

# Two-grid full-waveform Rayleigh-wave inversion via a genetic algorithm — Part 2: Application to two actual data sets

Zhen Xing<sup>1</sup> and Alfredo Mazzotti<sup>2</sup>

## ABSTRACT

We have applied our two-grid genetic-algorithm Rayleigh-wave full-waveform inversion (FWI) to two actual data sets acquired in Luni (Italy) and Grenoble (France), respectively. Because our technique used 2D elastic finite-difference modeling for solving the forward problem, the observed data were 3D to 2D corrected prior to the inversion. To limit the computing time, both inversions focused on predicting low-resolution, smooth models by using quite coarse inversion grids. The wavelets for FWI were estimated directly from the observed data by using the Wiener method. In the Luni case, due to the strong dispersion effects on the data, to strengthen the inversion, envelopes and waveforms were considered in the objective function and an offset-marching strategy was applied. Though no a priori information was exploited, the outcomes of the Luni and Grenoble data inversion were fair. The predicted Luni  $V_S$  model

indicates a strong velocity increase from approximately 3 to 6 m, and velocity inversions have been detected at approximately 2 and 9 m depths. Analyzing the dispersion spectra, it results that the predicted Luni data reasonably reproduced the waveforms related to the fundamental mode and, likely, a small part of those related to the first higher mode. Concerning the Grenoble example, the predicted  $V_S$  model coincides reasonably well with the long-wavelength structures presented in the  $V_S$  profiles obtained from nearby boreholes. The data reconstruction is generally satisfactory, and when mismatches occur between the predicted and observed traces, the phase differences are always within half-periods. The fair inversion outcomes suggest that the predicted Luni and Grenoble models would likely be adequate initial models for local FWI, which could further increase the resolution and the details of the estimated  $V_S$  models.

## INTRODUCTION

Multichannel analysis of surface waves (predominantly Rayleigh waves) is a technology that has brought major advancements in the reconstruction of S-wave velocities  $V_S$  of shallow layers (among others, Park et al., 1999; Xia et al., 1999; Bohlen et al., 2004; Socco and Strobbia, 2004; Cercato, 2009; Maraschini et al., 2010; Socco et al., 2010). Most of the used techniques assume a 1D-layered model and invert the dispersion curve of the fundamental mode, which has to be identified in the dispersion spectrum of the data.

In recent years, Rayleigh-wave full-waveform inversion (FWI) has been introduced (Schäfer et al., 2013; Tran et al., 2013; Groos

et al., 2014; Masoni et al., 2014, 2016) with the purpose of taking into account the whole waveform information present in the recorded data and to relax some of the assumptions inherent in the dispersion curve inversion. For instance, this new approach naturally supports the prediction of 2D or even 3D  $V_S$  models. However, to date, there have been only a few applications to actual Rayleigh-wave data (Schäfer et al., 2013; Tran et al., 2013; Groos et al., 2017). As a matter of fact, because of the strong nonlinearity of Rayleigh waves (Forbriger, 2003a, 2003b; Rix, 2004; Brossier et al., 2009; Schäfer et al., 2013), when a suitable initial model is not available, local, gradient-based Rayleigh-wave FWI could get trapped into local minima. A suitable initial model is supposed

Manuscript received by the Editor 19 November 2018; revised manuscript received 19 May 2019; published ahead of production 27 June 2019; published online 13 September 2019.

<sup>1</sup>Formerly University of Pisa, Department of Earth Sciences, Pisa, Italy; presently CGG, Crawley, UK. E-mail: zhnngx@outlook.com (corresponding author).

<sup>2</sup>University of Pisa, Department of Earth Sciences, Pisa, Italy. E-mail: alfredo.mazzotti@unipi.it.

© The Authors. Published by the Society of Exploration Geophysicists. All article content, except where otherwise noted (including republished material), is licensed under a Creative Commons Attribution 4.0 Unported License (CC BY-NC). See <http://creativecommons.org/licenses/by/4.0/>. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its digital object identifier (DOI). Commercial reuse is not permitted. The same license does not have to be used for derivative works.

to contain the long-wavelength structure of the investigated near-surface zone. It should also limit the cycle skipping between the observed and predicted data, particularly in the portions of data containing the fundamental mode. In actual cases, such strict requirements may be difficult to satisfy, especially when a priori information is not available or is too scarce.

To deal with this issue, in the companion paper (Xing and Mazzotti, 2019), we propose a new Rayleigh-wave inversion method, which we named two-grid genetic-algorithm Rayleigh-wave FWI. The proposed method makes use of a 2D elastic finite-difference modeling (FDM) code (Thorbecke and Draganov, 2011; Xing and Mazzotti, 2016) as the modeling engine and uses a global optimization technique, a genetic algorithm, as the inversion tool. The genetic algorithm, instead of being directed by local gradients of the objective function, stochastically explores a wide model space (Stoffa and Sen, 1991; Sen and Stoffa, 1992; Mallick, 1995, 1999). As a consequence, compared with gradient-based optimization techniques, it is much less vulnerable to local minima. In our inversion workflow, to further tackle the local minimum problem, we can choose to use a frequency-marching strategy (Bunks et al., 1995) and an offset-marching strategy, and the data misfit computation can include the waveforms, their envelopes, or both. To increase efficiency, we have implemented a two-grid scheme (Sajevo et al., 2014, 2016; Aleardi and Mazzotti, 2017; Mazzotti et al., 2017), which turns out to be indispensable for the practical application of our Rayleigh-wave inversion method. In fact, an exponential law links the number of unknowns to the computational time required by the genetic algorithm (Bellman, 1957); thus, we use a coarse grid in the inversion, with the number of unknowns proportional to the number of grid nodes, to balance between computational time and resolution of the predicted  $V_S$  model. Instead, a fine grid is used in forward modeling to guarantee the reliable computation of Rayleigh waves (Xing and Mazzotti, 2017a, 2018). Each predicted coarse-grid model is converted into its fine-grid equivalent by means of bilinear interpolation. To illustrate the feasibility of the proposed method, three synthetic examples with rather complex reference models have been considered. Through the examples, we have shown that, even in the case of null a priori information, our method was able to fairly predict the long-wavelength structures of the reference models, along with the quite satisfactory reconstructions of the “observed” data. However, several facilitating factors were present in the three synthetic examples: perfectly known wavelet, no (or limited) noise contamination, and the observed data and the predicted data were computed with the same 2D elastic FDM code. In the real world, these conditions are never met; thus, additional considerations need to be taken in the actual application of the proposed method.

