure 9. Details of the results for a complex case: misplacement of receivers overlapping changes in the physical properties of the reservoir. (a) Monitor survey. The three letters indicate different area with: [A] misplaced receivers; [B] misplaced sensors and changes in the reservoir, and [C] only changes in reservoir. (b) Mean posterior weight W_{ij} associated with each data point (i-th sample at the j-th trace). (c) Posterior standard deviation of W_{ij}. See Section 4.2 for details.

Figure 10. (a) Mean posterior weight W_{ij} associated with each data point (i-th sample at the j-th trace). Posterior standard deviation of W_{ij} is shown as red contour lines. See Section 4.2 for details. (b) Same as in Figure 2c, point-wise L_2 difference between monitoring and baseline surveys. (c) NRMS for each trace of the monitoring survey with respect to baseline survey. NRMS computed as in *Kragh and Christie* [2002]

Figure S1.Details on the trans-D sampling for the simple case: misplacement of receivers. (a) PPD for the number of sectors in the model. (b) Acceptance rate for the seven moves composing the recipe for the trans-D sampling. Outcomes for each move are labelled as: "+1", move has been accepted (candidate model improved the fit); "0", move has been rejected; "-1", move has been accepted, but the candidate does not improve the fit. (c) Variation of the number of cells in the sampled models for all five chains. A blue box indicates the "burn-in" period for which sampled models are not considered.

Figure 1.

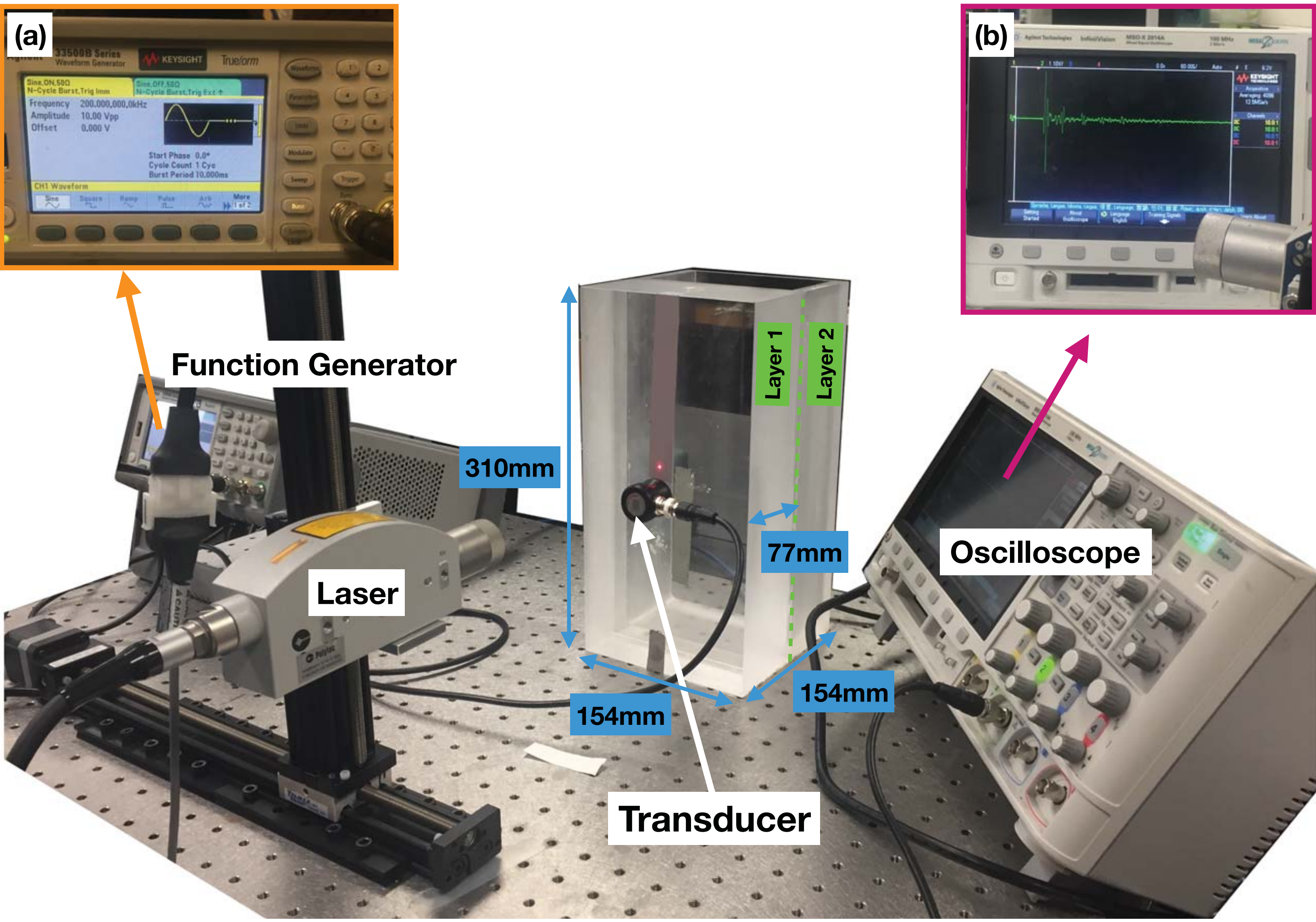

(a)
(b)
Function Generator
Laser
310mm
Layer 1
Layer 2
77mm
154mm
154mm
Transducer
Oscilloscope

Figure 2.

