Optimizing measurement geometry for seismic near-surface full waveform inversion

Geophysical Journal International

Geophys. J. Int. (2017) **210,** 1909–1921
Advance Access publication 2017 June 20
GJI Marine geosciences and applied geophysics

doi: 10.1093/gji/ggx267

Optimizing measurement geometry for seismic near-surface full waveform inversion

André Nuber, Edgar Manukyan and Hansruedi Maurer

ETH Zurich, Institute of Geophysics, Sonneggstrasse 5, 8092 *Zurich, Switzerland. E-mail: andre.nuber@erdw.ethz.ch*

Accepted 2017 June 19. Received 2017 June 5; in original form 2016 November 23

SUMMARY

Full waveform inversion (FWI) is an increasingly popular tool for analysing seismic data. Current practise is to record seismic data sets that are suitable for reflection processing, that is, a very dense spatial sampling and a high fold are required. Using tools from optimized experimental design (ED), we demonstrate that such a dense sampling is not necessary for FWI purposes. With a simple noise-free acoustic example, we show that only a few suitably selected source positions are required for computing high-quality images. A second, more extensive study includes elastic FWI with noise-contaminated data and free-surface boundary conditions on a typical near-surface setup, where surface waves play a crucial role. The study reveals that it is sufficient to employ a receiver spacing in the order of the minimum shear wavelength expected. Furthermore, we show that horizontally oriented sources and multicomponent receivers are the preferred option for 2-D elastic FWI, and we found that with a small amount of carefully selected source positions, similarly good results can be achieved, as if as many sources as receivers would have been employed. For the sake of simplicity, we assume in our simulations that the full data information content is available, but data pre-processing and the presence of coloured noise may impose restrictions. Our ED procedure requires an *a priori* subsurface model as input, but tests indicate that a relatively crude approximation to the true model is adequate. A further pre-requisite of our ED algorithm is that a suitable inversion strategy exists that accounts for the non-linearity of the FWI problem. Here, we assume that such a strategy is available. For the sake of simplicity, we consider only 2-D FWI experiments in this study, but our ED algorithm is sufficiently general and flexible, such that it can be adapted to other configurations, such as crosshole, vertical seismic profiling or 3-D surface setups, also including larger scale exploration experiments. It also offers interesting possibilities for analysing existing large-scale data sets that are too large to be inverted. With our methodology, it is possible to extract a small (and thus invertible) subset that offers similar information content as the full data set.

Key words: Inverse theory; Waveform inversion; Seismic tomography.

1 INTRODUCTION

Seismic full waveform inversion (FWI) is a very promising tool for obtaining high-resolution images of the subsurface. The expected resolution is in the order of half the minimum wavelength (Wu & Toksöz 1987). The theory behind FWI was already developed in the 1980s (Lailly 1983; Tarantola 1984; Mora 1987), but the method only became popular recently due to the enormous computational expenses. Comprehensive FWI overviews can be found in Plessix (2008), Buske *et al*. (2009), Virieux & Operto (2009) and Fichtner (2011). The method is nowadays applied to a broad range of scales, ranging from laboratory investigations in the submetre range (e.g. Pratt 1999; Bretaudeau *et al*. 2013), engineering and environmental applications (a few tens to a few hundreds of metres, e.g. Smithyman *et al*. 2009; Romdhane *et al*. 2011), exploration problems (a few kilometres, e.g. Jaiswal *et al*. 2009; Bleibinhaus & Hilberg 2012; Raknes *et al*. 2015), active source wide-angle surveys (a few tens of kilometres, e.g. Operto *et al*. 2004; Malinowski *et al*. 2011), crustal-scale passive seismic investigations (a few tens of kilometres, e.g. Fichtner *et al*. 2013) to whole Earth studies (a few hundreds to a few thousands of kilometres, e.g. French & Romanowicz 2014).

The FWI methodology applied to all these scales is similar, but the specific goals and constraints impose different challenges and require different experimental setups. For example, laboratory experiments may need to consider the finite size of the sensors; whereas whole Earth studies are restricted by logistical constraints for placing sensors (e.g. it is challenging to obtain a station density in oceanic areas that is comparable to regions on land). The survey

design of active seismic experiments at exploration scales are typically governed by the requirements of reflection seismics, where it has to be ensured that the spacing satisfies the Nyquist–Shannon sampling criterion for avoiding aliasing effects. Here, FWI is employed primarily for determining velocity models that are later used for pre-stack depth migration or reverse-time imaging (e.g. Rønholt *et al.* 2014).

In near-surface seismic data sets, it is challenging to isolate reflections from other wave types, such as surface waves and guided waves (e.g. Schmelzbach *et al.* 2005). FWI is expected to be particularly beneficial at such scales, because it requires no wave-type separation and high-resolution images can be obtained. The data sets to be acquired for FWI may not necessarily have to meet the criteria dictated by seismic reflection processing methodology.

Experimental design (ED) tools offer suitable means to set up an optimal survey or to choose an optimal subset of an existing dense data set (e.g. Maurer *et al.* 2010). The methodology has successfully been applied to electromagnetic problems (e.g. Maurer & Boerner 1998), geoelectrics (e.g. Stummer *et al.* 2004; Wilkinson *et al.* 2006), earthquake network design (e.g. Hardt & Scherbaum 1994) and seismic crosshole applications (e.g. Curtis 1999a). So far, optimized ED techniques have gained only little attention in the field of FWI. Djikpesse *et al.* (2012) formulated an efficient Bayesian ED methodology in order to optimize resolution in FWI and applied it to a crosshole example. Other attempts to optimize the survey design of FWI include the work of Romdhane *et al.* (2011) who directly compare inversion results from *x*- and *z*-directed receivers and decimate the number of sources employed, while keeping a regular spacing. Brenders & Pratt (2007) investigated the influence of the minimum frequency used and of spatial subsampling. Sirgue & Pratt (2004) designed optimal frequency selection schemes. Likewise, Maurer *et al.* (2009) developed efficient frequency and spatial sampling strategies for acoustic crosshole FWI. Manukyan *et al.* (2012) investigated the information content offered by multicomponent recordings for elastic crosshole FWI problems. Similar investigations by Vigh *et al.* (2014) also highlighted the importance of multicomponent recordings for elastic marine applications.

An important question that remains to be answered is if it is really necessary for near-surface seismic FWI purposes to acquire high-fold data sets, as required for reflection processing, or if much sparser data sets are sufficient. Initial investigations for crosshole surveys indicated that a dense spatial sampling is not required for FWI problems (Maurer *et al.* 2009), but it is unclear if this conclusion can be transferred to surface-based surveys in a straightforward manner. Here, sources and receivers are only located at or just below the surface, which yields a less-constrained one-sided inversion problem. More importantly, the free surface plays a very important role, because high-amplitude surface waves are superimposing reflected phases. In order to model surface waves, the acoustic approximation is insufficient, and elastic FWI is mandatory.

In this paper, we provide an in-depth study on how to design shallow seismic surveys, optimized for elastic FWI incorporating surface waves. After introducing the theoretical background of our FWI implementation, we present the basics of optimized ED and our specific algorithm, whose performance is illustrated with a simple acoustic example. This example is used for examining the dependency of ED on the underlying subsurface model. Then, we discuss the more realistic elastic case including noise, for which we inspect the importance of numerous recording parameters, such as receiver spacing, the choice of source and receiver types and optimal placement of sources. Based on our results, we ultimately offer specific guidelines for shallow seismic survey designs. Furthermore, we highlight potential problems associated with the assumptions made in our simulations.

2 THEORY

2.1 Full waveform inversion

The aim of FWI is to find a realistic subsurface model, for which forward modelled data $\underline{u}$ can be computed that match the observed data $\underline{d}$ within the data error bounds. This is typically achieved through a linearized inversion procedure, in which the model parameters $\underline{m}$ are updated successively, until the root-mean-square (rms) misfit between $\underline{u}$ and $\underline{d}$ is minimized, that is, until convergence is achieved (e.g. Tarantola 2005). Due to the dominance of surface waves in shallow seismic data, the acoustic approximation is not justified, but elastic FWI is required. We parametrize our subsurface models with *P*- and *S*-wave speeds $\underline{V_p}$ and $\underline{V_s}$ and density $\underline{\rho}$ (i.e. $\underline{m} = [\underline{V_p}, \underline{V_s}, \underline{\rho}]$), which are discretized on a regular 2-D grid. For the sake of simplicity, we do not consider anisotropy and anelastic effects, but it is conceptually possible to extend $\underline{m}$, such that these effects can be included.

In principle, any forward solver that predicts $\underline{u(m)}$ can be employed for an FWI algorithm. Here, we consider a frequency-domain finite-element approach (Min *et al.* 2003; Latzel 2010). As outlined by Pratt (1999), frequency-domain modelling is very efficient for multiple sources, and typically the response for only a few distinct frequencies needs to be computed. In finite-element modelling, the free surface is the natural boundary condition, such that accurate modelling of surface waves is possible at no extra costs. At the top of the modelling domain, the free surface is maintained, while at the other edges of the domain, it is replaced by perfectly matched layer boundary conditions (Zheng & Huang 2002; Basu & Chopra 2003).

For the solution of the inverse problem, also carried out in the frequency domain, we follow a Gauss–Newton approach (Pratt *et al.* 1998) using

$$\underline{m}^{i+1} = \left(\mathbf{J}^T\mathbf{J} + \alpha^2\mathbf{I} + \beta^2\mathbf{L}^T\mathbf{L}\right)^{-1}\left\{\mathbf{J}^T\left[(\underline{d} - \underline{u}) + \mathbf{J}\underline{m}^i\right] + \alpha^2\mathbf{I}\underline{m}^i\right\}. \quad (1)$$

The matrix $\mathbf{J}$ contains all sensitivities, $\mathbf{I}$ is the $N \times N$ identity matrix with N being the number of model parameters contained in $\underline{m}$, $\mathbf{L}$ is the Laplacian smoothing operator and i is the iteration index ($\underline{m}^{i=0}$ is the initial model). The scalars α and β determine the weights of regularization in form of damping and smoothing, which stabilize the inversion. These weights are adjusted to the various parameter classes (V_p, V_s and ρ) in order to balance the corresponding sensitivities and model updates (Manukyan *et al.* 2012). Damping is supposed to keep the model parameters close to a given reference model. We have chosen the model parameters of the previous iteration as a reference; therefore, damping essentially controls the step length in our formalism. Prior to each model update step, source functions are estimated using the approach described by Maurer *et al.* (2012).