In actual cases, first, to match the seismograms generated by the 2D forward modeling, observed data must be 3D to 2D corrected for compensating the different spreading factors. To this end, we use a correction algorithm proposed by Forbriger et al. (2014). Following Forbriger et al. (2014) and Schäfer et al. (2014), the traces at the near offsets, which are defined as the offsets shorter than the estimated dominant wavelength of the observed Rayleigh waves (Schäfer et al., 2014), are corrected by means of the single-layer transformation, whereas the rest of the traces are compensated via the multilayer transformation. Second, the wavelet needs to be estimated for actual FWI and the proper estimation is still an ongoing topic that draws much interest. In the two inversion cases that will

be presented, a significant effort was devoted to wavelet estimation, using different methods and performing several inversion tests. Some details will be given when discussing the first example on the Luni data.

In the inversion, we use elastic modeling, which is merely an approximation of the true anelastic wave propagation. Although the elasticity assumption is common for many Rayleigh-wave inversion approaches, there have been studies on the impact of neglecting attenuation on gradient-based Rayleigh-wave FWI (Groos et al., 2014; Dokter et al., 2017). Xing and Mazzotti (2017b) carry out an inversion test considering a synthetic model with S-wave quality factors as low as 3.75 and observe that neglecting such strong attenuation led indeed to the degradation of the model prediction, especially for what concerns the prediction of the velocity values, but the dominant subsurface shapes and velocity inversions were still predicted by our stochastic inversion method. On the other hand, adopting viscoelastic forward modeling (as performed by Bohlen, 2002) in our inversion scheme would cause a significant increase in computational time, due to the use of more complex modeling equations, the stricter requirements for guaranteeing the modeling stability, such as the demand for smaller spatial intervals in the modeling, and possibly the need of including quality factors as unknowns. Therefore, with our inversion approach, we know that the predicted velocities may be imprecise and affected by an error whose magnitude depends on the severity of attenuation, but the velocity trends should be still fairly predicted. Thus, other less computer intensive inversion approaches, starting from the predicted long-wavelength model, may follow and be able to take into account the quality factors, either as unknown or, more likely, as a priori information.

In the following sections, we illustrate the inversion of two field data cases. The first data set was acquired near the archaeological site of Luni (Italy). The second data set was acquired in Grenoble (France) in the framework of the InterPACIFIC project documented in Garofalo et al. (2016a, 2016b).

In each case, first, we will describe the data acquisition and the few processing steps that we performed and, next, we will show how we dealt with the wavelet estimation. Then, the inversion specifications, such as the chosen objective function, genetic-algorithm controlling parameters, frequency- and/or offset-marching scheme, modeling and inversion grid sizes, and so forth, will be given. Finally, we will illustrate and discuss the inversion outcomes. In both cases, no a priori information was used to guide the inversion; that is, the models of the starting population were randomly distributed within search ranges constant with depth and lateral distance. Coarse inversion grids were used with the aim of deriving long-wavelength structures of the subsurface.

## LUNI DATA INVERSION

### Luni data acquisition and processing

The Luni field data were acquired in March 2017 in an area with a horizontal topographic surface. A total of 36 vertical-component geophones, 4.5 Hz natural frequency, were laid out inline at 1 m intervals. The energy source was a seismic gun. Three shot gathers, of which one was split-spread and two were off-end, were used in our FWI. The maximum offset of the split-spread and the off-end shot gathers are 17.5 and 38 m, respectively. The shot gathers are displayed in Figure 1 (one trace every two has been plotted) and

show a general good quality, although at early times, they are contaminated by a strong air-blast event, which likely interferes with the Rayleigh waves.

The few processing steps we carried out are geometry assignment; air-blast attenuation; band-pass filtering with corner frequencies of 5, 8, 25, and 30 Hz; top and bottom muting to remove the pre-first-break noise and to isolate the Rayleigh wavetrain; and 3D to 2D correction (Forbriger et al., 2014).

The air blast was attenuated by applying an inverse amplitude scalar to the air-blast energy. The inverse amplitude was estimated dynamically based on the average amplitude of the signals in the windows above and below the air-blast event.

The choice of 5–8 Hz as the low corner frequencies of the band-pass filter was dictated by the results of a frequency analysis that showed that at less than 5 Hz, no significant Rayleigh-wave energy is present. The choice of the high corner frequencies (25–30 Hz) was taken to limit the computing time to acceptable levels, although Rayleigh wave energy is also present at higher frequencies.

The top and bottom mutes were applied to remove noise, such as pre-first-break noise and possible seismic gun baseplate oscillations. It should be noted that the very same mute functions were also applied to all of the predicted seismograms, before the computation of the data misfit.