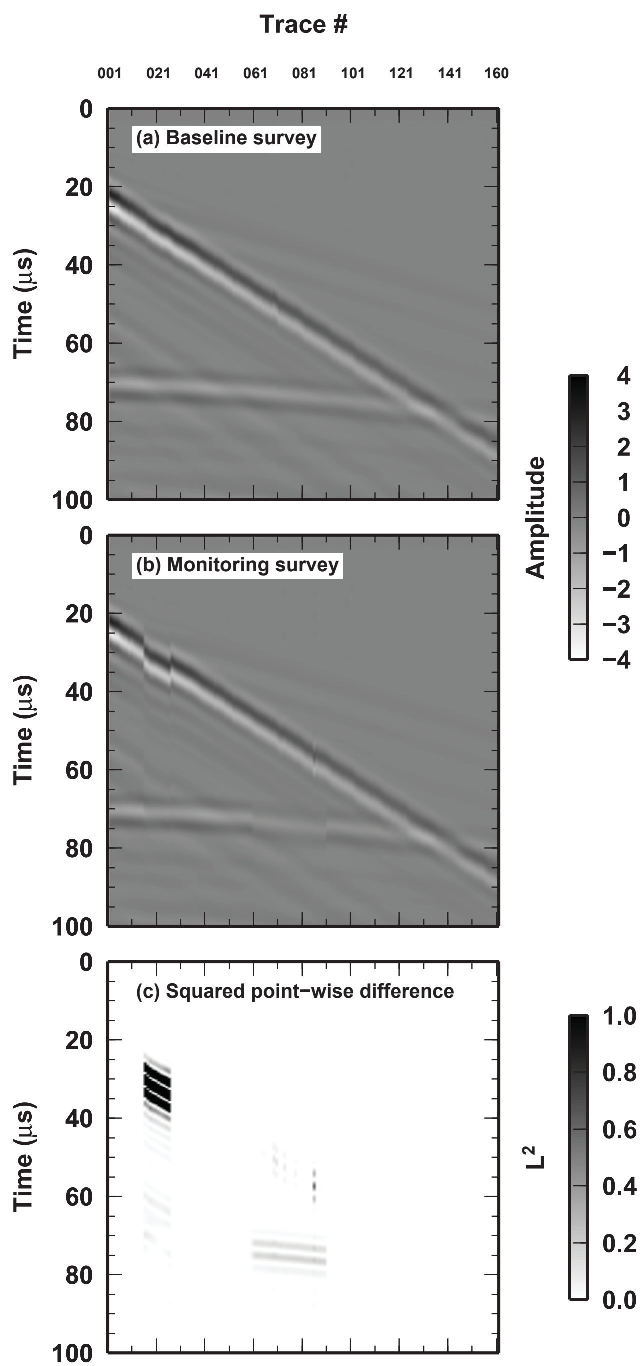

Trace #
001 021 041 061 081 101 121 141 160
(a) Baseline survey
(b) Monitoring survey
(c) Squared point−wise difference
Time (µs)
0 20 40 60 80 100
Amplitude
4 3 2 1 0 −1 −2 −3 −4
L^2
1.0 0.8 0.6 0.4 0.2 0.0

Figure 3.

