Seismic characterization of waste disposals by full waveform inversion: possibilities and challenges

Thomas Bohlen, Yudi Pan, Nikolaos Athanasopoulos, Lingli Gao

August 21, 2020

1 Introduction

Geophysical characterisation of potential geological sites for nuclear waste disposal is important to develop a repository design adequate to ensure safety and waste isolation. The specific site properties are generally very heterogeneous. Ground water flow, water chemistry, geologic structure as well as thermal and structural stability are very much site dependent. Noninvasive geophysical methods can provide important in-situ information related to these key components of site characterization. However, different geophysical methods have varying responses and resolution. Therefore, a combination of different geophysical methods calibrated by in-situ borehole information and laboratory results is generally desirable.

At the expense of high efforts in both acquisition and processing, reflection seismic methods yield perhaps the highest resolution and lowest ambiguities in imaging geologic structures. The main product of reflection seismics are detailed images of seismic reflectivity which display the location of geological discontinuities. These seismic images contribute significantly to the structural and tectonic interpretation and the overall characterization of the disposal. The petrophysical characterization of individual rocks, however, is still a challenge for classical reflection seismic methods as the important information on petrophysical properties is "hidden" in the amplitudes of the reflected signals. The seismic properties can be classified to some extend by analyzing the reflected amplitudes versus offset (AVO) for specific events.

Seismic full waveform inversion (FWI) aims at a complete seismic characterization by exploiting the full signal content of seismic measurements. The aim of FWI is to fit iteratively the relevant signals (amplitude and phase) by numerical solutions of the full wave equation. The main product of FWI are multi-parameter models of acoustic/elastic/viscoelastic/anisotropic material properties which can explain the traveltime and amplitudes of (selected) recorded signals. Such multi-parameter models can help to improve the petrophysical characterization of geological site properties in-situ. In combination with the complementary information from reflection seismic images, FWI can thus help to characterize and monitor the in-situ conditions of rocks and therefore contribute to the appraisal of disposals.

A second fundamental advantage of FWI is the possibility of utilizing various wave phenomena. In principle, all wave types predictable by the full wave equation can be exploited. This can include reflected waves, multiples, refracted waves, guided waves, mode conversions between P- and S-waves, (tunnel) surfaces waves, borehole guided waves etc. Unconventional wavefields often appear in complex environments in the presence of strong material discontinuities. Although each wave phenomena will require a special FWI workflow, they can be exploited to image specific features which conventional waves don't "see".

Despite its interesting potential in exploiting unconventional waves and subtle wavefield changes, applications of FWI for characterizing radioactive waste repositories are still relatively sparse.

This might also be due to the high computational requirements of FWI and the lack of experience and (free) software. Zhang and Juhlin (2014) use 2D FWI on a crooked line land seismic reflection seismic data to evaluate possible transport routes for radionuclides along water filled fractures zones in the planned repository Forsmark in eastern central Sweden. Bretaudeau et al. (2014) applied 2D acoustic FWI in a transmission configuration between two galleries to image faults in clay rock medium. They report that FWI allows to get images with higher resolution than traveltime tomography and that FWI is less sensitive to incomplete illumination because also diffracted waves are used. Similarly, Manukyan and Maurer (2018) successfully detect a micro-tunnel by 2D elastic VTI FWI of cross-hole data acquired in the highly anisotropic Opalinus clay formation in the Mont Terri underground rock laboratory. Bentham et al. (2018) verify by synthetic reconstruction tests that 3D acoustic FWI of surface seismic data together with downhole (tunnel) recordings has the potential to map changes within excavated distributed zones around tunnels in granite host rock at 1km depth.

In this paper we discus some aspects and ingredients of FWI relevant to assess possible future applications of FWI for site characterization. We hope to give some ideas and guidelines that help to decide whether the application of FWI can be beneficial and the efforts are worth it. For this purposes we first summarize some fundamental aspects of current FWI technology. Afterwards we illustrate the potential by showing applications to marine and land seismic data. Finally we give a short outlook on possible future developments.