For large-scale inversion problems, it can be challenging to compute and store the approximate Hessian matrix $\mathbf{J}^T\mathbf{J}$. Therefore, alternative options, such, as non-linear conjugate gradients or L-BFGS methods (e.g. Nocedal & Wright 1999) have to be considered. However, for our subsequently described ED procedure the explicit computation of $\mathbf{J}^T\mathbf{J}$ is required. Since this matrix has to be established anyway, we employ the Gauss–Newton algorithm in this study. We

use the expressions by Zhou & Greenhalgh (2010) for calculating the sensitivities.

2.2 Experimental design

The goal of ED is to set up a survey or choose data from a large data set, such that benefit is optimized, while acquisition and/or computational costs are minimized. For that purpose, we first need to specify the terms 'benefit' and 'cost'. The costs of a seismic survey depend on several factors, such as accessibility, manpower and many more. We assume that the costs linearly scale with the number of sources employed. We restrict our definition of costs to be a function of the number of sources only, although placing receivers can be challenging too (e.g. three-component receivers or ocean bottom sensors). It is important to note that the methodology is sufficiently general, such that receivers could be included into the design process too.

The survey benefit can be defined via the information content offered by the data set. This can be quantified by means of measures from linear inverse theory. As discussed by Curtis (1999a), a variety of options exists to quantify the goodness of a particular data set. Here, we consider measures that are based on the approximate Hessian matrix $\mathbf{J}^T\mathbf{J}$. The reliability of the model reconstruction depends on the ability to invert $\mathbf{J}^T\mathbf{J}$ (Maurer *et al.* 2010). Without the regularization terms ($\alpha = \beta = 0$), this matrix is usually singular. The regularization terms are therefore essential, but the 'goodness' of $\mathbf{J}^T\mathbf{J}$ shall be maximized, such that the contribution of regularization is minimized. The sensitivities, contained in the Jacobian matrix $\mathbf{J}$, are governed primarily by the survey design. We can therefore maximize the 'goodness' of $\mathbf{J}^T\mathbf{J}$ by choosing appropriate source–receiver configurations. As a consequence of the non-linearity of the FWI problem, it is important to note that the sensitivities contained in $\mathbf{J}$ depend on the model parameters. As further discussed in the sequel of the paper, this needs to be considered when setting up ED.

A conceptual example is shown in Fig. 1. We assume that we have M sources available in total, from which we would like to choose a useful subset of source positions (while keeping all the receivers active). Based on an *a priori* subsurface model, we compute the sensitivities for all possible source–receiver configurations. A data set that includes all possible source–receiver configurations will be subsequently referred as a comprehensive data set Ω_M. For displaying the 'goodness' of the comprehensive matrix $\mathbf{J}^T\mathbf{J}$, we show its eigenvalue spectrum (eigenvalues sorted by their magnitude and normalized by the maximum eigenvalue; solid line in Fig. 1a). In theory, the 'goodness' is proportional to the number of non-zero eigenvalues. Due to finite numerical precision, the eigenvalues rarely equal zero. We therefore introduce a threshold, below which an eigenvalue is considered insignificant. Here, and throughout the paper, we define this threshold to be 10^{-10} times the largest eigenvalue. As shown in Fig. 1(a), the intersection of the comprehensive eigenvalue spectrum with the threshold level is at about 45 per cent of the eigenvalues. Maurer *et al.* (2009) defined this intersection to be the 'relative eigenvalue range (RER)'. It is a measure of the resolved portion of the model space. The choice of the threshold value is not critical; it only scales the RER values. We have repeated our experiments with a range of threshold values and the results were essentially identical.

For obtaining an optimized survey layout, we employ a greedy algorithm (e.g. Coles *et al.* 2015). Initially, we compute eigenvalue spectra and the corresponding RER values for subsets including data from only one source at a time. The data from the source associated with the largest RER will be chosen to form Ω_1. Then, the next source is chosen using

$$\max_{\text{sources in } \Omega_M/\Omega_k} \text{RER}_{k+1} \quad (k = 1 \ldots M-1) \tag{2}$$

with $k+1$ being the index of the next selected source. For displaying the performance of the algorithm, we construct benefit-cost curves as shown in Fig. 1(b). The horizontal cost axis ranges from 0 (no source) to 1 (all sources). The benefit, indicated on the vertical axis, is displayed by means of a normalized RER (nRER), which is defined as $\text{RER}_k/\text{RER}_M$. In this example, 80 per cent (nRER = 0.8) of the maximally resolvable portion of the model space can be resolved using only 10 per cent of all sources.

Our ED algorithm requires a large number of eigenvalue spectra to be computed. This can be computationally prohibitive for realistically sized FWI problems. Therefore, we have additionally considered an alternative goodness function. The diagonal elements of $\mathbf{J}^T\mathbf{J}$ include the squared column sums of $\mathbf{J}$. Meles *et al.* (2010) showed that the (absolute) column sums of $\mathbf{J}$ are a good proxy for the diagonal elements of the model resolution matrix (e.g. Menke 2012), which is also a measure of the 'goodness' of a particular

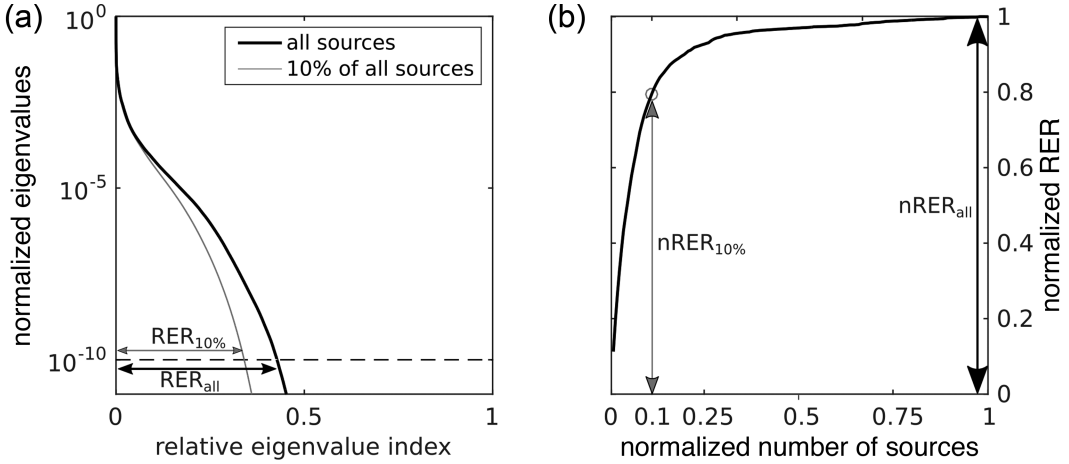

Figure 1. (a) Logarithmic eigenvalue spectra and RER values for a survey when using either 10 per cent or all (100 per cent) sources. (b) Benefit-cost curve for a seismic experiment. Costs are defined by the normalized number of sources (1 = all sources), and benefit is defined as normalized RER (nRER) (1 = benefit from all sources). The vertical double arrows specify the benefit-cost pairs related to the eigenvalue spectra in (a).

survey layout. Therefore, we can define a new measure g_k that offers similar information as the RER, but is much cheaper to compute:

$$g_k = \sum_{i=1}^{N} \frac{D_i^{\Omega_M}}{D_i^{\Omega_k} + \delta}, \tag{3}$$

with $D^{\Omega} = \mathrm{diag}(\mathbf{J}^{\mathrm{T}}\mathbf{J})$. The parameter δ is a small positive number that stabilizes the procedure in the presence of very small $D_i^{\Omega_k}$ values. Consequently, we can substitute eq. (2) by

$$\min_{\text{shots in } \Omega/\Omega_k} g_{k+1} \tag{4}$$

We have benchmarked the approximate measure in eq. (4) against the term in eq. (2) using a small data set, and found that the design results were quite similar for both measures. Therefore, we have employed eq. (4) within the optimization procedure, but we still use the corresponding *n*RER values for displaying the final benefit-cost curves, because it is a measure for the resolved model space (only a few eigenvalue decompositions need to be performed for that purpose).

3 NUMERICAL EXAMPLES

With a series of numerical examples we aim to find configurations for acquiring near-surface seismic data optimized for FWI. After introducing the general setup, we start with a simple noise-free acoustic example to investigate the model dependency of our ED results. We then move on to more realistic elastic investigations including noise, with which we seek (i) a suitable receiver spacing, (ii) optimized combinations of *x*- and *z*-directed sources and receivers and (iii) the minimal number of sources employed and their appropriate positions. Finally, we test the robustness of our findings on a different subsurface model.

3.1 Experimental setup

Fig. 2 shows the true and the initial models employed for the synthetic experiments; for the elastic case, V_s is obtained from V_p by applying a constant Poisson's ratio of 0.29, density ρ is obtained from V_p using Gardner's relation (Gardner *et al.* 1974). A 24 Hz Ricker wavelet was used for producing the synthetic data set. Seven inversion frequencies are considered accordingly: 6.4, 9.6, 12.8, 19.2, 25.6, 36.8 and 48.0 Hz, in order to cover the full wavenumber domain (Sirgue & Pratt 2004). The amplitude spectrum of the Ricker wavelet further determines the weight, with which the frequencies contribute to ED and FWI. This is favourable, because in observed data, frequencies towards both ends of the spectrum are expected to exhibit low signal-to-noise ratios. The maximum frequency $f_{\max}$ governs the minimum wavelength $\lambda_{\min}$:

$$\lambda_{\min} = \frac{V_{s,\min}}{f_{\max}} = \frac{734\,\mathrm{m\,s^{-1}}}{48\,\mathrm{Hz}} = 15.3\,\mathrm{m}. \tag{5}$$

In order to obtain good numerical accuracy, the size of the forward modelling cells is set to 1 m. The expected spatial resolution of FWI is in the order of half the minimum wavelength. Therefore, 25 forward modelling cells are merged into one inversion cell of size 5×5 m, which is still below the expected spatial resolution.