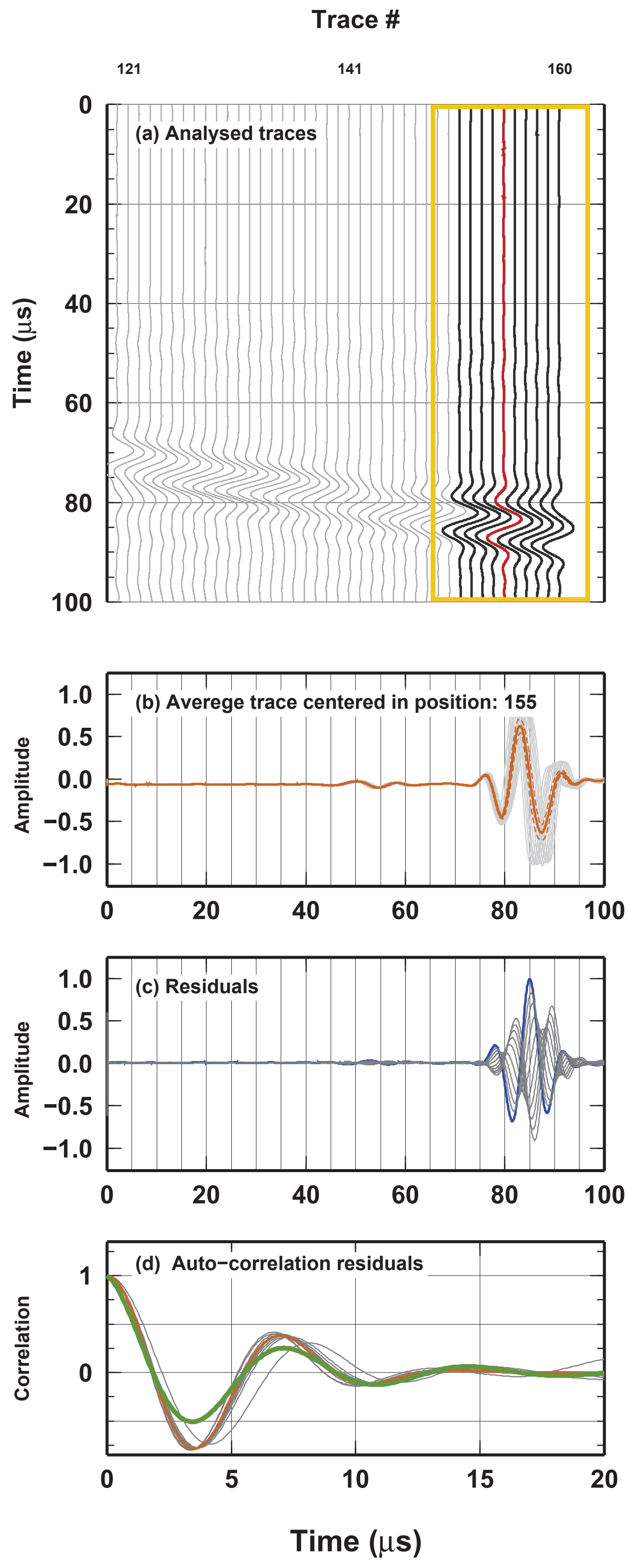
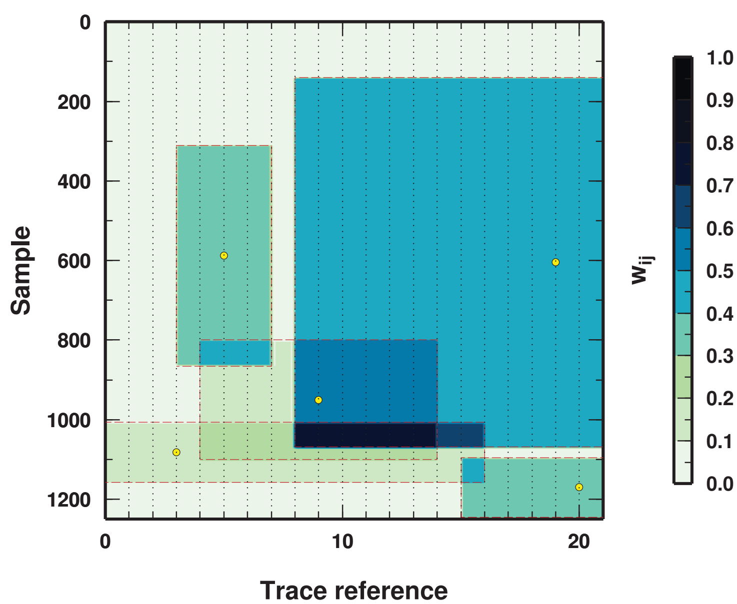