The three shot gathers after the processing are presented in Figure 2. The strong air blast has been well-attenuated. The

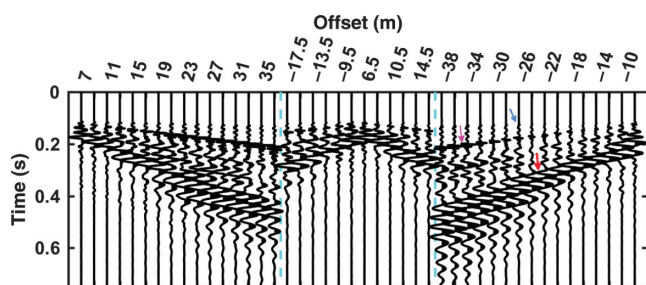


Figure 1. Raw seismic data acquired in Luni, Italy. One of every two traces is plotted. The dashed cyan lines delimit the left, middle, and right shot gathers. The seismograms have been trace-by-trace normalized. A strong air-blast event (pointed out by the magenta arrow) is present just below the first breaks (indicated by the blue arrow) due to the use of a seismic gun as the energy source. Most of the clearly shown events in the seismograms are Rayleigh waves, e.g., the one pinpointed by the red arrow.

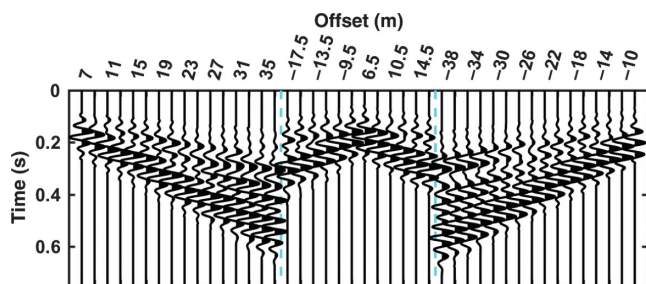


Figure 2. The Luni field data of Figure 1 after processing and the 3D to 2D correction (Forbriger et al., 2014). The maximum frequency is 30 Hz. The dashed cyan lines delimit the left, middle, and right shot gathers. The seismograms have been trace-by-trace normalized. Most of the clearly shown events are Rayleigh waves.

Rayleigh-wave events are clearly recognizable and show a high degree of dispersion, aggravating the complexity of the observed waveforms and thus the difficulty of the inversion (Bunks et al., 1995; Cercato, 2011), especially if no a priori information is exploited.

### Luni wavelet estimation

Wavelet estimation is essential for FWI due to its direct impact on the data simulation in the forward modeling and consequently on the computed data misfits. We tried different ways to estimate the wavelet from the field data, considering only the first-break body waves, or taking into account the whole data, and using different estimation approaches, such as the Wiener method, the Hilbert transform technique, and the method based on singular-value decomposition (SVD). The Wiener method and the Hilbert transform technique estimate the wavelet from the spectrum of the data and assume a minimum phase (see, e.g., Claerbout, 1985). Instead, the SVD approach computes the eigenimages for the wavelet estimation (Ulrych et al., 1988). All of the methods that we tried estimate the wavelet from the seismic data only and do not make use of any, yet approximate, subsurface model. At present, given that we do not consider any a priori information about the model, estimating the wavelet from the seismic data only is a necessity. However, because models are being predicted along the inversion, at certain steps of the evolution they could be used in a model-based wavelet estimation approach to improve the wavelet. Or even, the wavelet updates could be considered as additional unknowns in the inversion. These possible improvements need to be further investigated for their inclusion in a stochastic inversion procedure such as the one we propose.

Finally, the wavelet we chose for the inversion was an average of the wavelets estimated trace by trace on the full data set by using the Wiener method. In fact, this wavelet led to reasonable results and better data matching in the inversion and resulted as very similar to the wavelet estimated through SVD on the data of a nearby roll-along data set, although the SVD approach does not require the minimum phase assumption.

Figure 3 shows the trace-by-trace estimated wavelets that are rather similar among each other, especially at early times. The weighted mean of these wavelets yields the final wavelet used for the inversion. Higher weights were given to the near-offset wavelets. The weights are plotted in Figure 4.

Figure 5 also shows the wavelet (in red) estimated using SVD on the near-offset traces of a roll-along data set acquired by using the

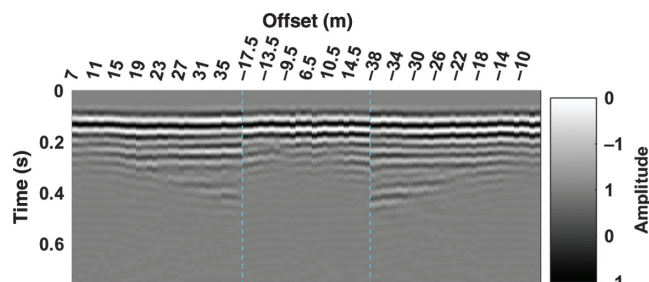


Figure 3. Wavelets estimated trace by trace from the Luni field data by the Wiener method. The dashed cyan lines delimit the wavelets estimated from the left, middle, and right shot gathers. The wavelets are very similar among each other, especially at the early times.

same energy source and receivers at the same site on the same day as the acquisition of the data input to the inversion shown in Figure 1.

Though the different estimation methods and the different amount of data are used for the estimation, the two wavelets in Figure 5 are quite similar, further corroborating the use of the wavelet (in black) in Figure 5 as the input wavelet for the Luni data inversion.

### Luni data inversion specifications

We inverted for  $V_S$ , P-wave velocities  $V_P$ , and densities  $\rho$ . Based on the scientific literature (among others, Xia et al., 1999) and our experience on synthetic data inversion,  $V_P$  and  $\rho$  are to be considered more as dummy parameters due to their much lower influence on Rayleigh waves compared to  $V_S$ .