For the inversions, the frequencies are gathered into three partly overlapping groups: (i) 6.4, 9.6 and 12.8 Hz; (ii) 12.8, 19.2 and 25.6 Hz and (iii) 25.6, 36.8 and 48.0 Hz. Each inversion is started

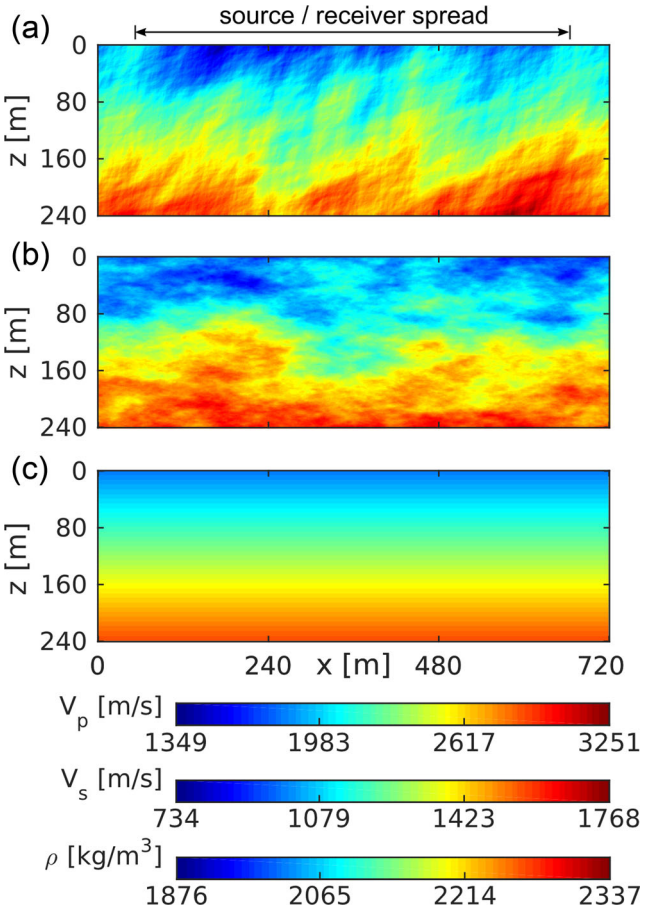

Figure 2. Models considered in the synthetic studies. V_s is obtained from V_p by applying a fixed Poisson's ratio of 0.29. Density ρ is obtained from V_p with Gardner's relation. (a) True model A; (b) true model B; (c) initial model.

with the lowest frequency group and progressed to larger frequencies as the model converges, in order to prevent cycle skipping. In our case, only six iterations per frequency group are needed until convergence is achieved, that is, until the rms deviation between $\underline{d}$ and $\underline{u}$ did not change anymore by more than 3 per cent. For consistency, all inversions are run with the same frequency schedule and the same parameters, leaving the choice of the regularization weight to a line search algorithm, which ensures that the rms is efficiently minimized (after Nocedal & Wright 1999). Two additional constraints are given: (i) the weight of damping is four times higher than the weight of smoothing ($\alpha = 2\beta$ in eq. 1), and (ii) damping of the model update in ρ is five times higher than damping in V_p and V_s, because it was observed that large model updates in ρ can make the inversion unstable. All the inversion runs converged to a comparable rms level.

3.2 Experimental design for acoustic FWI

We consider noise-free acoustic data modelled for 61 pressure receivers located at $x = 60, 70, \ldots, 660$ m at 1 m depth and 62 potential source locations placed at $x = 55, 65, \ldots, 665$ m at 1 m depth, such that they do not coincide with the receivers.

For determining an optimal experimental layout, we need to specify a subsurface model, with which the sensitivities, contained in $\mathbf{J}^{\mathrm{T}}\mathbf{J}$, can be computed. Ideally, one would employ the true model. However, the true model is unknown prior to the seismic

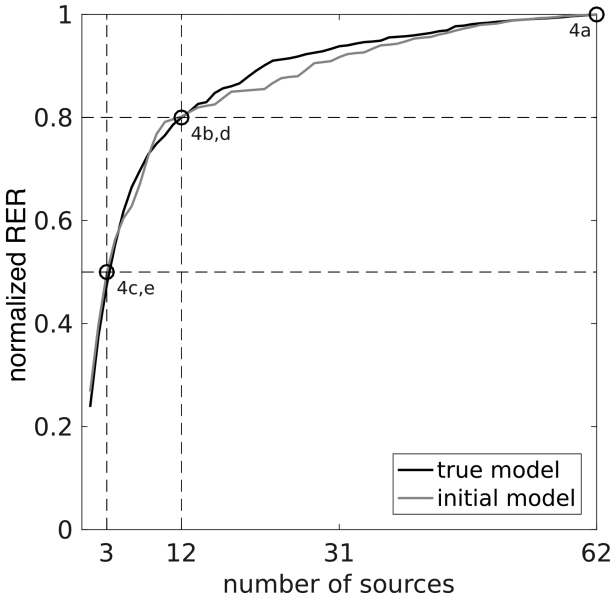

Figure 3. Benefit-cost curves for the acoustic experiment. With 3 and 12 sources, an *n*RER value of 0.5 and 0.8 could be achieved, respectively. Open dots indicate data sets for which inversion was carried out. The corresponding images are shown in the figures indicated besides the open dots.

survey. Typically, all the *a priori* information available is included in the initial model. Often, the initial model is obtained from traveltime tomography, which ensures that it is sufficiently close to the true model (e.g. Malinowski *et al.* 2011). The initial model is therefore an obvious choice for ED. However, it is unclear, if the discrepancies between using the initial or the true model could affect the design process in a negative way (due to the high non-linearity of the FWI problem). Therefore, we repeat our ED procedure, using either the initial or the true model, and compare the results.

We have mimicked the initial model with the vertical gradient model shown in Fig. 2(c). Such a gradient model is relatively easy to obtain, because it is characterized by three parameters only—V_p at the surface, the velocity gradient and a fixed Poisson's ratio. In fact, it can still be obtained from the sparse data sets acquired with optimized measurement geometry. If the traveltime tomography result shall be used as initial model, the measurement geometry shall additionally fulfill the corresponding needs. Using tools of ED, we have found that mainly the far offsets are important in this case.

The benefit-cost curves for these two scenarios are shown in Fig. 3. Both curves show a rapid increase, when only a few source positions are included. At about 12 sources, they both start to flatten out and they reach the area of diminishing returns, that is, it becomes very expensive to increase the information content of the data set. Between 12 and 30 sources, the two benefit-cost curves show some discrepancies, but overall they match very well. With both designs, it is possible to get about 50 per cent of the full information content with a mere 3 out of 62 sources, and 80 per cent of the information can be obtained with only 12 sources.

It is noteworthy that the maximum RER reached with the true model ($RER_{max} = 0.25$) is 14 per cent larger than the maximum RER reached with the initial model ($RER_{max} = 0.22$). This can be explained by the fact that the stochastic fluctuations in the true model lead to a better illumination of the subsurface, because the waves are scattered at the heterogeneities and illuminate the subsurface structures from different angles. However, the normalized curves *n*RER versus costs, as shown in Fig. 3, are comparable. For the crosshole case, the effect of scattering was further illustrated by Maurer *et al.* (2009), who plot sensitivities and the spatial distribution of the diagonal elements of the model resolution matrix for two models with various roughness.

To verify that the designs with the initial and true models are indeed similar, we compare the images from FWI of the subsets obtained from the source selections based on either the true or the initial model. Inversions are performed for the comprehensive data set including all 62 sources (Fig. 4a) and for subsets with 12 or 3 sources, chosen with ED based on either the initial (Figs 4b and c) or the true model (Figs 4d and e). For the comprehensive data set (Fig. 4a), all the important features are well reproduced, such as the shallow high-velocity anomaly around $(x, z) = (480\ \text{m}, 20\ \text{m})$ and the deep fault-like structure around $(x, z) = (450\ \text{m}, 200\ \text{m})$. With 12 sources (80 per cent benefit), the model can be resolved equally well, regardless whether the sources are chosen based on the initial model (Fig. 4b) or based on the true model (Fig. 4d). Even with three sources only (50 per cent benefit), the main features are still well recognizable, although slightly blurred (Figs 4c and e). Besides visual comparison of the inversion results, we have quantified the tomogram quality by the rms deviations between the true model and the inversion results. We have also tested the cross-correlation measure as applied by Reiser *et al.* (2017), which led to the same conclusions. For this first acoustic example, all model rms values lie within a very narrow window of about 3 per cent, indicating that all setups yield reasonable inversion results.

The source patterns are indicated with open dots above the subsurface images in Fig. 4. For the initial model, the selection follows the symmetry of the model, such that the patterns of selected sources are symmetric. First, a few sources around the centre of the model

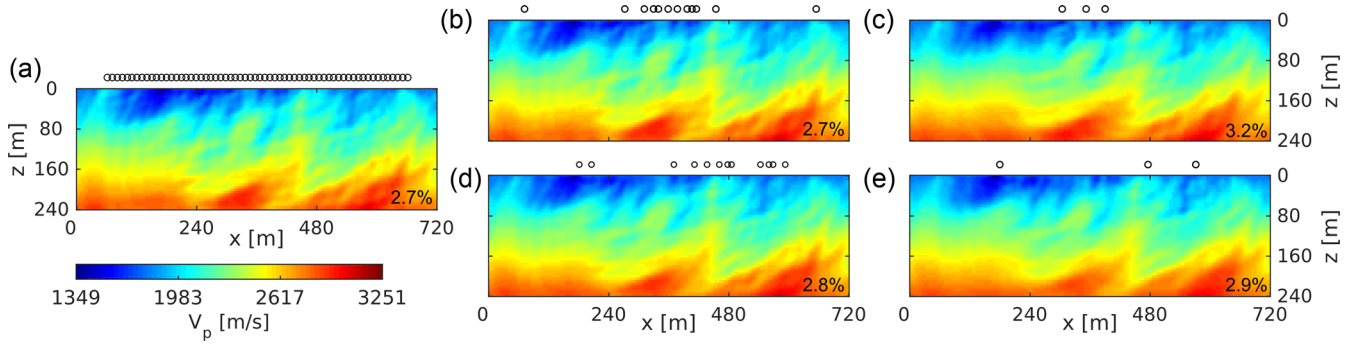

Figure 4. Acoustic FWI experiment with model A (Fig. 2a); images from (a) inverting data sets including all 62 sources, (b) 12 sources selected with the initial model and (d) the true model, and (c) for three sources selected with the initial model and (e) the true model. The open dots above the images indicate the pressure source positions employed. The rms deviations to the true model are indicated in the bottom right corners.

are selected (Fig. 4c), but soon the sources at the edges of the spread are added in (Fig. 4b), which ensures a regular illumination of the model. The symmetry disappears when selecting the sources based on the true model. Here, the source position selection is guided by the pattern of the stochastic fluctuations in the model; the sources are preferably located above the shallow high-velocity anomaly around $x = 480$ m (Fig. 4d).