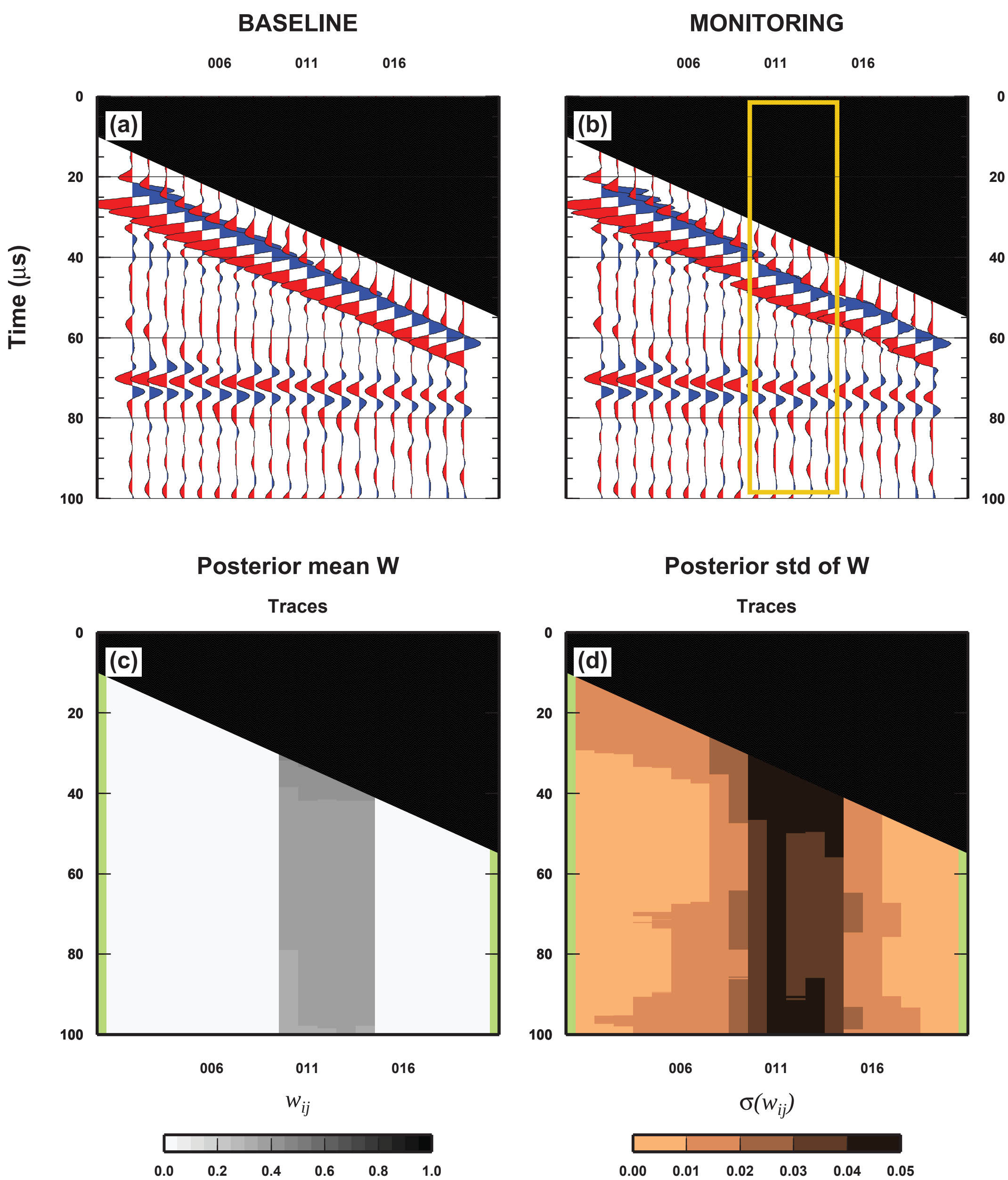