2 Full Waveform Inversion (FWI)

The ultimate goal of seismic full waveform inversion (FWI) is to find all possible earth models that predict all observed wiggles (all wave phenomena) by solutions of the full wave equations. The manifold of possible earth models would then allow to estimate uncertainties for each reconstructed material parameter. Today, we have not yet reached this ambitious goal. Applications of FWI are considered to be successful if one discrete numerical model has been found which explains selected signals at low frequencies. In marine environment FWI is often performed using the acoustic approximation as the wavefield is dominated by compressional waves. Recent applications of FWI to marine data have been very impressive. The resolution of the retrieved P-wave velocity models could be increased significantly compared to traveltime tomography reconstructions (Operto et al., 2015). In contrast, applications to land seismic data are much more challenging. The reasons are (1) the higher complexity of land seismic data due to the presence of P-, S-, and surface waves and (2) current limitations of the FWI technology in recovering multi-parameter models in viscoelastic media.

2.1 Workflow

In a nutshell FWI is a PDE constrained local optimization procedure to iteratively fit observed data. The widely applied FWI workflow is illustrated in Figure 1. Starting with an initial model we iteratively calculate model updates utilizing the gradient of the misfit function which can be effectively calculated by the adjoint-state method. Without processing gradient based FWI would converge to the next local minimum of the misfit function which should be avoided. Most strategies change the topology of the misfit function during iterations by various means such as windowing the data in time and frequency, smoothing the model or the gradient, or changing the misfit function itself during iterations. The most important technique is called multi-scale approach which uses low frequencies (long wavelength) first and the gradually increases the bandwidth of the data to resolve finer details making the inverse problem gradually more non-linear. The features of the FWI workflow must be adapted to the problem under investigation.

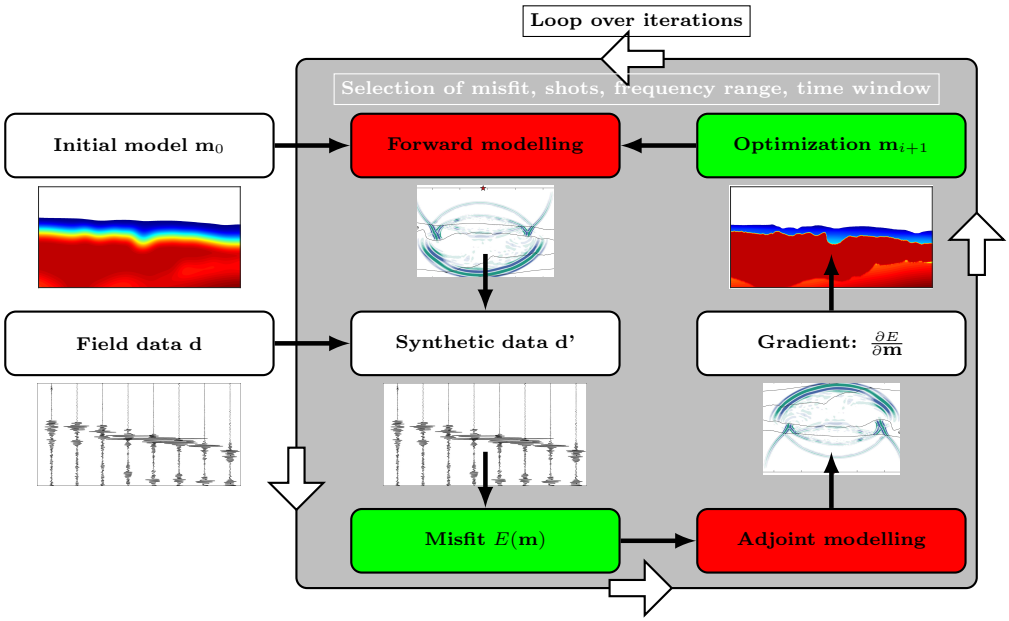

Figure 1: *A typical workflow of gradient-based FWI. An initial model that already predicts the (time windowed) relevant signals at the lowest frequencies within half a cycle is required. For this model the synthetic seismic response for one or several shots is calculated and compared with the observed signals. Different misfit functions can be defined at different iterations to quantify the difference between observed and synthetic signals in an appropriate way. The misfit definition is crucial and drives the "adjoint" modelling where the corresponding adjoint wave equations are solved. The correlation of forward and adjoint wavefield provide the gradients for all model parameters simultaneously. This gradient gives the direction to update the model to reduce the chosen misfit. Higher order optimization methods that employ second order derivatives can improve the convergence and reduce parameter trade offs. The procedure is repeated for other selected shots, higher frequencies, other signals until the procedure has reached a sufficient local minimum of misfits. The white boxes indicate task may require large storage of wavefield in space and time. The red boxes indicate the modelling tasks which may demand high computing time. The green boxes indicate the "critical" choices of misfit function and optimization procedure that mainly steer the convergence towards the local minimum.*