From this first acoustic simulation, we can already derive a few interesting conclusions. Most importantly, it is found that only a few sources are required for constraining the V_p model very well. Secondly, ED based on the initial model does yield a highly optimized source selection, as long as the initial model is sufficiently close to the true model (which is further a pre-requisite for FWI). Although the source patterns for the two design models are different, the benefit-cost curves and the resulting images are similar.

3.3 Experimental design for elastic FWI

The acoustic example described above offered interesting insights, but for designing realistic near-surface surveys, we move on to the elastic case with free-surface boundary conditions at the top of the modelling and inversion domain. Incorporating surface waves is crucial in near-surface FWI. Furthermore, we try to make our simulations more realistic by adding considerable noise to the waveform data. In the time domain, the standard deviation of the seismograms was calculated and 30 per cent white noise relative to it was added. Finally, we also include the receiver spacing into our ED considerations. We invert for V_p, V_s and density, but we restrict the images shown here to the wave speeds. It has been previously shown that it is difficult to recover meaningful density images due to parameter trade-offs (e.g. Operto *et al.* 2013). Therefore, we have artificially damped the model update in density, such that the corresponding images stay close to the initial model.

The full data space includes the same 62 source and 61 receiver positions as already employed for the acoustic simulations. Instead of pressure sources and receivers, we consider *x*- and *z*-directed source and receiver components. For the sake of simplicity and computational efficiency, we restrict ourselves to pure 2-D problems. Therefore, no source and receiver components perpendicular to the tomographic planes are modelled. In real-data applications, out-of-plane heterogeneity will also play an important role on the quality of the model reconstruction. This aspect was studied by Butzer *et al*. (2013) and placing sensors out of plane could be part of ED, but this would be beyond the scope of this study.

In the following, we denote a particular source-receiver configuration as '*src_type-rec_type*', where *src_type* and *rec_type* can be either *x*, *z* or *xz*. For example, *z*-directed sources and multicomponent receivers with *x*- and *z*-components are denoted as '*z–xz*'. It is important to note that for the source type *xz* not necessarily both components at a particular position must be activated. Instead, *x*- and *z*-source directions can be chosen individually.

In principle, pressure (explosion) sources could also be included in the design process. In this case, since pressure sources do not produce primary shear waves, shear wave energy only stems from *P*-to-*S* converted waves (e.g. at the surface). We have experimented with pressure sources, and we found that the resulting images are often prone to artefacts because *S*-wave speeds are less constrained. We conclude that elastic near-surface FWI surveys should be performed preferably with directed sources, which produce primary shear waves; and we have thus not considered pressure sources in our elastic simulations.

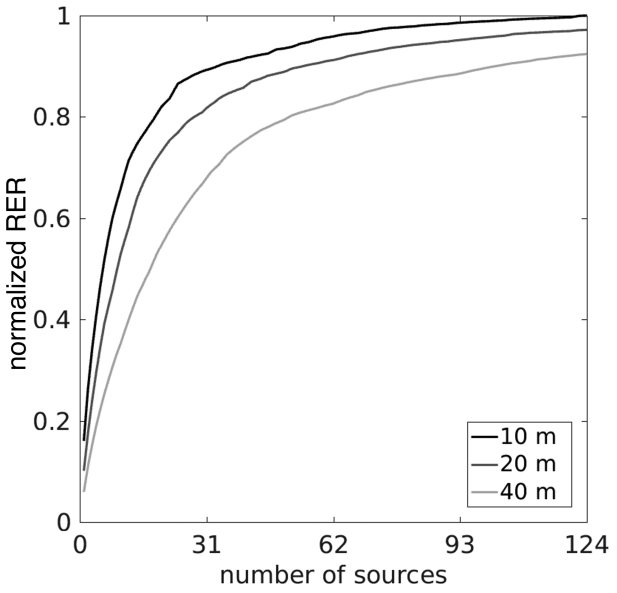

Figure 5. Benefit-cost curves for 10 m receiver spacing (black line, 61 receiver positions), 20 m spacing (dark grey line, 31 positions) and 40 m spacing (light grey line, 16 positions).

With the acoustic example, we have demonstrated that the choice of the design model is not overly critical. We have performed similar tests for the elastic case, and we found that this conclusion remains valid for elastic data as well. Therefore, we base our elastic experiments on the *a priori* known initial model shown in Fig. 2(c).

3.3.1 Suitable receiver spacing

Due to the design of receiver cables and other logistical constraints, it seems unpractical to design receiver layouts with an irregular spacing (although wireless receivers become available; e.g. Savazzi & Spagnolin 2009). Instead, we compare three receiver deployments with regular spacings. Shannon's sampling criterion dictates spatial sampling of half the minimum wavelength, which corresponds to $\Delta_{\text{samp}} = 28.7$ m for the first frequency group (up to 12.8 Hz). However, Brenders & Pratt (2007) obtain satisfactory waveform tomography images with source spacings well beyond Shannon's criterion. We therefore test three different receiver spacings: $\Delta r_1 = 10\ \text{m} = 0.35 \times \Delta_{\text{samp}}$ (61 receivers), $\Delta r_2 = 20\ \text{m} = 0.70 \times \Delta_{\text{samp}}$ (31 receivers) and $\Delta r_3 = 40\ \text{m} = 1.40 \times \Delta_{\text{samp}}$ (16 receivers).

For the experimental procedure, we consider *xz–xz* configurations, which include 124 possible sources (62 *x*-directed and 62 *z*-directed sources). The resulting benefit-cost curves are shown in Fig. 5. The curves for 10 and 20 m spacing are relatively close to each other, thereby indicating that doubling the number of receivers offers only marginal benefits (only 3 per cent for the full cost experiment including 124 sources). Using a very coarse receiver spacing of 40 m still produces good results, but the loss of information is significantly larger (8 per cent at the full cost level and up to 21 per cent in the critical zone, where the curves enter into the area of diminishing returns (around 31 sources). From these observations, we conclude that a receiver spacing of 20 m offers a reasonable compromise between acquisition efforts and survey benefits, and we have included only 31 receivers into our comprehensive data set used for the subsequent simulations. This receiver spacing furthermore conforms Shannon's sampling criterion, which is desirable

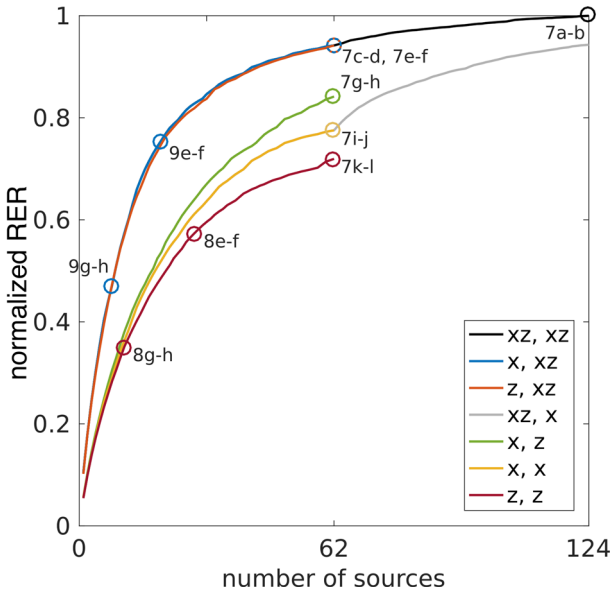

Figure 6. Benefit-cost curves for selected source–receiver combinations. Open dots refer to the figures with the corresponding FWI images (colour coding according to the legend in the bottom right corner).

when later minimizing the number of sources used. In an acoustic waveform tomography study, Brenders & Pratt (2007) have found that the image quality is satisfactory until the receiver spacing approaches Δ_{samp}, while using twice as large source spacing.

3.3.2 *Suitable source–receiver configurations*

Obviously, most subsurface information can be retrieved, when considering *xz–xz* configurations, but such a survey may be labour intense and costly. Therefore, the majority of near-surface surveys are acquired with *z–z* configurations using hammer, weight-drop or vibrator sources. With our ED computations, we make an attempt to quantify the value of the options available, and we check whether other configurations can produce similarly good results as the *xz–xz* configuration.

Fig. 6 shows the benefit-cost curves for selected configurations. The data set *xz–xz*, serving as a reference, is indicated by the black line. In Fig. 6, the *n*RER for all curves is obtained by normalizing all RER values by the maximum RER value from the comprehensive data set (i.e. all sources from the *xz–xz* configurations). This allows cross-comparison of various configurations.

As expected, best results are obtained for the comprehensive *xz–xz* data set (Figs 7a and b). The level of detail in V_s is eye-catching, and the spatial resolution is excellent; the model rms deviation to the true model accounts for 2.5 per cent only. Due to the larger *P* wavelengths, V_p is less well resolved, therefore, the rms deviation is slightly larger (3.2 per cent).

We continue our analysis by omitting more and more components. First, we omit one of the source directions (*x–xz* and *z–xz* configurations). The resulting benefit-cost curves almost coincide with those of the *xz–xz* configurations (Fig. 6). Although, *x–xz* and *z–xz* configurations include only 62 sources, a maximum *n*RER of 0.94 can be achieved. From a practical view, vertical sources (e.g. hammer blow or drop weight) are more common, because it is more difficult to design horizontal sources that transmit sufficient energy into the ground (e.g. Knödel *et al.* 2004). Therefore, *z–xz* seems to be a very favourable configuration. However, it is interesting to note that the full-cost *x–xz* images (Figs 7c and d) are slightly better than the corresponding *z–xz* images (Figs 7e and f), although the model rms deviations are equal. Minor artefacts are visible in the V_p image of the latter configuration. The *x–xz* images can hardly be distinguished from the inversion of the comprehensive data set (Figs 7a and b), but the model rms deviations are slightly larger (3.7 per cent and 2.8 per cent for V_p and V_s, respectively).