Trace #
121
141
160
Time (µs)
(a) Analysed traces
(b) Averege trace centered in position: 155
Amplitude
(c) Residuals
Amplitude
(d) Auto−correlation residuals
Correlation
Time (µs)

Figure 4.



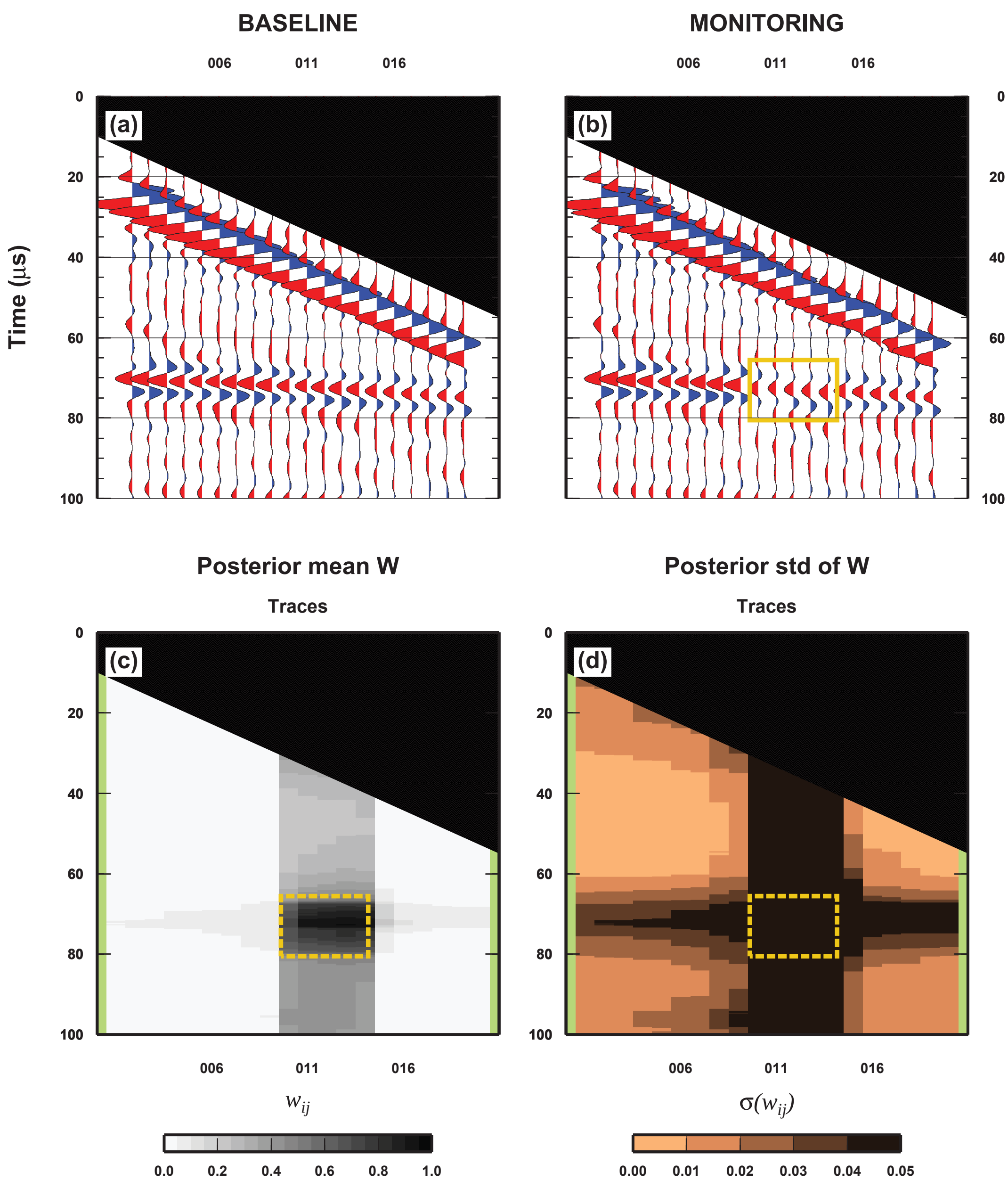
0
200
400
600
800
1000
1200
Sample
0
10
20
Trace reference
w_{ij}
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0

Figure 5.



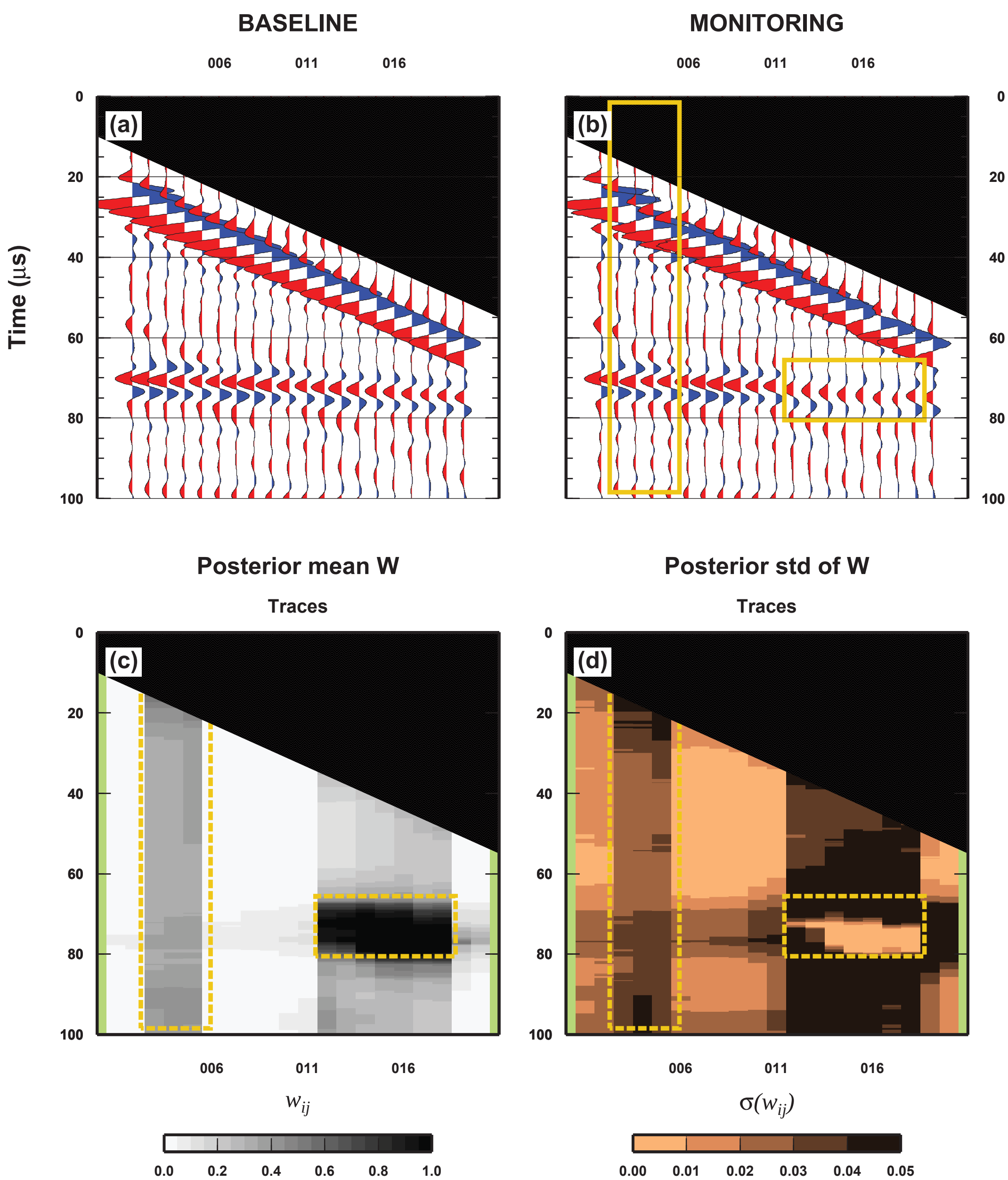
BASELINE
MONITORING
006
011
016
(a)
(b)
Time (µs)
0
20
40
60
80
100
Posterior mean W
Posterior std of W
Traces
(c)
(d)
w_{ij}
$\sigma(w_{ij})$
0.0
0.2
0.4
0.6
0.8
1.0
0.00
0.01
0.02
0.03
0.04
0.05