2.2 Starting model

The initial model is perhaps the most important decision as all gradient based optimization methods cannot change the initial information significantly. The initial model should contain a priori information for example the location of known strong contrast geological discontinuities, e.g. mining galleries. It should otherwise be smooth to allow FWI to introduce small-scale updates during iterations. The initial model should already predict the arrival times at the smallest frequency of the data within half a period to avoid cycle skipping and to ensure convergence. For this reason the acquisition of low frequencies is very important. The lower the frequency content of the observed data the lesser a priori information must be included in the initial model. This trade-off between low frequency information in the data and required resolution of the initial model may be relaxed if "adapted" misfit functions are used which can correct for cycle skipping by a filtering procedure (Warner and Guasch, 2016).

2.3 Misfit function

Another important choice to be made is the misfit function. Here we generally experience a trade-off between robustness (smoothness) and resolution. On the one hand we want to avoid to converge to a local minima too early, and on the other hand it is desirable to have high resolution and resolve small-scale model features. It may reasonable to start with a robust misfit function and then change during iterations to misfits with higher sensitive to subtable changes of signals, such as the L2 norm. A more automatic approach is decide on a suite of robust and sensitive misfit functions and then change the misfit function randomly at very iteration step (Pan et al., 2020a).

2.4 Input data

As described above the most important feature of the seismic input data is a good signal to noise ratio at low frequencies. The lower the smallest frequency of the recorded data the better to avoid cycle skipping of the relevant signals. The minimum frequency required for FWI does also depend on the long wavelength content of the initial model which can be constructed for example by traveltime tomography.

The acquisition should allow a sufficient wavefield illumination of the target where also unconventional wavefields (mode conversions, multiples, guided waves) may contribute. This illumination can be investigated by modelling experiments in which difference wavefields are calculated to emphasize the wavefield footprint of the target (Athanasopoulos et al., 2020).

The importance of multi-component recordings using 3-C geophones is not yet fully understood. Butzer et al. (2013) revealed in a modelling study that 3D elastic FWI can use the directional information of 3-C recordings to locate anomalies in the vicinity of the direct wave path. This was recently confirmed by Irnaka et al. (2019) who showed in a shallow seismic field application that the inversion of 3-C data can better resolve near surface anomalies.

2.5 Higher order optimization

FWI is a highly nonlinear inverse problem and is usually solved iteratively with a local descent optimization algorithm such as the steepest descent method and the nonlinear conjugate gradient method. These methods only use the first-order derivative of the misfit function (gradient) to estimate the descent direction and have a linear convergence rate. By incorporating the second-order derivative of the misfit function, namely the Hessian matrix, super-linear convergence can be obtained. Because of computational limitations the approximated inverse Hessian instead of the full Hessian is often calculated for gradient preconditioning. Besides faster convergence rate, second-order optimization is important in multi-parameter FWI since the Hessian matrix contains information related to parameter scalings and inter-parameter trade-off also called cross-talk. Cross-talk refers to the observation that a perturbation in one parameter produces artifacts in a second parameter because of similarities in their seismic (scattering) response. Our recent work indicates that for near surface applications the footprint between certain groups of material parameters can be reduced by using second-order optimization (Gao et al., 2020a). Although second-order methods converge faster, the overall computational costs are significantly higher because an additional Newton or Gauss-Newton equation must be solved at every iteration.

2.6 Estimation of uncertainties

The output of FWI is generally one deterministic 'best-fit' model which can explain the relevant signals at low frequencies. Because of the high degrees of freedom (high number of model param-

eters) a very good data fit can be obtained in many applications. The data fit alone is thus not sufficient to evaluate the reliability of the reconstructed multi-parameter model.

One formal way to quantify uncertainties of inverted model parameters is the framework of Bayesian inference (Tarantola, 2005), which requires an adequate sampling of the model space to estimate a posterior Bayesian probability. Another approach is to estimate the uncertainty by using the ensemble transform Kalman filter (Thurin et al., 2019) or the square-root variable metric (Liu et al., 2019). However, all these methods require significant additional computational cost which can be even more expensive than the cost of running a conventional FWI. A computationally more efficient approach is to estimate the Hessian matrix using the inverse of the posterior covariance (Fichtner and van Leeuwen, 2015). This method, however, requires that the inverted model is close to the global minimum which is unclear in many practical applications where we have limited a priori information and observe a high complexity of the unknown (residual) wavefield (Figure 3). A new more general approach which is computationally still affordable is to calculate the distribution of Pareto optimal models in the model space in a multi-objective FWI framework (Shigapov, 2019; Pan et al., 2020b).