If the survey shall be restricted to one component on both the source and the receiver side, the benefit-cost curves in Fig. 6 indicate that it is most beneficial to use the *x–z* configuration (or the reciprocal configuration *z–x*). They reach an *n*RER of 0.84 when using all *x*-directed sources and *z*-directed receivers. In contrast, the *x–x* configurations achieve only a maximum *n*RER of 0.78 and the *z–z* configurations provide the lowest value (0.72). These results are partially confirmed by the corresponding FWI images. For the *x–z* case, a high-quality image for V_s can be obtained, but the inversion is prone to artefacts, again mainly in V_p (Fig. 7g). FWI of *x–x* components still yields a satisfactory image for V_s but the spatial resolution in V_p at greater depths is poor (Figs 7i and j). The opposite is the case for *z–z* (Figs 7k and l): A relatively good image for V_p can be obtained, but the V_s image appears rather blurred, mainly at greater depths. The comparison of the model rms deviations shows that *z–z* yields the best image for V_p, while *x–z* yields the best image for V_s in case of using one source and one receiver component.

3.3.3 *Optimizing the number of sources*

First, we consider the most commonly (and cheapest) employed configurations *z–z*. For that purpose, we extract data sets offering 80 per cent and 50 per cent of the information relative to the corresponding full cost experiments (Fig. 6).

For the *z–z* case, 28 sources (45 per cent) are required to reach the 80 per cent benefit level and 11 sources (18 per cent) to reach the 50 per cent benefit level (Fig. 6), respectively. The FWI images at the 80 per cent level (Figs 8e and f) are only slightly degraded compared with the *z–z* full-cost experiment (Figs 8c and d). At the 50 per cent level, the corresponding images are degraded significantly (Figs 8g and h), which also manifests itself in a strong increase of the rms values.

Next, we repeat the analysis with our preferred *x–xz* configuration. Only 20 sources (32 per cent) are needed for obtaining 80 per cent benefit (Fig. 6). This reduction affects the reconstruction of V_p, but the image obtained for V_s is still good (Figs 9e and f). With only eight sources (13 per cent) the 50 per cent benefit can be reached and the quality of the V_s image is still good (Figs 9g and h). The image for V_p is somewhat blurred, especially at greater depths.

The source patterns are displayed on top of the corresponding images in Figs 8 and 9. For all simulations, the algorithm initially selects more or less regularly distributed source locations, but at a later stage, the sources start clustering in the central parts of the profile. It is noteworthy that the outermost source positions do not seem to be of particularly high importance, and they are not selected for reaching the 50 per cent benefit (Figs 8 and 9g and h). Interestingly, the opposite is observed, when similar ED computations are performed for refraction traveltime tomography experiments. In that case, the algorithm picks preferably source locations towards both ends of the profile for obtaining sufficient depth penetration.

The source patterns selected with ED are irregular. Nevertheless, a regularly spaced source deployment may offer logistical advantages. We have tested this option, and it turned out that 50 per cent

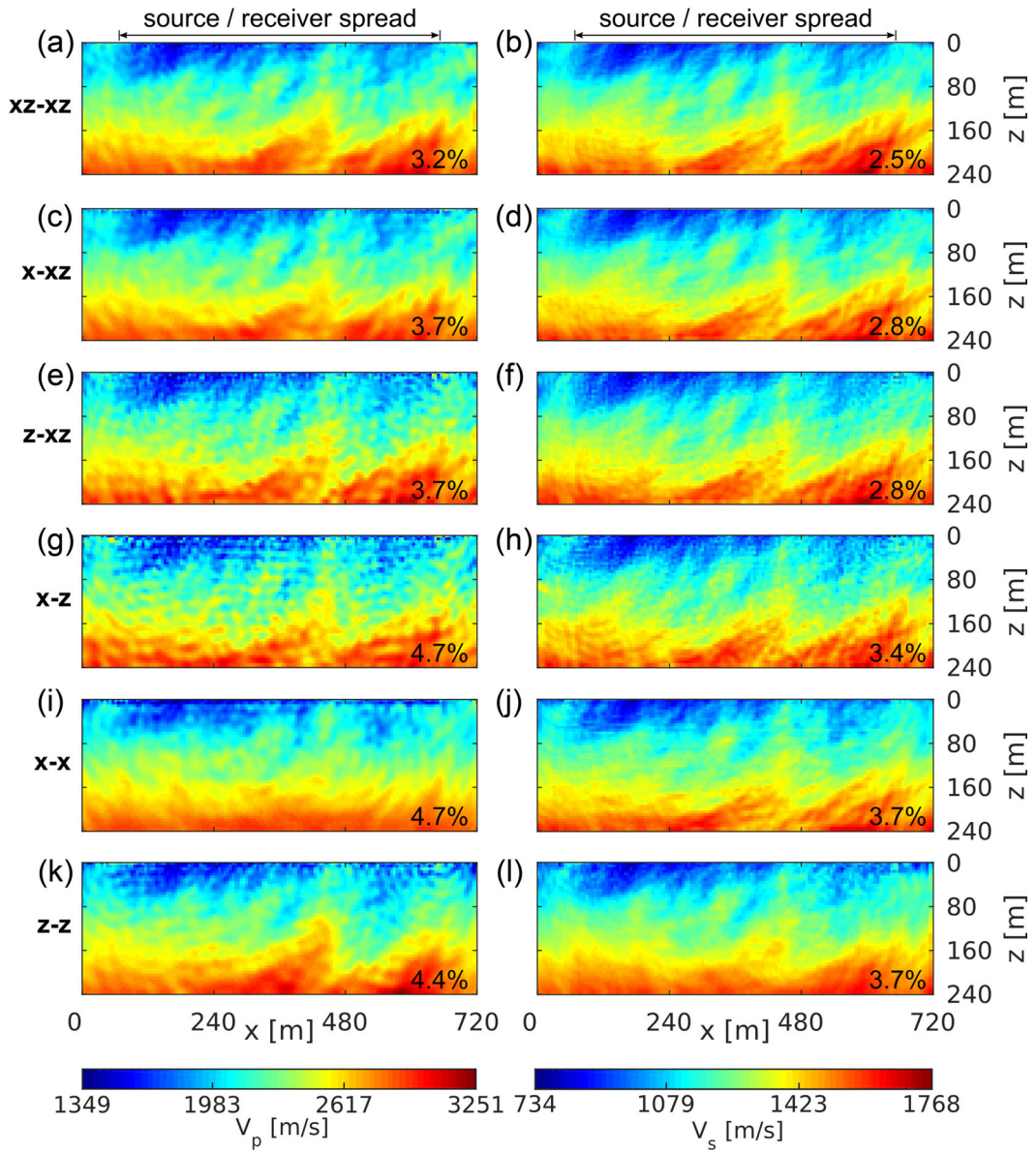

Figure 7. Model A: FWI images (left: V_p and right: V_s) for selected source–receiver combinations (see labels at the left). All source positions were considered for computing the images. The rms deviations to the true model are indicated in the bottom right corners.

benefit can be reached with the same number of regular spaced sources as with the sources selected by the ED algorithm. This indicates that the ED problem includes ambiguities, that is, there may exist several source configurations that lead to similar benefits. Choosing 11 (*z–z* case) or 8 (*x–xz* case) regularly distributed sources at once seems as optimal (in terms of benefit) as subsequently selecting them with the greedy algorithm. We have repeated the simulations, shown in Figs 8 and 9(g) and (h), with regular source spacing. In the *z–z* case, choosing regular source spacing slightly decreases the rms values, but the images are still prone to artefacts (Figs 8i and j). In the *x–xz* case, the rms and the artefacts are slightly increased for regular source spacing, indicating that the source selection from ED is superior to the regular spacing.

3.3.4 Checking the model dependency

Since our ED is based on the initial model, the results should be valid for the inversion of any other models, for which this initial model is appropriate. We test this hypothesis by repeating the inversions experiments, shown in Fig. 9, with model B (Fig. 2b).

The images obtained with the comprehensive data set (Figs 10a and b) and the increased rms values indicate that model B is more challenging to recover. We observe a few artefacts around the source positions, mainly visible in V_p. Results obtained with the *x–xz* configurations (Figs 10c and d) are, however, very comparable to those in Figs 10(a) and (b), just like it is the case for model A. With 20 sources (80 per cent benefit; Figs 10e and f), the model for V_p appears rather blurred and it is difficult to distinguish artefacts from structure. The reconstruction of V_s is still very good; it only degrades significantly, when employing not more than eight sources (50 per cent benefit; Figs 10g and h). The decrease of image quality when omitting sources, are comparable to model A; thereby indicating a fairly general validity of our findings. Likewise, the source patterns are comparable to those observed in Figs 8 and 9, and the conclusions with regard to optimized versus regular source patterns also apply to model B (Figs 8–10g–j).