Figure 6.



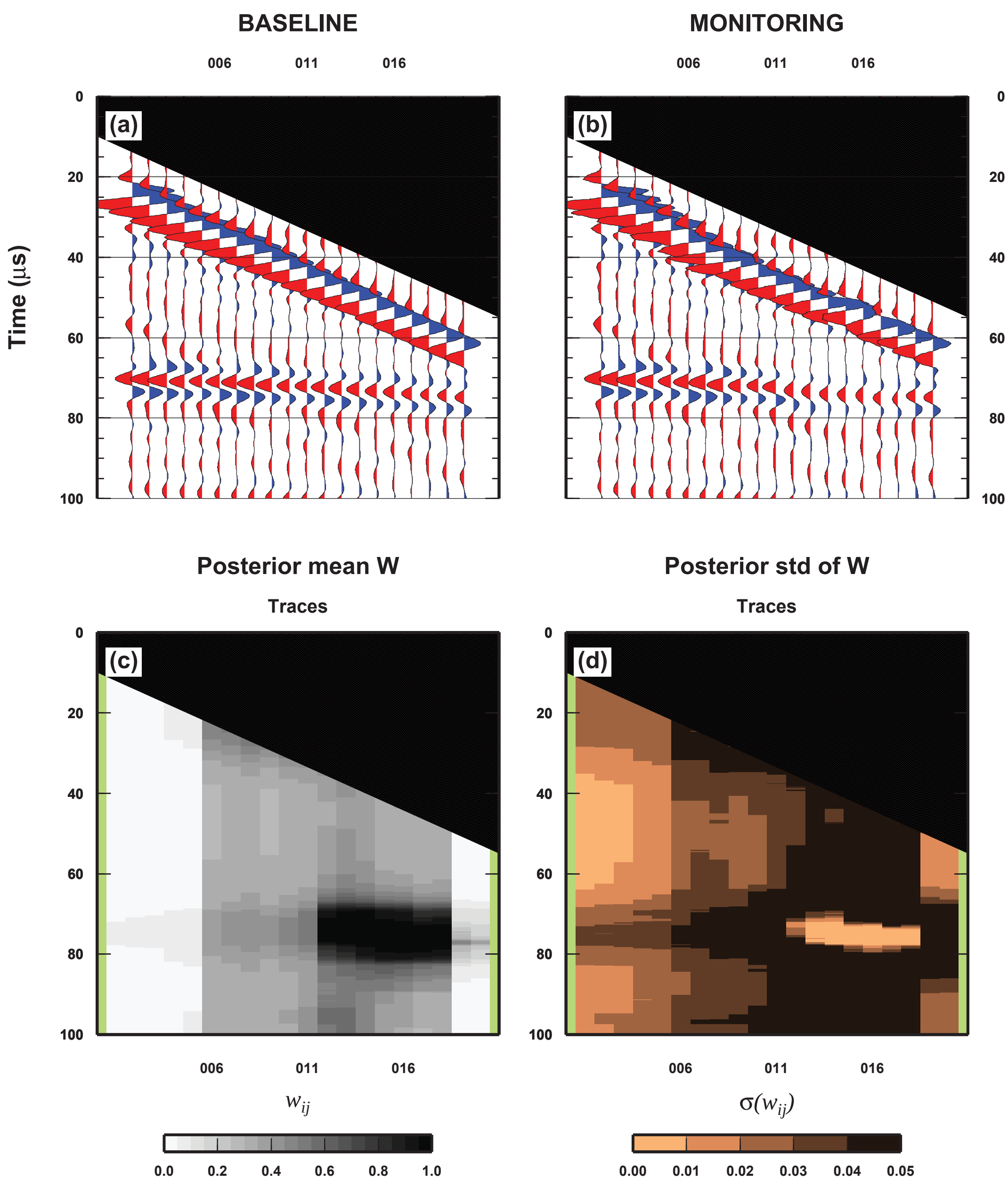
BASELINE
MONITORING
006
011
016
Time (µs)
0
20
40
60
80
100
(a)
(b)
Posterior mean W
Posterior std of W
Traces
(c)
(d)
w_{ij}
$\sigma(w_{ij})$
0.0
0.2
0.4
0.6
0.8
1.0
0.00
0.01
0.02
0.03
0.04
0.05