2.7 Computational resources

The computational requirements of FWI are demanding. The resources are mainly consumed by the forward and adjoint simulations (see red boxes in Figure 1) which have to be performed for each shot. In addition, the forward wavefield must be stored in order to perform the wavefield correlation to obtain the gradients.

In the following we roughly estimate the run time for gradient-based viscoelastic FWI in two and three dimensions. This might be helpful to estimate very roughly the required computational resources for a planned land seismic FWI project and decide whether FWI is generally affordable. For simplicity we calculate the run times for the forward and adjoint modellings only and ignore all other parts of the workflow such as filtering, optimization procedure, correlation etc. This is plausible as most of the resources are required for wavefield modelling. We further consider that the wavefield correlation is performed "on the fly" without storing the wavefield which requires one additional forward calculation for each shot. This results in 3 simulations per shot. The estimation of the step length by parabolic line search requires 3 additional forward simulations. We therefore assume a total number of 6 simulations per shot. We further assume that the wavefield is calculated by viscoelastic Finite-Difference (FD) modelling which is computationally relatively efficient especially on HPC systems (Bohlen, 2002). Also the required computing times for FD modelling can be estimated quite accurately. In oder to derive simple equations for estimating computing times we assume that the model volume has dimension $D \in \{2, 3\}$ with N_λ wavelength in each direction. Each wavelength is discretized with G grid points per wavelength. The total number of shots is N_λ^{D-1}/X_{shot}. This means that the shots completely cover on side of the model, e.g. the surface, with an average spacing of X_{shot} wavelength. With these assumptions we can estimate the computing time of viscoelastic FWI in dimensions $D \in \{2, 3\}$ to (Bohlen, 1998) :

$$T \approx \frac{6}{X_{shot}^{D-1}} \cdot \frac{OP(D)}{P} G^D N_T I_T N_\lambda^{2D-1} \sim \frac{G^D N_\lambda^{2D-1}}{X_{shot}^{D-1} P}. \tag{1}$$

The performance of the HPC system P is measured in FLoating point OPerations per Second (FLOPS), and mainly dependents on the number of cores being used.

We now calculate expected run times for the following typical parameters: $N_T = 10^4$ (number of FD time steps), $I_T = 100$ (number of FWI iterations), $OD(2, 3) \approx 150, 300$ (number of floating point operations per grid point in 2D and 3D), $X_{shot} = 2$ (distance between shots in wavelength). With these parameters and basic assumptions (cubic model and only considering floating point operations of 6 modellings per shot) we estimate the run times for viscoelastic FWI when a model size of N_λ wavelength in each model direction is considered. In Figure 2 (left) we show expected

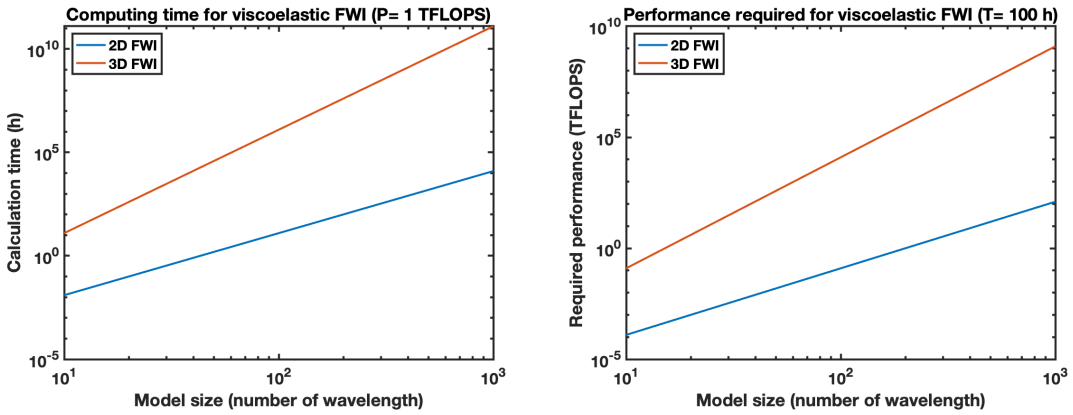

Figure 2: *Estimation of required computing times for viscoelastic FWI in 2D and 3D using equation 1. Size of model is N_λ wavelength in each direction. Shots are placed every second wavelength along one side of model. Left: Estimated run time (T) at $P = 10^{12}$ FLOPS. Right: Required performance (P) to finish viscoelastic FWI within $T = 100h$.*