The genetic-algorithm search ranges for  $V_S$ ,  $V_P$ , and  $\rho$  were constant with depth and laterally and wide enough to include reasonable velocity and density values for shallow sedimentary layers. Sequentially for  $V_S$ ,  $V_P$ , and  $\rho$ , the search range was set as 100–500 m/s, 600–2000 m/s, and 1500–1920 kg/m<sup>3</sup>. A further constraint is given by setting the admissible  $V_P/V_S$  ratios between 1.5 and 10.

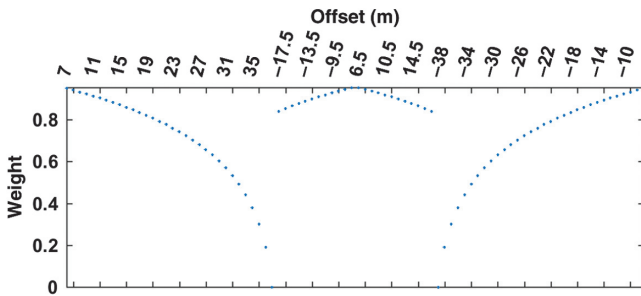


Figure 4. The weights applied to the wavelets displayed in Figure 3 for estimating the final wavelet used in our Luni data inversion. Higher weights were given to the near-offset wavelets.

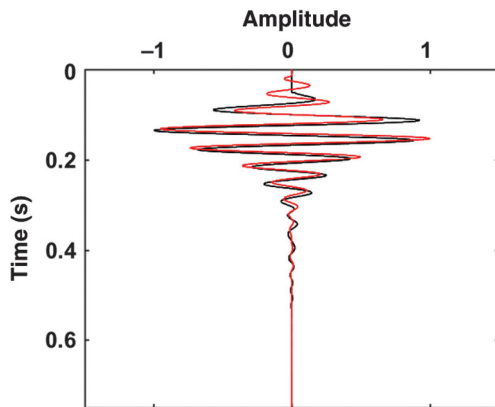


Figure 5. The wavelet computed as the weighted mean of the wavelets presented in Figure 3 (the black curve). The wavelet obtained from a roll-along data set through the SVD approach (the red curve). Though the two wavelets have been estimated by using different techniques, they are rather similar to each other.

Based on the defined velocity search ranges and the maximum frequency of the input data, the fine modeling grid, which guaranteed 20 points per minimum wavelength, was built with 31,913 nodes, i.e., with a grid spacing of 0.125 m. Under the two-grid scheme, we constructed the coarse inversion grid with 45 nodes, which resulted in  $45 \times 3 = 135$  unknowns. Because no information was available to suggest velocity variations, except that the vertical velocity variations should be greater than the horizontal variations, a rectangular 10.25–1.5 m inversion grid was used. The coarseness of the inversion grid indicates that the predicted model will be a “macro” model with only the long-wavelength structure of the investigated zone reproduced.

We implemented the possibility of using different objective functions and different strategies of inversion, such as frequency marching, time stripping, and offset marching. In the specific example of Luni, the objective function ( $\chi$ ) that we used is composed of two parts: The first takes into consideration the waveform of the data, and the second considers the envelope of the data:

$$\chi = \frac{\sum_{n_x=1}^{N_x} \sum_{n_t=1}^{N_t} |\mathbf{D}_{n_t, n_x}^w - \mathbf{P}_{n_t, n_x}^w|}{\sum_{n_x=1}^{N_x} \sum_{n_t=1}^{N_t} |\mathbf{D}_{n_t, n_x}^w|} + \alpha \times \frac{\sum_{n_x=1}^{N_x} \sum_{n_t=1}^{N_t} |\mathbf{D}_{n_t, n_x}^e - \mathbf{P}_{n_t, n_x}^e|}{\sum_{n_x=1}^{N_x} \sum_{n_t=1}^{N_t} |\mathbf{D}_{n_t, n_x}^e|}, \quad (1)$$

where  $\mathbf{D}$  and  $\mathbf{P}$  are the trace-by-trace normalized observed and predicted data, respectively, superscripts  $w$  and  $e$  denote the waveform and the enveloped traces, respectively,  $n_t$  and  $n_x$  are the time and the trace sampling index, respectively, and  $N_t$  and  $N_x$  indicate the maximum number of time samples and traces, respectively. The symbol  $\alpha$  is a predefined weight that can be varied with generations. According to our tests, equiweighting the waveform misfit and the envelope misfit generally led to reasonable inversion results. Thus, in the Luni case,  $\alpha$  has been kept constant and is equal to one. As is well-known, the use of enveloped traces (Chi et al., 2014; Wu et al., 2014) further alleviates the cycle-skipping problem and helps the inversion converge toward more promising zones of the model space, especially at the early generations of the genetic-algorithm evolution. The inversion of the enveloped traces is less dependent on the estimated input wavelet.

It should be noted that in equation 1, no regularization terms in the model space are present because it would require the knowledge of at least some characteristics of the model, a case we are not considering here. However, regularization in the model space was implicitly applied in our inversion by the smoothing of the model through bilinear interpolation from the coarse inversion grid to the fine modeling grid.

Offset marching was also used in the inversion, starting from short offset (up to 10.5 m) and then including longer offset traces every 40 generations until reaching the maximum offset of 38 m.

As shown in Figure 2, the dispersion of the Rayleigh waves at the near offset is not fully developed yet. Consequently, the near-offset traces are much less complex than those at the far offset; thus, we can suppose that including in the inversion the short-offset traces only, the corresponding error function is less complicated as well.