Figure 7.



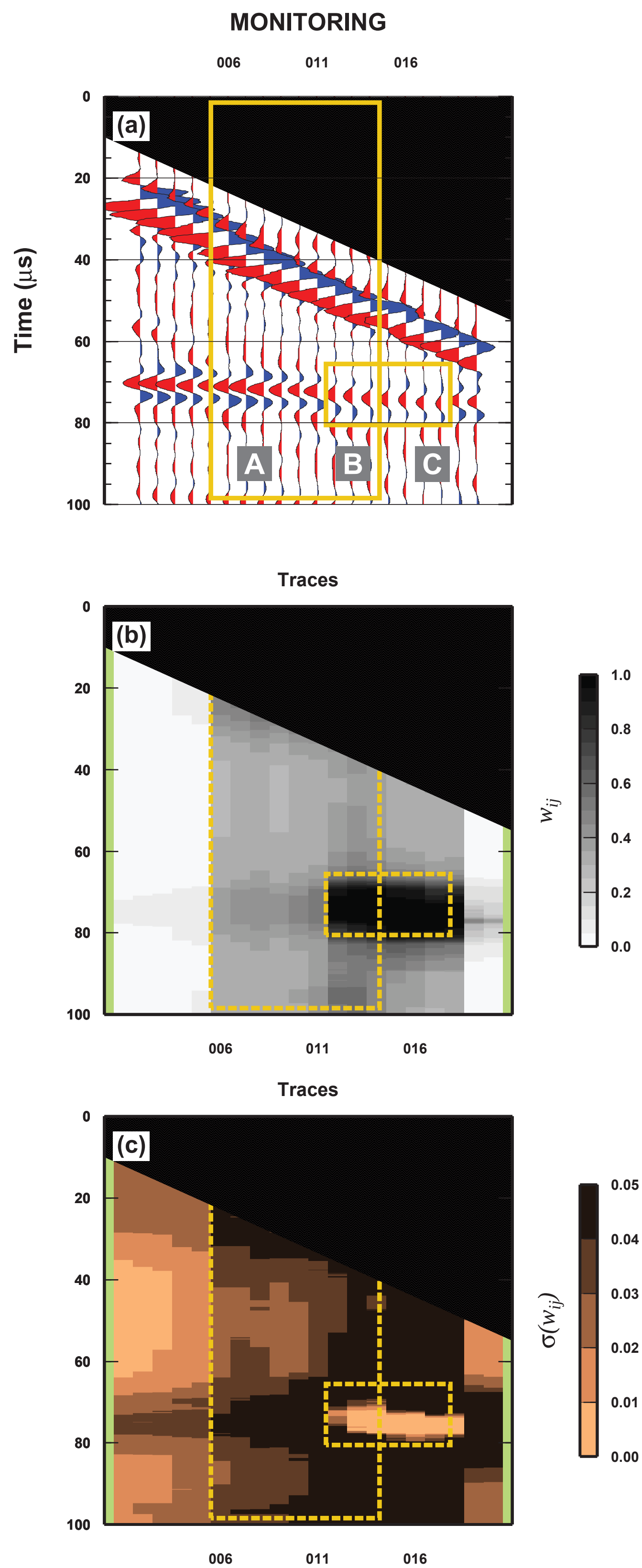
BASELINE
MONITORING
006
011
016
Time (μs)
(a)
(b)
0
20
40
60
80
100
Posterior mean W
Posterior std of W
Traces
(c)
(d)
w_{ij}
$\sigma(w_{ij})$
0.0
0.2
0.4
0.6
0.8
1.0
0.00
0.01
0.02
0.03
0.04
0.05

Figure 8.