run times on a moderate HPC system with $P = 1TFLOPS$. We see that viscoelastic FWI for typical model sizes of 50-100 wavelength requires less than one hour. In 3D, however, run times increase by a factor GN_λ^2/X_{shot} (equation 1) compared to the 2D case. For a model size of $N_\lambda = 50$ wavelength this results in a factor $10 \cdot 50^2/2 \approx 10^4$ between 2D and 3D. Viscoelastic FWI in 3D is thus considerable more expensive but computationally feasible on modern HPC resources (see Figure 2 (right)).

3 Applications

In recent 20 years FWI has received great attention and has been applied sucessfully to a broad range of spatial scales and wave types (Figure 3). In the following we examplarily present two of our recent applications of 2D FWI. The first example is a marine data set where 2D acoustic FWI have been used to improve the resolution of the P-wave velocity model of shallow marine sediments. In our second example we present one of our first applications of 2D viscoelastic multi-parameter FWI to shallow land seismic data dominated by Rayleigh waves to retrieve models of seismic wave velocities and attenuation within the upper 10 meters of the subsurface.

3.1 Application to marine data

To illustrate the potential of FWI in providing valuable complementary subsurface information using unconventional wavefields we show an application of 2D acoustic FWI applied to dispersive guided wave modes acquired in shallow water by an Ocean-Bottom-Cable (OBC) (Figure (4). The dispersive acoustic modes contain valuable information about the P-wave velocity structure which improves the resolution of the initial traveltime tomography model. The anomalies in the final FWI model indicate shallow gas accumulations which are also visible as high reflectivity anomalies in the migrated image of reflected waves (which have not been used in FWI).

3.2 Application to land data

Our second application shows the first succesful application of viscoelastic 2D FWI to retrieve multi-parameter models of wave velocities and attenuation structure from land seismic data (Figure 5). The inferred parameter variations appear realistic and can explain the observed waveforms.

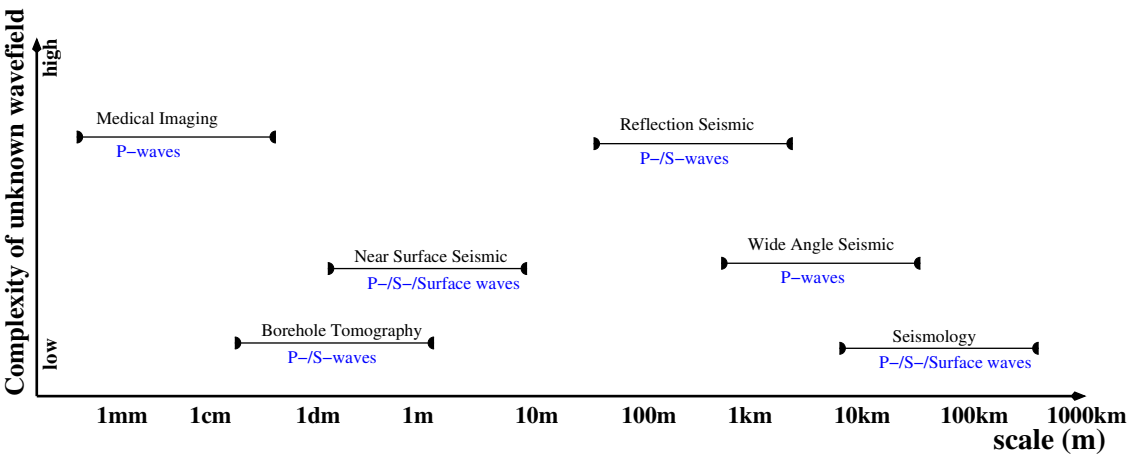

Figure 3: *Today, 2D acoustic/elastic FWI has become feasible on a wide range of applications/scales covering nine orders of magnitude. Different wave types and acquisition configurations (transmission versus reflection) are applied. Applications in which the wavefield is composed of multiple scattered/reflected waves which is the case in medical imaging and reflection seismics are most challenging.*

However, further investigations are necessary to evaluate the reliability of recovered models especially with respect to attenuation.