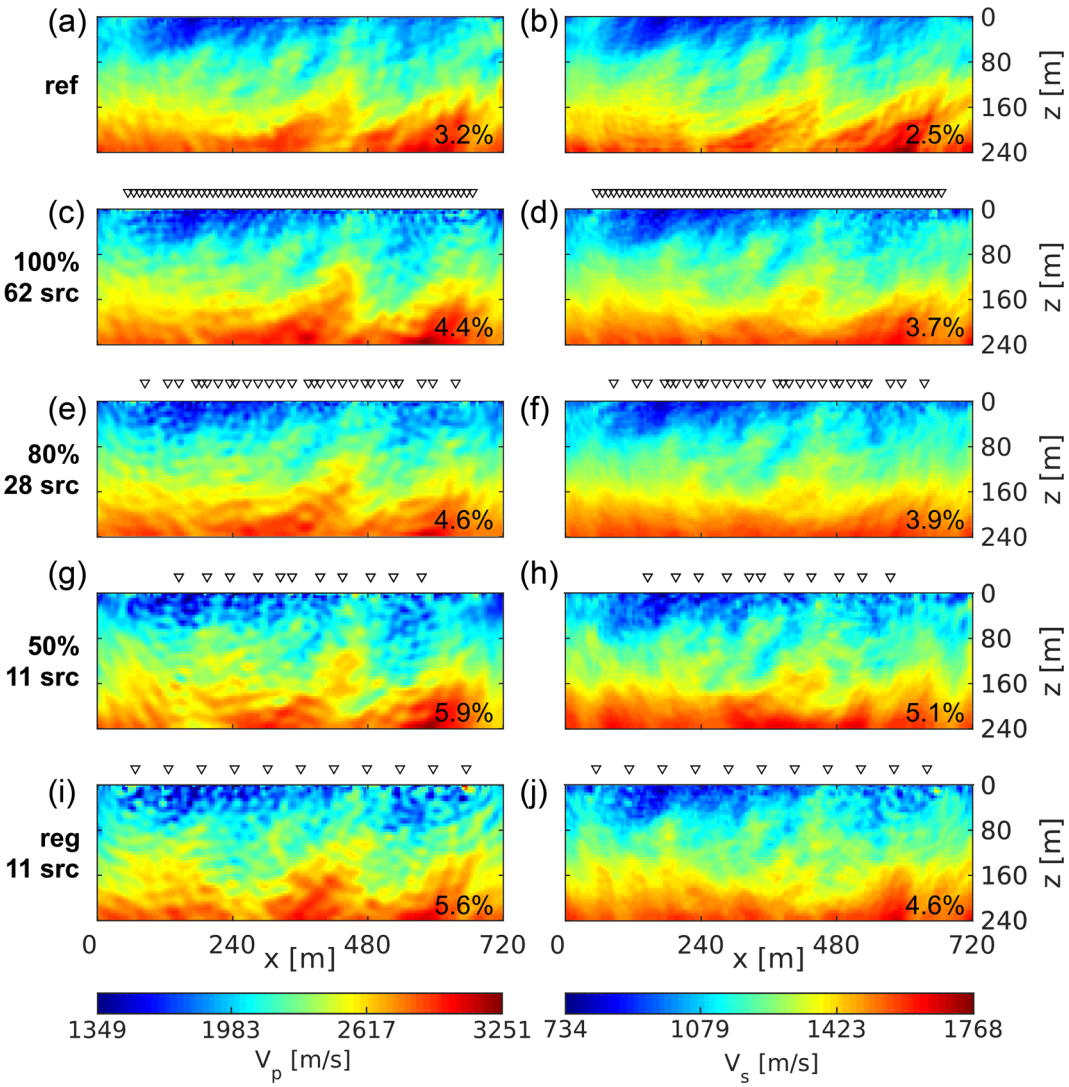

Figure 8. Model A: FWI images (left: V_p and right: V_s) for optimized z–z configurations. On the left-hand side, the number of sources employed and the corresponding percentage of the full-cost experiment are displayed (see also Fig. 6). Selected z-directed sources are indicated above the images with downward pointing triangles. The uppermost panels (a) and (b) show inversion results of the comprehensive data set (xz–xz) as a reference. The lowermost panels (i) and (j) display the tomographic results obtained with the same number of sources as in (g) and (h) but with regular spacing. The rms deviations to the true model are indicated in the bottom right corners.

4 DISCUSSION

The simulations presented above demonstrated the usefulness of our ED approach, but we also would like to emphasize a few challenges related to the approach. Probably the most severe issue of the FWI method in general, is its strong non-linearity. This problem is not alleviated by our approach. We make the assumption that the choice of the initial model, the frequency schedule and the regularization scheme allow a full exploitation of the data information content, that is, convergence to the global minimum can be achieved. This is often, but not always the case. For example, there are artefacts observed in some FWI images that were computed with data sets exhibiting a relatively high *n*RER value (e.g. results for z–xz configurations in Figs 7e and f). The results for these inversions may be improved by fine-tuning the regularization parameters. However, for the sake of consistency, we kept the inversion strategy unchanged throughout all inversions, and we left the choice of the regularization weight to our line search algorithm. Addressing the non-linearity is an issue that must be always considered in FWI problems. Our ED approach does not contribute directly to the solution of this problem, but it offers improved subsurface information, when the non-linearity problem has been addressed adequately. A possible option to address the problem is non-linear ED, where an optimized experimental setup is determined that is suitable for a range of likely subsurface models (Maurer *et al*. 2010).

A potential limitation of our ED algorithm is imposed by the simplified goodness function (eq. 3). This is demonstrated in Fig. 6 by means of the rarely used xz–x configuration (grey curve). In order to minimize the alternative goodness functional (eq. 3), it seems beneficial to select all x-directed sources (i.e. overlap with the x–x curve until source 62) prior to selecting the first z-directed source. The kink and the sudden increase of the slope in the xz–x curve beyond source 62 indicate that it would likely be beneficial to select z-directed sources at an earlier stage for maximizing the RER. This problem is caused by the generally larger magnitudes of the sensitivities of x-directed sources compared with z-directed sources. The simplified goodness function in eq. (3) is more strongly

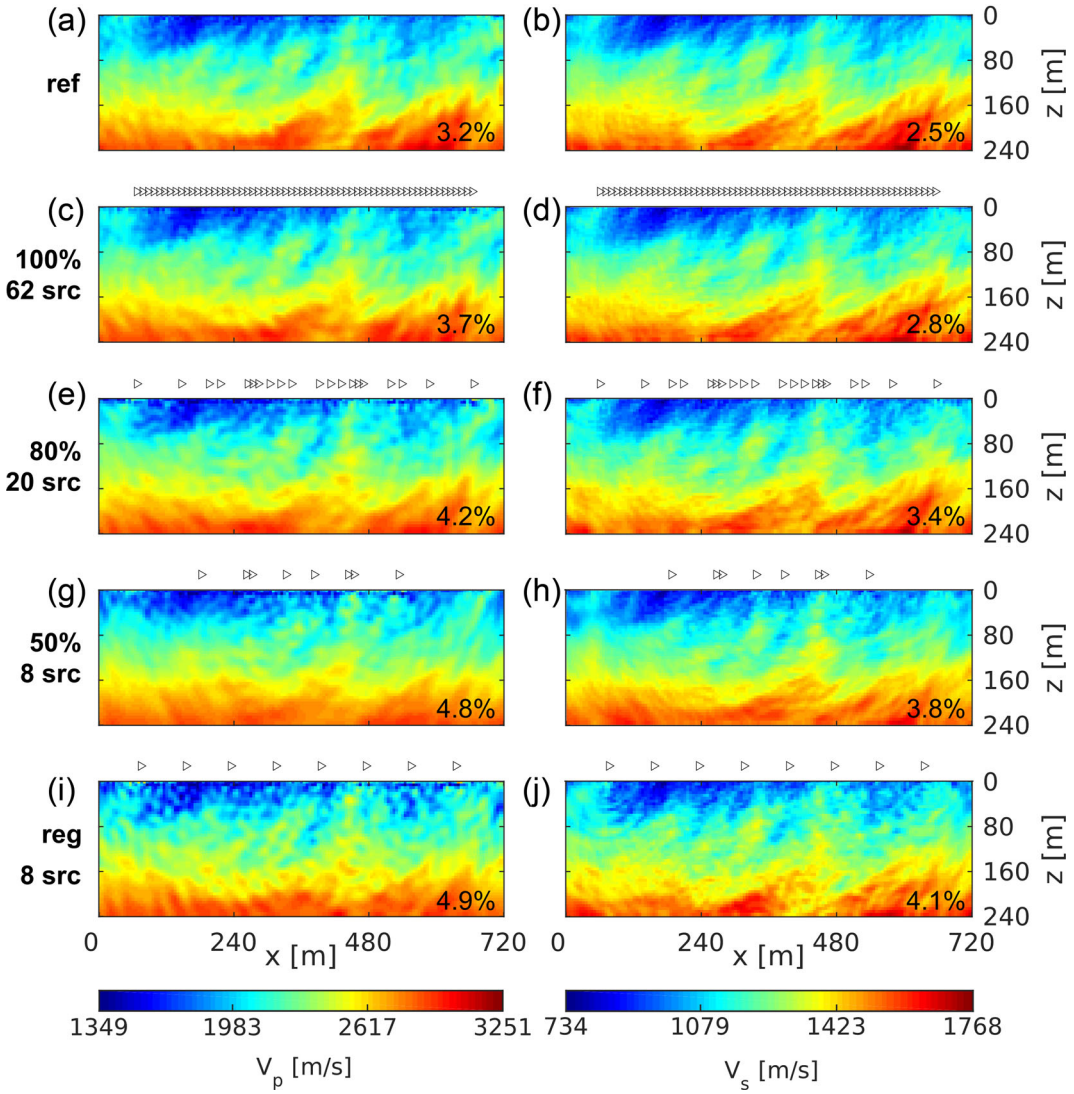

Figure 9. Model A: FWI images (left: V_p and right: V_s) for optimized x–xz configurations. On the left-hand side, the number of sources employed and the corresponding percentage of the full-cost experiment are displayed (see also Fig. 6). Selected x-directed source positions are indicated above the images with triangles pointing to the right. The uppermost panels (a) and (b) show inversion results of the comprehensive data set (xz–xz) as a reference. The lowermost panels (i) and (j) display the tomographic results obtained with the same number of sources as in (g) and (h), but with regular spacing. The rms deviations to the true model are indicated in the bottom right corners.

governed by the sensitivity magnitudes than the eigenvalues related with the RER measure. Interestingly, this problem does not seem to affect the xz–xz curve, and it is of course not an issue when only a single source component is available.

Further limitations can be imposed by data pre-processing that may be necessary for obtaining stable inversion results. Pre-processing typically includes frequency filtering and/or extracting selected portions of the seismograms. Both options will affect the information content offered by the comprehensive data set and may thus have an impact on the ED results.

Frequency filtering is related to an optimal choice of frequencies for FWI experiments. This topic has been discussed by Sirgue & Pratt (2004) and Maurer *et al*. (2009). Likewise, selecting suitable time windows can be also a topic of ED, and first attempts have been presented by Bernauer *et al*. (2014). In this contribution, we have assumed a prescribed frequency range computed with the full seismograms. It will be a topic of future research to combine the identification of optimized source–receiver combinations with the choice of suitable frequencies and/or time windows.