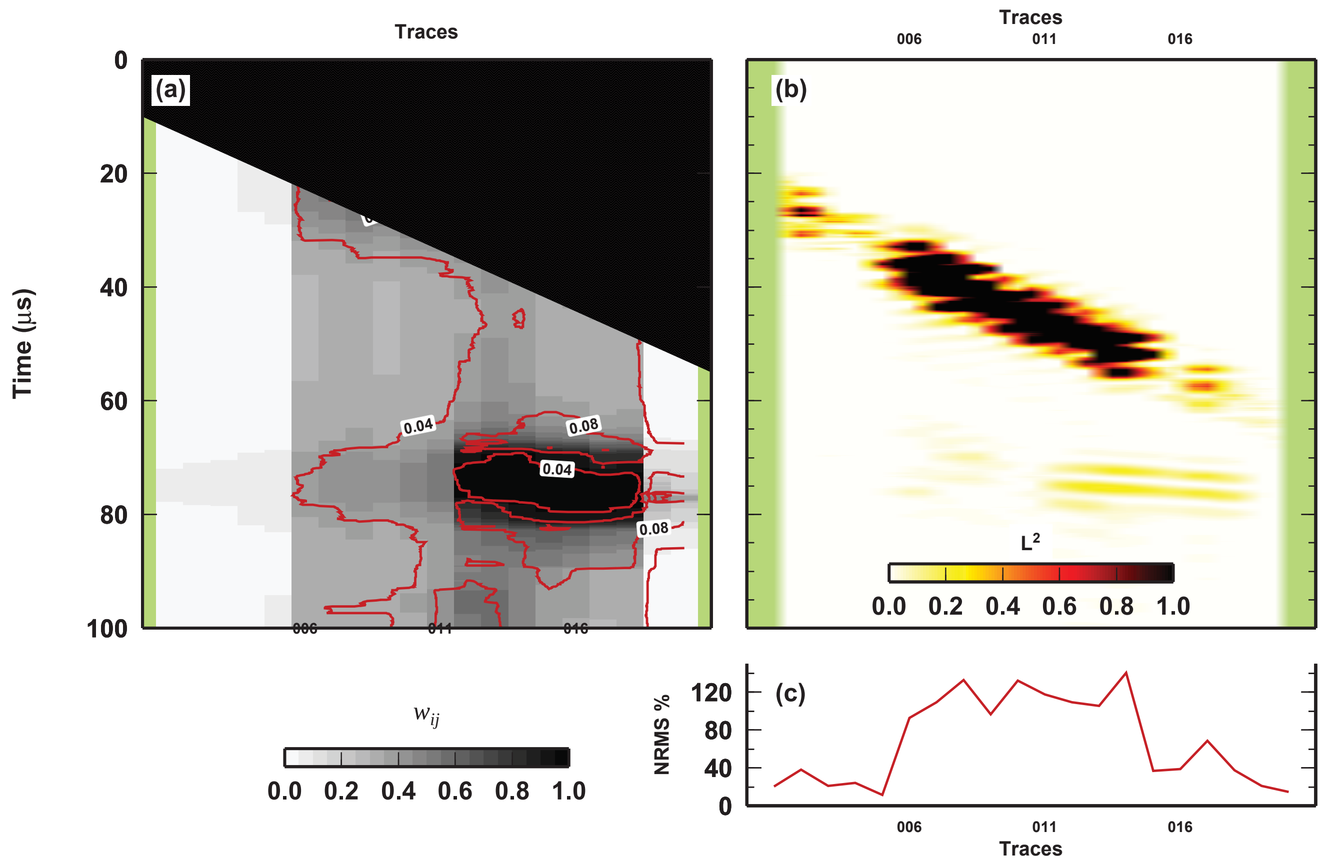
BASELINE
MONITORING
006
011
016
(a)
(b)
Time (µs)
0
20
40
60
80
100
Posterior mean W
Posterior std of W
Traces
(c)
(d)
w_{ij}
$\sigma(w_{ij})$
0.0
0.2
0.4
0.6
0.8
1.0
0.00
0.01
0.02
0.03
0.04
0.05

Figure 9.



MONITORING
006
011
016
0
20
40
60
80
100
Time (μs)
(a)
A
B
C
Traces
(b)
1.0
0.8
0.6
0.4
0.2
0.0
w_{ij}
Traces
(c)
0.05
0.04
0.03
0.02
0.01
0.00
$\sigma(w_{ij})$

Figure 10.



Traces
(a)
Time (µs)
0
20
40
60
80
100
0.04
0.08
0.04
0.08
006
011
016
w_{ij}
0.0 0.2 0.4 0.6 0.8 1.0
Traces
006 011 016
(b)
L^2
0.0 0.2 0.4 0.6 0.8 1.0
(c)
NRMS %
120
80
40
0
006 011 016
Traces