4 Outlook

First applications revealed that gradient-based FWI has a great potential and is applicable on different wave types acquired on a broad range of spatial scales covering 9 orders of magnitude. However, recent results and current problems also indicate that we are still in the early stage of the development of this technology. Despite some impressing applications to mainly marine seismic data, many fundamental aspects are not yet satisfactorily solved. This includes for example the reliable reconstruction of multi-parameter models and the efficient quantification of model parameter uncertainties. Furthermore, robust workflows for the routine application to 3D land reflection seismic data are not yet available. In addition, 3D land reflection seismic data will generally require 3D viscoelastic forward modelling which is computationally very demanding even on large HPC systems.

A promising direction to increase the robustness with respect to the initial model and also to mitigate parameter trade-offs is to introduce a greater flexibility in the choice of misfit functions. Proper misfit functions might help to mitigate local minima and at the same time better isolate certain parameter pertubations.

By performing a few iterations on randomly choosen data subsets, e.g. shot gathers, different models can be obtained without significant additional computational cost. These models allow to define model uncertainties which correspond to the variation of illumination for different subsets of the data.

Recent synthetic studies indicate a clear benefit of higher order optimization methods: (1) they allow a physical scaling of each material parameter and (2) cross-talk between parameters can be predicted and then used to precondition the gradient and reduce cross-talk in the model updates. Unfortunately, higher order methods require significantly more forward simulations and will thus increase the computational demands of FWI even more.

5 Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, Project-ID 258734477-SFB 1173) and by the Federal Ministry of Education (BMBF, project WAVE, grant no. 01IH15004). The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS at Jülich Supercomputing Centre (JSC).

References

Athanasopoulos, N., Manukyan, E., Maurer, H., and Bohlen, T. (2020). Time-frequency windowing in multiparameter elastic fwi of shallow seismic wavefield. *Geophysical Journal International*, submitted.

Bentham, H. L. M., Morgan, J. V., and Angus, D. A. (2018). Investigating the use of 3-D full-waveform inversion to characterize the host rock at a geological disposal site. *Geophysical Journal International*, 215(3):2035–2046.

Bohlen, T. (1998). *Viskoelastische FD-Modellierung seismischer Wellen zur Interpretation gemessener Seismogramme*. PhD thesis, Christian-Albrechts-Universität zu Kiel.

Bohlen, T. (2002). Parallel 3-D viscoelastic finite difference seismic modelling. *Computers and Geosciences*, 28(8):887 – 899.

Bretaudeau, F., Gélis, C., Leparoux, D., Brossier, R., Cabrera, J., and Côte, P. (2014). High-resolution quantitative seismic imaging of a strike-slip fault with small vertical offset in clay rocks from underground galleries: Experimental platform of tournemire, france. *GEOPHYSICS*, 79(1):B1–B18.

Butzer, S., Kurzmann, A., and Bohlen, T. (2013). 3D elastic full-waveform inversion of small-scale heterogeneities in transmission geometry. *Geophysical Prospecting*, 61(6):1238–1251.

Fichtner, A. and van Leeuwen, T. (2015). Resolution analysis by random probing. *Journal of Geophysical Research: Solid Earth*, 120(8):5549–5573.

Gao, L., Pan, Y., Rieder, A., and Bohlen, T. (2020a). Multi-parameter viscoelastic shallow seismic full waveform inversion with the truncated Newton method. In *82nd EAGE Conference and Exhibition 2019*, volume 2020. European Association of Geoscientists & Engineers.

Gao, L., Yudi, P., and Bohlen, T. (2020b). Reconstructing 2D near-surface models via viscoelastic full waveform inversion of shallow-seismic surface wave. *submitted to Geophysical Journal International*.

Habelitz, P. M. (2017). 2D akustische Wellenforminversion geführter Wellen im Flachwasser. Master's thesis, Karlsruhe Institute of Technology.

Irnaka, T. M., Brossier, R., Métivier, L., Bohlen, T., and Pan, Y. (2019). Uncovering the effect of multi-component data on 9C 3D Elastic FWI: Ettlingen Line Case Study. In *AGU Fall Meeting 2019*. AGU.

Kunert, M., Kurzmann, A., and Bohlen, T. (2016). Application of 2D Acoustic Full Waveform Inversion to OBC-data in Shallow Water. In *78th EAGE Conference and Exhibition 2016*. EAGE.