On the basis of the experience acquired in our tests with synthetic data, the parameters of the genetic algorithm optimization were set as follows: maximum generation, 200; number of individuals, 2000; number of subpopulations, 5; selection rate, 0.8; reinsertion rate,



0.6; and mutation rate, 0.1. The selection pressure was increased from 1 to 2 along with generations. The first migration happened at the 30th generation. Then, it occurred every 20 generations. The migration rate was 0.2. This setting of parameters was also adopted for the Grenoble data inversion that will be discussed later.

### Luni data inversion results

Figure 6 exhibits the evolution of the minimum and the mean data misfit. Within each offset-marching phase, the decrease in the data misfits is evident. The data misfits in equation 1 have been normalized separately within each offset-marching step; thus, the actual misfit values cannot be compared among different steps. However, at generations in which the inclusion of new traces in the observed data occurred, a sudden increase in the data misfit may truly take place. At the last generations within each offset range, the improvement of the minimum data misfit was very slow and converged to the mean data misfit, indicating that continuing the inversion would lead to negligible improvement.

Figure 7 shows the predicted  $V_S$  macromodel showing smooth velocity variations, characterized by a significant velocity increase between approximately 3 and 6 m and by velocity inversions at approximately 2 and 9 m.

Figure 8 shows the mean of the  $V_S$  models at the last generation of the inversion. This model is encouragingly quite similar to the predicted model in Figure 7. The long-wavelength structure reconstructed by the inversion (Figures 7 and 8) is predominantly one dimension. When the lateral velocity variations in the subsurface are not very

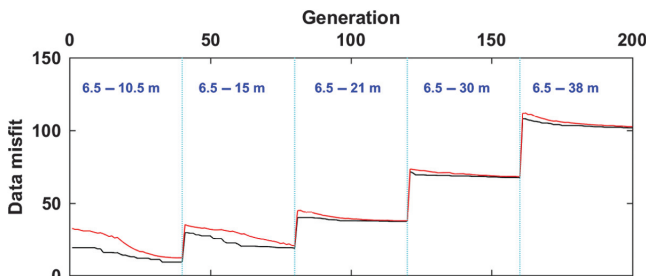


Figure 6. The evolution of the minimum (black) and mean (red) data misfit in the Luni data inversion, with offset marching. The annotations in blue bounded by the dashed cyan lines indicate the absolute-offset ranges of the inverted data in the offset-marching scheme. In each offset-marching phase, the data misfits decrease evidently.

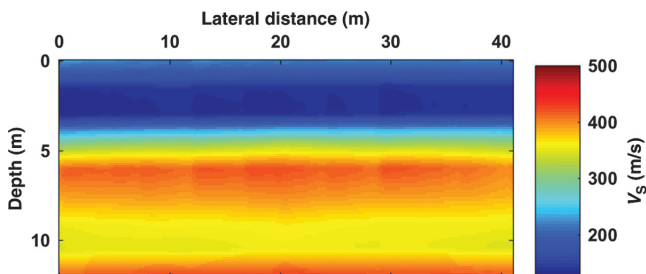


Figure 7. The  $V_S$  model predicted from the Luni data inversion, in which offset marching has been applied. The model shows a strong velocity increase from approximately 3 to 6 m. Additionally, velocity inversions at approximately 2 and 9 m have been detected.

significant, the smoothing of the model, the coarseness of the inversion grid, and the fact that genetic-algorithm optimization is driven by average data misfits computed over all the shots, often render macrosurface models with a principally 1D structure.

Figure 9 shows the evolution of the predicted  $V_S$  model with offset marching. In particular, the models at the last generation of the second, third, and fourth offset-marching phases are shown in Figure 9a–9c, respectively. The predicted model in Figure 9a is already very similar to the shallow part of the final model in Figure 7, down to a depth of approximately 4 m. Figure 9b indicates that

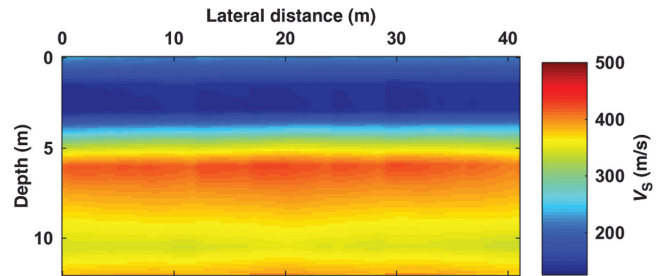


Figure 8. The mean of the  $V_S$  models at the last generation of the Luni data inversion, in which offset marching has been applied. This model is encouragingly quite similar to the model shown in Figure 7.

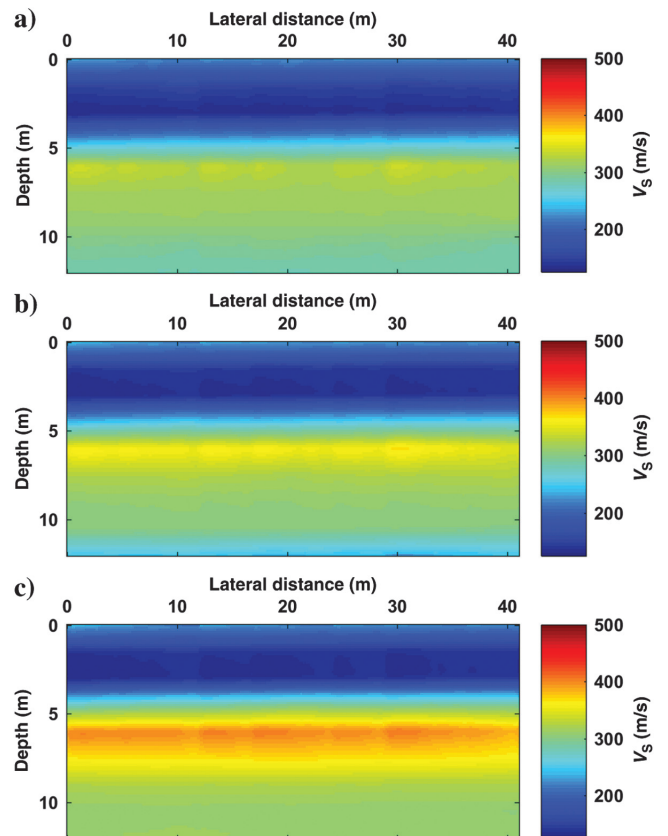


Figure 9. The predicted models at the last generation of the second, third, and fourth offset-marching phase in the Luni data inversion. The inverted traces were within the absolute offset ranges of (a) 6.5–15 m, (b) 6.5–21 m, and (c) 6.5–30 m.