Besides a greedy algorithm, which chooses sources one after another, a global algorithm, which chooses the desired number of sources at once, could yield superior source distributions. This is most obvious when looking at the extreme case of using two sources only. The greedy algorithm chooses the first source preferably around the middle of the profile and the second source such that it complements the first source. In contrast, we expect a global algorithm to choose two sources that are distributed more regularly. However, global algorithms are still prohibitively expensive for such studies, even when using simplified goodness functions.

Goodness measures that are based on eigenvalue spectra, such as the RER, are a valuable option, because they allow straightforward quantification of the resolved model space and the null space. A potential problem is that such measures provide no information on particular model parameters. This is, for example, visible in Fig. 7: The V_p image obtained with z–z configurations is superior to most

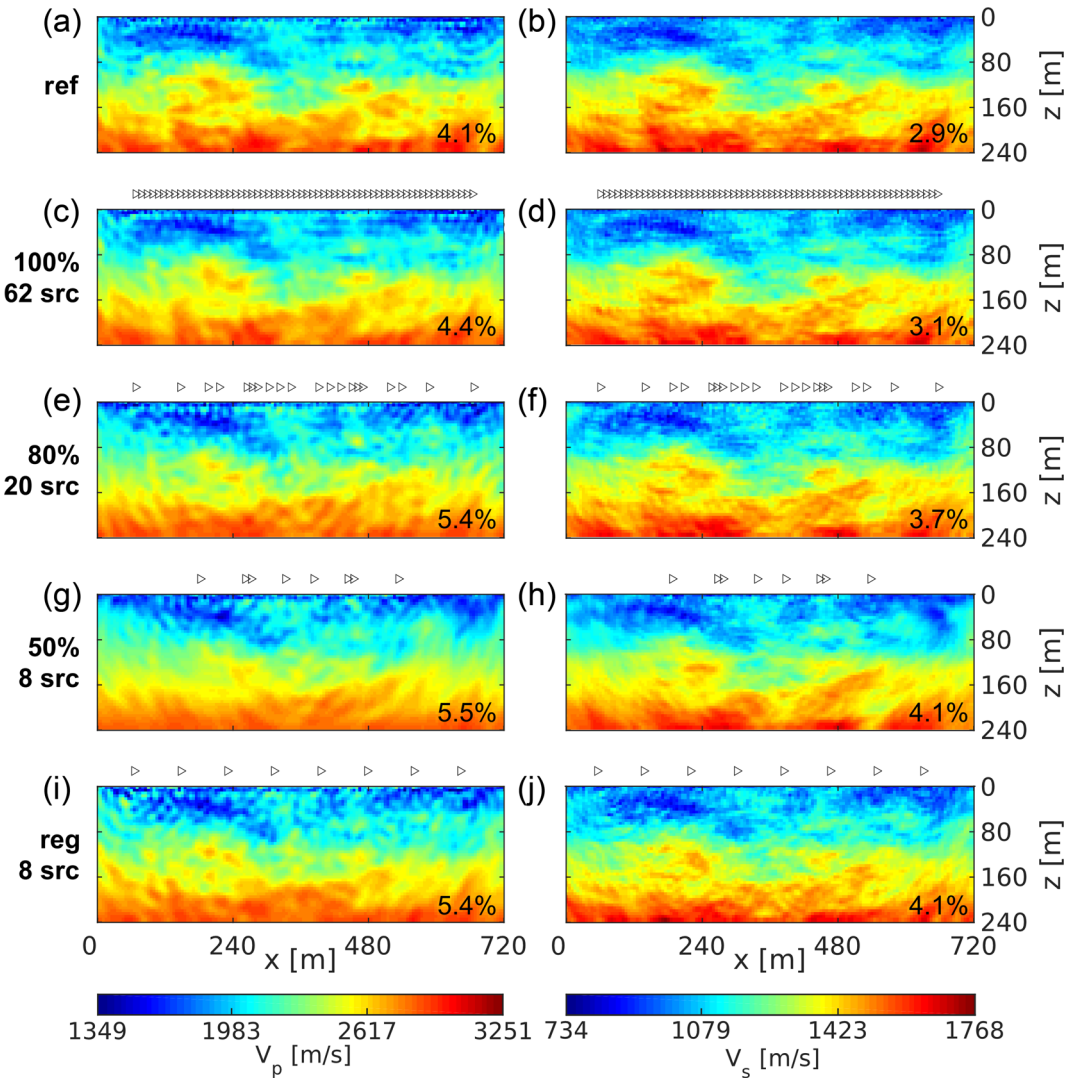

Figure 10. Model B: FWI images (left: V_p and right: V_s) for optimized x–xz configurations. On the left-hand side, the number of sources employed and the corresponding percentage of the full-cost experiment are displayed (see also Fig. 6). Selected x-directed sources positions are indicated above the images with triangles pointing to the right. The uppermost panels (a) and (b) show inversion results of the comprehensive data set (xz–xz) as a reference. The lowermost panels (i) and (j) display the tomographic results obtained with the same number of sources as in (g) and (h), but with regular spacing. The rms deviations to the true model are indicated in the bottom right corners.

other V_p images computed with data sets related with larger nRER values. The poor nRER score of the z–z configuration is caused by its inability to produce good V_s images (Fig. 7l).

A possible extension of our current methodology is to include goodness functions that maximize the resolution of a particular parameter type (V_p, V_s or ρ). One could even focus on certain subsurface regions that are of special interest. This is the topic of focused ED, and can be achieved with benefit measures that are based on the model resolution matrix (e.g. Curtis 1999b; Wagner *et al.* 2015). In fact, our simplified goodness function, defined in eq. (3), could also be employed for that purpose by summing over the sensitivities with respect to the model parameters of interest only.

The benefit-cost curves in Figs 3, 5 and 6 exhibit a relatively smooth transition into the area of diminishing returns, and it is difficult to identify a clear kink in the curves, where the benefit-cost ratio is optimal. Therefore, we discuss the choice of a suitable data set using the percentages relative to the full cost experiments. The noise-free acoustic example indicates that choosing only a small number of sources, that is, 50 per cent of the full cost experiment or even fewer sources, still leads to very good results. We have also experimented with noise-free elastic data and came to similar conclusions. However, the results obtained with noise-contaminated elastic data indicate that acquiring data sets at a level of at least 80 per cent of the corresponding full cost experiments is advisable, depending on the extent of the full cost experiment.

For the sake of simplicity, we have considered only white noise in our simulations, but in a realistic scenario the presence of coloured noise must be expected. This survey-specific problem may be partially alleviated with an increased data redundancy. Therefore, it seems advisable to acquire slightly more data than indicated by the ED procedure.

Our results will be most beneficial for designing shallow seismic experiments, where the tomograms constitute the final product. Seismic surveys performed on hydrocarbon exploration scales are typically designed such that the data sets produce optimal pre-stack depth migration or reverse-time migration images. Therefore, one may conclude that our ED strategy is of lesser importance for such

surveys. However, it is nowadays standard practice that FWI is employed for establishing velocity models that serve as input for the subsequently applied migration algorithms. Even with substantial computing resources, it is still challenging to apply FWI to large and densely sampled data sets. Therefore, strategies have been developed for alleviating the computational costs. For example, source encoding techniques were devised, with which several sources can be simulated simultaneously (e.g. Krebs *et al.* 2009). Alternatively, posteriori ED could be applied, that is, our ED strategy could be employed for selecting a small subset out of a large data volume for performing FWI.

5 CONCLUSIONS

We have employed optimized ED techniques for delineating useful source–receiver configurations that are amenable for acoustic and elastic FWI. It was found that it is not necessary to employ a dense spatial sampling and to achieve a high fold, as it is dictated by Shannon's sampling criterion and required for seismic reflection processing techniques. Instead, a relatively coarse spatial sampling is sufficient for obtaining detailed FWI images. More specifically, we suggest designing shallow seismic surveys on the basis of the following considerations:

(1) We recommend horizontal x-directed sources. In combination with multicomponent receivers, they offer similar information content as multicomponent sources (i.e. x- and z-directed sources). Using only z-directed sources (which may be easier to implement in practise) is another reasonable strategy.

(2) For obtaining high-resolution V_s images, multicomponent geophones clearly outperform single-component receivers. Although the former are more expensive, we recommend using such devices, whenever possible.

(3) When only single-component sources and receivers are available, it is recommended to use z-directed sources and x-directed receivers or vice versa.

(4) A receiver spacing of the order of the minimum shear wavelength is judged to be sufficient.

(5) The optimal number of sources and receivers and their locations can be determined with benefit-cost curves, as shown in Figs 3, 5 and 6. The resulting source patterns generally exhibit a denser spacing in the central parts of the profile, but regularly spaced sources offer similarly good results.

The number of sources used can be optimized with our ED algorithm. However, for real data one should not overly tweak the number of sources. With high levels of correlated and uncorrelated noise and/or unexpected earth features, additional sources can increase the stability of the inversion. Likewise, all the other limitations of our ED algorithm, mentioned in Section 4, must be considered during the ED process.

Our ED methodology is sufficiently general, such that it can be adapted to any type of seismic survey (large-scale seismic surveys, crosshole applications, vertical seismic profiling, anisotropy, etc.). For example, for including anisotropy, it will be necessary to consider more subsurface parameters compared with isotropic problems, and it has to be made sure that all the parameters can be resolved independently. The corresponding survey layouts will likely differ from those derived for isotropic cases, but the eigenvalue and RER analysis, described in this paper, will be essentially identical. Methodological FWI developments are progressing rapidly, but so far, only little efforts have been made to establish optimized input data sets. Our approach can fill this gap, and make FWI attractive to a broader range of applications.

ACKNOWLEDGEMENTS

We thank Cedric Schmelzbach and Marlies Vasmel, as well as two anonymous reviewers and the editor for their constructive feedback. Furthermore, we thank the members of the 'CARNEVAL' consortium, Nagra, OMV and Schlumberger Gould Research, as well as the Swiss National Science Foundation (grant no. 200020_121707) for financial support.

REFERENCES

Basu, U. & Chopra, A.K., 2003. Perfectly matched layers for time-harmonic elastodynamics of unbounded domains: theory and finite-element implementation, *Comput. Methods Appl. Mech. Eng.*, **192**, 1337–1375.

Bernauer, M., Fichtner, A. & Igel, H., 2014. Optimal observables for multiparameter seismic tomography, *Geophys. J. Int.*, **198**(2), 1241–1254.