Liu, Q., Peter, D., and Tape, C. (2019). Square-root variable metric based elastic full-waveform inversion–part 1: theory and validation. *Geophysical Journal International*, 218(2):1121–1135.

Manukyan, E. and Maurer, H. (2018). *Imaging of radioactive-waste repository with vertically transversely isotropic full-waveform inversion*, pages 4743–4747. SEG Technical Program Expanded Abstracts.

Operto, S., Miniussi, A., Brossier, R., Combe, L., Metivier, L., Monteiller, V., Ribodetti, A., and Virieux, J. (2015). Efficient 3-d frequency-domain mono-parameter full-waveform inversion of ocean-bottom cable data: application to Valhall in the visco-acoustic vertical transverse isotropic approximation. *Geophysical Journal International*, 202(2):1362–1391.

Pan, Y., Gao, L., and Bohlen, T. (2020a). Random-objective waveform inversion of shallow-seismic SH and Love waves. In *82nd EAGE Conference and Exhibition 2020*. EAGE.

Pan, Y., Gao, L., and Shigapov, R. (2020b). Multi-objective waveform inversion of shallow seismic wavefields. *Geophysical Journal International*, 220(3):1619–1631.

Shigapov, R. (2019). *Probabilistic waveform inversion: Quest for the law*. PhD thesis, Karlsruher Institut für Technologie (KIT).

Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. Society for Industrial and Applied Mathematics.

Thurin, J., Brossier, R., and Métivier, L. (2019). Ensemble-based uncertainty estimation in full waveform inversion. *Geophysical Journal International*, 219(3):1613–1635.

Warner, M. and Guasch, L. (2016). Adaptive waveform inversion: Theory. *GEOPHYSICS*, 81(6):R429–R445.

Zhang, F. and Juhlin, C. (2014). Full waveform inversion of seismic reflection data from the forsmark planned repository for spent nuclear fuel, eastern central sweden. *Geophysical Journal International*, 196(2):1106–1122.