extending the offset range of the data to 21 m, the inversion has reconstructed deeper parts (down to 6 m) of the model. Furthermore, extending the offset range of the traces considered in the inversion to 30 m (Figure 9c) further deepened the reconstructed model to 7.5 m. That is, the offset marching approach from near offset to far offset progressively reconstructs deeper parts of the model.

The predicted seismograms (the red traces), overlaid on the observed data (the black traces), are displayed in Figure 10. The blue lines in the figure roughly separate the bottom part, which has been quite fairly predicted, from the top part in which the data prediction was not satisfactory. In fact, the top part of the predicted data shows extremely small amplitudes compared with the observed data

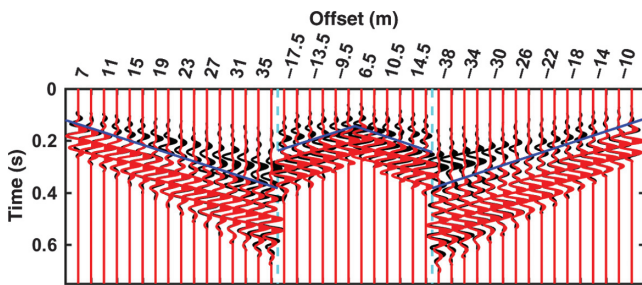


Figure 10. The predicted (red) and observed (black) Luni data in the inversion, in which offset marching has been applied. The dashed cyan lines delimit the left, middle, and right shot gathers. The seismograms have been trace-by-trace normalized. According to the dispersion spectra that will be presented from Figure 11 to Figure 13, waveforms mainly associated with the fundamental mode, which is primarily visible below the blue lines, have been fairly predicted. The prediction of the events above the blue lines has not been as satisfactory, particularly for the too-low amplitudes.

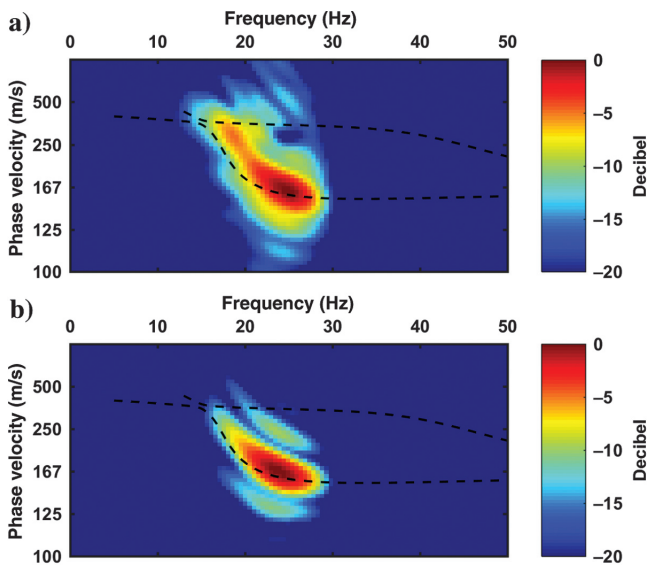


Figure 11. The dispersion spectra of (a) the observed and (b) the predicted Luni data presented in Figure 8. The dashed black curves represent the dispersion curves of the fundamental and the first higher mode computed on the 1D model obtained as the mean of the predicted 2D Luni model shown in Figure 7. As shown in the figure, most of the energy in the observed and predicted data is related to the fundamental mode.

amplitudes. To search for a possible explanation of this mismatch, we carried out an analysis of the dispersion spectra of the whole seismograms and of the two parts separately. The dispersion spectra were computed using the linear Radon transform followed by the temporal Fourier transform.

Sequentially from Figures 11a, 12, and 13b, the dispersion spectra are shown for the whole observed data, the whole predicted data, the bottom part of the observed data, the bottom part of the predicted data, the top part of the observed data, and the top part of the predicted data. The decibel scale is used to represent the amplitudes. The dashed black curves superimposed on these figures correspond to the dispersion curves of the fundamental and the first higher mode computed on the pseudo-1D  $V_S$  model obtained as the mean of the predicted  $V_S$  model of Figure 7.

Figure 11a suggests that the main energy in the whole observed data corresponds to the fundamental mode. The energy related to the first higher mode is also present but with lower amplitudes. Figure 11b shows that our inversion mainly predicted the fundamental-mode waveforms, whereas the predicted amplitudes of the first higher mode were so low as to be invisible in this plot. Using Figure 11a as the reference, Figure 12a and 12b confirms that almost all of the energy in the bottom part of the observed data coincides with the fundamental mode, and that has been fairly reproduced by the bottom part of the predicted data. Coming to the analysis of the top parts, Figure 13a reveals that most of the energy of the observed data is associated with the first higher mode, although some energy of the fundamental mode is present but it is less dominant. On the contrary, as unveiled in Figure 13b, only part of the energy connected with the first higher mode is predicted and some energy predicted by the inversion still lies around the fundamental mode.