Bleibinhaus, F. & Hilberg, S., 2012. Shape and structure of the Salzach Valley, Austria, from seismic traveltime tomography and full waveform inversion, *Geophys. J. Int.*, **189**, 1701–1716.

Brenders, A.J. & Pratt, R.G., 2007. Efficient waveform tomography for lithospheric imaging: implications for realistic, two-dimensional acquisition geometries and low-frequency data, *Geophys. J. Int.*, **168**, 152–170.

Bretaudeau, F., Brossier, R., Leparoux, D., Abraham, O. & Virieux, J., 2013. 2D elastic full-waveform imaging of the near-surface: application to synthetic and physical modelling data sets, *Near Surf. Geophys.*, **11**(3), 307–316.

Buske, S., Lecomte, I., Nemeth, T., Operto, S. & Sallares, V., 2009. Imaging and inversion—introduction, *Geophysics*, **74**, WCA1–WCA4.

Butzer, S., Kurzmann, A. & Bohlen, T., 2013. 3D elastic full-waveform inversion of small-scale heterogeneities in transmission geometry, *Geophys. Prospect.*, **61**, 1238–1251.

Coles, D.A., Prange, M.D. & Djikpesse, H.A., 2015. Optimal survey design for big data, *Geophysics*, **80**, P11–P22.

Curtis, A., 1999a. Optimal experiment design: Cross-borehole tomographic examples, *Geophys. J. Int.*, **136**, 637–650.

Curtis, A., 1999b. Optimal design of focused experiments and surveys, *Geophys. J. Int.*, **139**, 205–215.

Djikpesse, H.A., Khodja, M.R., Prange, M.D., Duchenne, S. & Menkiti, H., 2012. Bayesian survey design to optimize resolution in waveform inversion, *Geophysics*, **77**, R81–R93.

Fichtner, A., 2011. *Full Seismic Waveform Modelling and Inversion*. Springer.

Fichtner, A., Saygin, E., Taymaz, T., Cupillard, P., Capdeville, Y. & Trampert, J., 2013. The deep structure of the North Anatolian Fault Zone, *Earth planet. Sci. Lett.*, **373**, 109–117.

French, S.W. & Romanowicz, B.A., 2014. Whole-mantle radially anisotropic shear velocity structure from spectral-element waveform tomography, *Geophys. J. Int.*, **199**, 1303–1327.

Gardner, G.H.F., Gardner, L.W. & Gregory, A.R., 1974. Formation velocity and density—the diagnostic basics for stratigraphic traps, *Geophysics*, **39**, 770–780.

Hardt, M. & Scherbaum, F., 1994. The design of optimum networks for aftershock recordings, *Geophys. J. Int.*, **117**, 716–726.

Jaiswal, P., Zelt, C.A., Dasgupta, R. & Nath, K.K., 2009. Seismic imaging of the Naga thrust using multiscale waveform inversion, *Geophysics*, **74**, WCC129–WCC140.

Knödel, K., Krummel, H. & Lange, G., 2004. *Handbuch zur Erkundung des Untergrundes von Deponien*, Band 3, Geophysik, Springer.

Krebs, J.R., Anderson, J.E., Hinkley, D., Baumstein, A., Lee, S., Neelamani, R. & Lacasse, M.-D., 2009. Fast full wave seismic inversion using source encoding, in *SEG Technical Program Expanded Abstracts 2009*, pp. 2273–2277, chairperson Castagna, J., Society of Exploration Geophysicists.

Lailly, P., 1983. The seismic inverse problem as a sequence of before-stack migrations, *Conference of Inverse Scattering: Theory and Application*, eds Bednar, J.B., Robinson, R.R.E. & Weglein, A., SIAM.

Latzel, S., 2010. 2.5D elastic modelling and sensitivity investigations directed towards near-surface seismic waveform analysis, *PhD thesis*, ETH Zurich, Zurich, Switzerland.

Malinowski, M., Operto, S. & Ribodetti, A., 2011. High-resolution seismic attenuation imaging from wide-aperture onshore data by visco-acoustic frequency-domain full-waveform inversion, *Geophys. J. Int.*, **186**, 1179–1204.

Manukyan, E., Latzel, S., Maurer, H., Marelli, S. & Greenhalgh, S.A., 2012. Exploitation of data-information content in elastic-waveform inversions, *Geophysics*, **77**, R105–R115.

Maurer, H. & Boerner, D.E., 1998. Optimized and robust experimental design: a non-linear application to EM sounding, *Geophys. J. Int.*, **132**, 458–468.

Maurer, H., Curtis, A. & Boerner, D.E., 2010. Recent advances in optimized geophysical survey design, *Geophysics*, **75**, 177–194.

Maurer, H., Greenhalgh, S.A. & Latzel, S., 2009. Frequency and spatial sampling strategies for crosshole seismic waveform spectral inversion experiments, *Geophysics*, **74**, WCC79–WCC89.

Maurer, H., Greenhalgh, S.A., Manukyan, E., Marelli, S. & Green, A.G., 2012. Receiver-coupling effects in seismic waveform inversions, *Geophysics*, **77**, R57–R63.

Meles, G.A., Van der Kruk, J., Greenhalgh, S.A., Ernst, J.R., Maurer, H. & Green, A.G., 2010. A new vector waveform inversion algorithm for simultaneous updating of conductivity and permittivity parameters from combination crosshole/borehole-to-surface GPR data, *IEEE Trans. Geosci. Remote Sens.*, **48**, 3391–3407.

Menke, W., 2012. *Geophysical Data Analysis: Discrete Inverse Theory*, Elsevier/Academic Press.

Min, D.-J., Shin, C., Pratt, R.G. & Yoo, H.S., 2003. Weighted-averaging finite-element method for 2D elastic wave equations in the frequency domain, *Bull. seism. Soc. Am.*, **93**, 904–921.

Mora, P., 1987. Nonlinear two-dimensional elastic inversion of multioffset seismic data, *Geophysics*, **52**, 1211–1228.

Nocedal, J. & Wright, S.J., 1999. *Numerical Optimization*, Springer.

Operto, S., Gholami, Y., Prieux, V., Ribodetti, A., Brossier, R., Métivier, L. & Virieux, J., 2013. A guided tour of multiparameter full-waveform inversion with multicomponent data: from theory to practise, *Leading Edge*, **32**, 1040–1054.

Operto, S., Ravaut, C., Improta, L., Virieux, J., Herrero, A. & Dell'Aversana, P., 2004. Quantitative imaging of complex structures from dense wide-aperture seismic data by multiscale traveltime and waveform inversions: a case study, *Geophys. Prospect.*, **52**, 625–651.

Plessix, R.-É., 2008. Introduction: towards a full waveform inversion, *Geophys. Prospect.*, **56**, 761–763.

Pratt, R.G., 1999. Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model, *Geophysics*, **64**, 888–901.

Pratt, R.G., Shin, C. & Hicks, G.J., 1998. Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion, *Geophys. J. Int.*, **133**, 341–362.

Raknes, E.B., Arntsen, B. & Weibull, W., 2015. Three-dimensional elastic full waveform inversion using seismic data from the Sleipner area, *Geophys. J. Int.*, **202**, 1877–1894.

Reiser, F., Schmelzbach, C., Maurer, H., Greenhalgh, S. & Hellwig, O., 2017. Optimizing the design of vertical seismic profiling (VSP) for imaging fracture zones over hardrock basement geothermal environments, *J. appl. Geophys.*, **139**, 25–35.

Romdhane, A., Grandjean, G., Brossier, R. & Virieux, J., 2011. Shallow-structure characterization by 2D elastic full-waveform inversion, *Geophysics*, **76**, R81–R93.

Rønholt, G. *et al.*, 2014. High-fidelity complete wavefield velocity model building and imaging in shallow water environments—a North Sea case study, *First Break*, **32**, 127–131.

Savazzi, S. & Spagnolin, U., 2009. Synchronous ultra-wide band wireless sensors networks for oil and gas exploration, in *2009 IEEE Symposium on Computers and Communications*, pp. 907–912, IEEE.

Schmelzbach, C., Green, A.G. & Horstmeyer, H., 2005. Ultra-shallow seismic reflection imaging in a region characterized by high source-generated noise, *Near Surf. Geophys.*, **3**(1), 33–46.

Sirgue, L. & Pratt, R.G., 2004. Efficient waveform inversion and imaging: a strategy for selecting temporal frequencies, *Geophysics*, **69**, 231–248.

Smithyman, B.R., Pratt, R.G., Hayles, J. & Wittebolle, R., 2009. Detecting near-surface objects with seismic waveform tomography, *Geophysics*, **74**, WCC119–WCC127.

Stummer, P., Maurer, H. & Green, A.G., 2004. Experimental design: electrical resistivity data sets that provide optimum subsurface information, *Geophysics*, **69**, 120–139.

Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics*, **49**, 1259–1266.

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

Vigh, D., Jiao, K., Watts, D. & Sun, D., 2014. Elastic full waveform inversion application using multicompoment measurements of seismic data collection, *Geophysics*, **79**, R63–R77.

Virieux, J. & Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics, *Geophysics*, **74**, WCC127–WCC152.

Wagner, F.M., Günther, T., Schmidt-Hattenberger, C. & Maurer, H., 2015. Constructive optimization of electrode locations for target-focused resistivity monitoring, *Geophysics*, **80**, E29–E40.

Wilkinson, P.B., Meldrum, P.I., Chambers, J.E., Kuras, O. & Ogilvy, R.D., 2006. Improved strategies for the automatic selection of optimized sets of electrical resistivity tomography measurement configurations, *Geophys. J. Int.*, **167**, 1119–1126.

Wu, R.-S. & Toksöz, M.N., 1987. Diffraction tomography and multisource holography applied to seismic imaging, *Geophysics*, **52**, 11–25.

Zheng, Y. & Huang, X., 2002. Anisotropic perfectly matched layers for elastic waves in cartesian and curvilinear coordinates anisotropic PML for elastic waves, Massachusetts Institute of Technology, Cambridge, MA. Available at: http://dspace.mit.edu/bitstream/handle/1721.1/67857/ZHENG.PDF?sequence=1.

Zhou, B. & Greenhalgh, S.A., 2010. Computing the sensitivity kernels for 2.5-D seismic waveform inversion in heterogeneous, anisotropic media, *Pure appl. Geophys.*, **168**, 1729–1748.