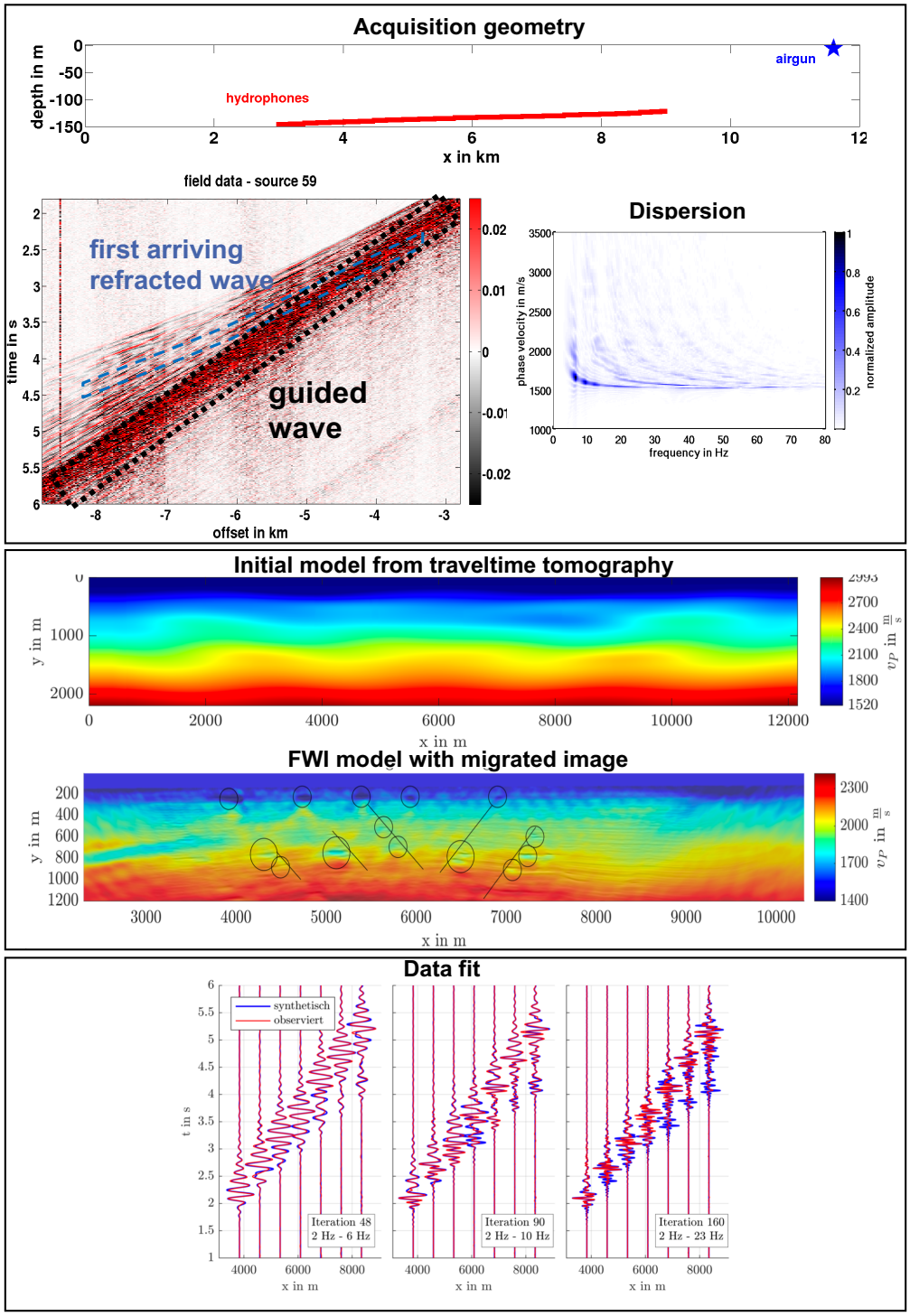

Figure 4: *Application of 2D acoustic FWI to marine OBC data (Kunert et al., 2016; Habelitz, 2017). Top box: The field data is recorded in shallow water with a 6km long OBC cable. The wavefield is composed of refracted waves and high amplitude dispersive guided waves with several modes. Middle box: Initial model obtained by traveltime tomography of the first arrivals and final FWI model underlain by migrated image of reflected waves. The improved resolution of FWI velocity model allows to identify shallow gas accumulations and faults. Bottom box: The final FWI model can predict the guided modes and refracted waves up to approximately 10 Hz.*

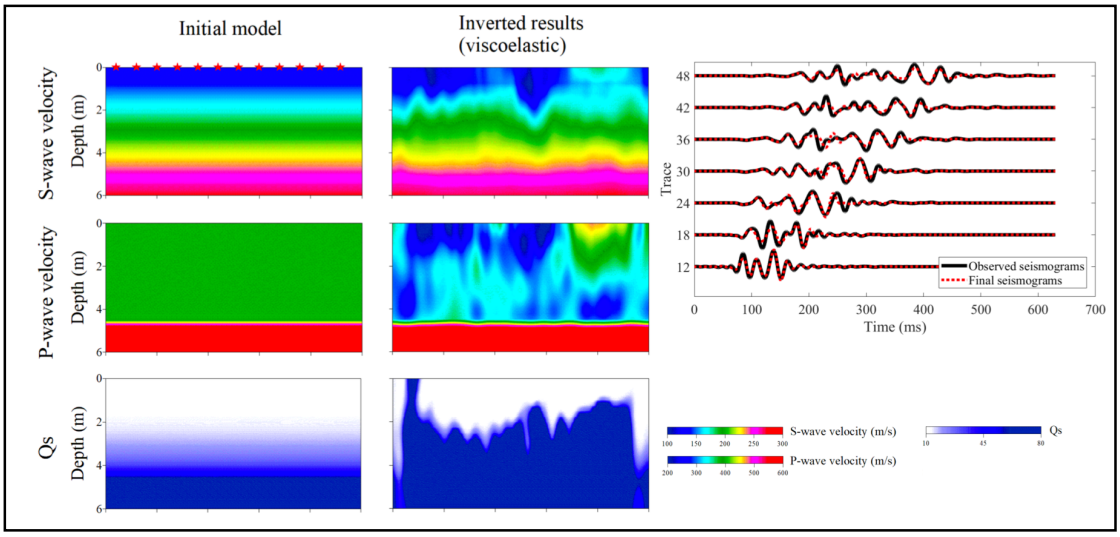

Figure 5: *Application of 2D viscoelastic FWI to shallow seismic land data (Gao et al., 2020b). The recovered multi-parameter models of seismic wave velocities and S-wave attenuation are plausible and can explain the full information content of the seismic data.*