In summary, the analysis of the dispersion spectra tells us two facts about the results of the Luni data inversion. First, we have

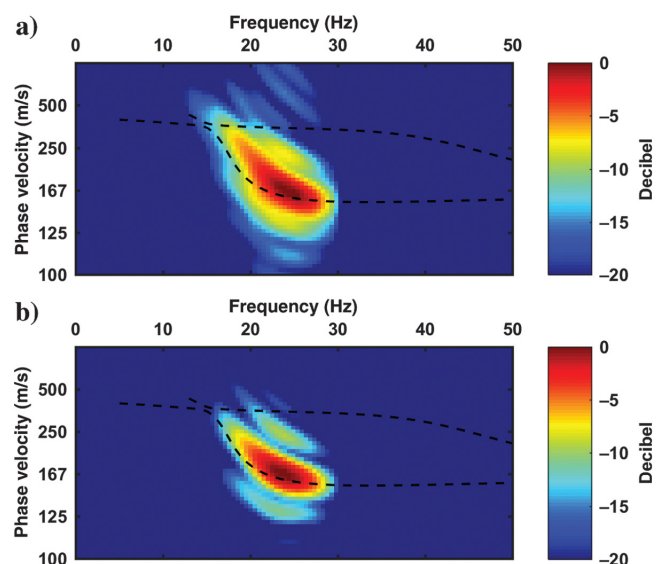


Figure 12. The dispersion spectra of (a) the observed and (b) the predicted data below the blue lines of Figure 10. As in Figure 11, the dashed black curves represent the fundamental and the first higher mode. As shown in the figure, the portions of the observed and predicted seismograms mainly correspond to the fundamental mode, which has been fairly predicted.

fairly predicted the seismogram portions that approximately correspond to the fundamental mode in the observed data. Second, the mismatched waveforms in the top parts of Figure 10 are mostly related to the higher modes. These mismatches are mainly amplitude mismatches caused by the too-low amplitudes of the predicted waveforms, although phase mismatches are also present.

The same Luni data have been used to perform other inversion tests, changing the inversion strategy but keeping constant the inversion grid and all the other controlling parameters of the genetic algorithm. For instance, in one test, we applied offset and frequency marching (for instance, as is done with synthetic data, by Masoni et al., 2016). The predicted model and data are presented in Figures 14 and 15, respectively. They are very similar to the already presented inversion results in Figures 7 and 10. In another test, we split the observed data along the blue lines in Figure 10 into a top part, containing the higher velocity events, and a bottom part, including the lower velocity events, and the two parts were separately included in the objective function of equation 1 and differently weighed. Again, we obtained very similar results. At last, it was

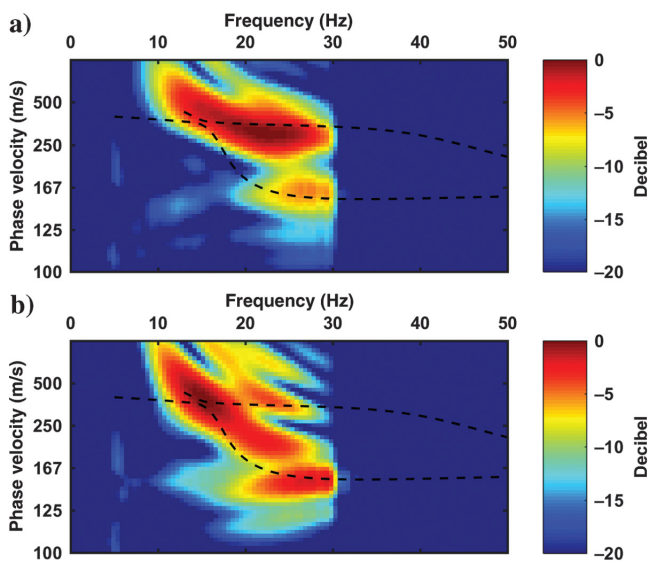


Figure 13. The dispersion spectra of (a) the observed and (b) the predicted data above the blue lines of Figure 10. The dashed black curves are the same as those shown in Figures 11 and 12. As shown in the figure, most of the energy in the observed data corresponds to the first higher mode, which has been only partially reproduced by the predicted data.

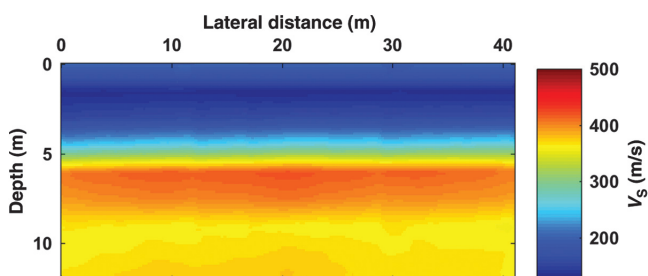


Figure 14. The model predicted from the Luni data inversion, in which the frequency and the offset marching have been applied. This result is similar to that displayed in Figure 7.

never possible to satisfactorily match both parts, and the bottom part, mostly dominated by the fundamental mode, was always better matched than the top part, resulting in very similar model and data predictions.

One possible explanation for the much lower quality of the prediction of the higher modes energy is the too-coarse grid used in the inversion, which only allows for the reconstruction of a smooth model. In fact, as shown by many authors (among others, Yilmaz, 2015), sharp velocity contrasts lead to strong dispersion and high-energy higher modes, which cannot be predicted by using a coarse inversion grid.

However, the smooth model of Figure 7 fairly explains large parts of the observed waveforms, particularly those associated with the fundamental mode. A further gradient-based Rayleigh-wave FWI starting from the model in Figure 7 might render more details of the Luni area. In fact, based on synthetic tests we carried out making use of IFOS2D, an FWI code made available by the toolbox for applied seismic tomography project (Bohlen, 2002; Köhn et al., 2012; Groos et al., 2014), even in the case in which the genetic algorithm FWI has already attained a very good matching between the observed and predicted data, gradient-based FWI, because of the use of a much finer inversion grid, to the inclusion of higher frequencies in the observed data and to its much more rapid convergence, is able to increase the details and to sharpen the contrasts.

## GRENOBLE DATA INVERSION

### Grenoble data acquisition and processing

The Grenoble seismic data set was acquired in September 2013 in the framework of the InterPACIFIC project (Garofalo et al., 2016a, 2016b). Borehole data were also acquired, which now are a reference to check our inversion results. The data acquisition was carried out on a flat topographic surface using 48 vertical-component geophones with 4.5 Hz natural frequency, laid out inline at 1 m intervals. The energy source was an 8 kg sledgehammer. Analogously to the previous example, we used three shot gathers in the inversion: One was split-spread, and two were off-end. The maximum offset of the split-spread and the off-end shot gathers are 23.5 and 51 m, respectively. The three shot gathers are shown in Figure 16. Different from the Luni data, no visible air-blast event is present; thus, no specific processing for its removal is needed. Therefore, the

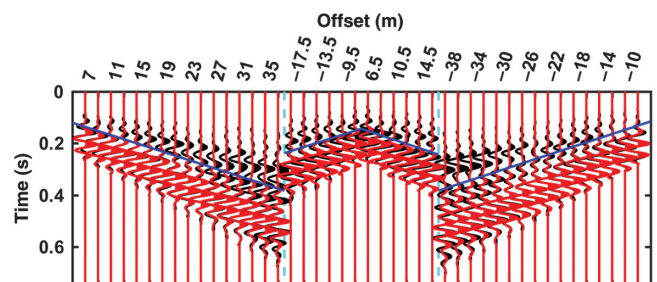


Figure 15. The data predicted (the red traces) from the Luni data inversion, in which the frequency and the offset marching have been applied. The black traces are the observed data. The dashed cyan lines delimit the left, middle, and right shot gathers. The blue lines have been drawn manually to separate the bottom part that has been well-predicted from the top part that has not been predicted as satisfactorily. The seismograms have been normalized trace by trace. This result is similar to that displayed in Figure 10.



processing that we carried out included: geometry assignment; band-pass filtering with corner frequencies 5, 8, 20, and 30 Hz; top and bottom muting to remove the pre-firstbreak noise and to isolate the Rayleigh wavetrain; and 3D to 2D correction. The reasons that lead to this choice of the band-pass filter were the same as in the Luni case; that is, no coherent Rayleigh waves were evident below 5 Hz and we kept the highest frequency to 30 Hz to limit the forward-modeling computing time.

The processed data are displayed in Figure 17. Compared with the processed Luni data (Figure 2), the dispersion effects are much less pronounced and the Rayleigh wavetrain is more compact.

### Grenoble wavelet estimation

The wavelet estimation process was the same as that in the Luni case. We first estimated a wavelet from each trace of the seismic data via the Wiener method. Then, the final wavelet to be used in the inversion was obtained as the weighted mean of the wavelets estimated trace by trace. Same as the Luni case, the weights honored more the near-offset traces. The final wavelet is displayed in Figure 18, and the same comments made for the Luni data wavelet estimation do apply to this case as well.

### Grenoble data inversion specifications

Like the Luni case, no a priori information was exploited in the inversion for the  $V_S$ ,  $V_P$ , and  $\rho$  values at the nodes of the coarse grid. The search range for each unknown was set constant with depth and with lateral distance and was wide enough to include what we thought were reasonable values. Sequentially for  $V_S$ ,

$V_P$ , and  $\rho$ , the search range was 200–500 m/s, 600–2000 m/s, and 1500–1920 kg/m<sup>3</sup>.

Based on the minimum velocity defined in the search range and on the maximum frequency to be modeled, we constructed a fine modeling grid of 26,015 nodes, corresponding to a grid spacing of 0.25 m, which guaranteed to include 20 points per minimum wavelength in the finite difference (FD) computation. Under the two-grid scheme, we built a coarse inversion grid containing 65 nodes, resulting in  $65 \times 3 = 195$  unknowns. Because no a priori information was used to suggest the velocity variation, a rectangular 13.375–2.5 m inversion grid was constructed. The large spacing of the coarse grid again implies that the sought  $V_S$  models would only reproduce the smooth, long-wavelength structure of the investigated zone.

Given the relatively “simple” waveforms in the observed Grenoble data, compared to those of the Luni example, in the inversion, we neither considered envelope traces nor applied offset marching. Thus, the data misfits were calculated just using the L1 norm of the observed and the predicted waveforms; that is, we used only the first term of equation 1. Frequency marching was adopted with the following scheme: We started the inversion from 5 to 15 Hz band-pass filtered data and then every 40 generations, a band of 5 Hz was added to the inverted data until the frequency band expanded from 5 to 30 Hz. A final additional run of 40 generations was carried out with the full-band 5–30 Hz data.

The settings of the genetic-algorithm control parameters were exactly the same as those given in the Luni case. In particular, the maximum number of generations was 200, and the number of individuals per generation was 2000.

### Grenoble data inversion results

Figure 19 shows the evolution of the minimum and mean data misfit. Similar to the Luni case, comparing the data misfits calculated in different frequency-marching phases has no use, due to the change of the normalization term in the objective function after the addition of new frequencies.

Inside each frequency-marching phase, the decreases in the data misfits are evident. In the last generation, the minimum and the corresponding mean data misfits are very close, and the decreases in the data misfits are quite slow. It is mainly due to the loss of the gene

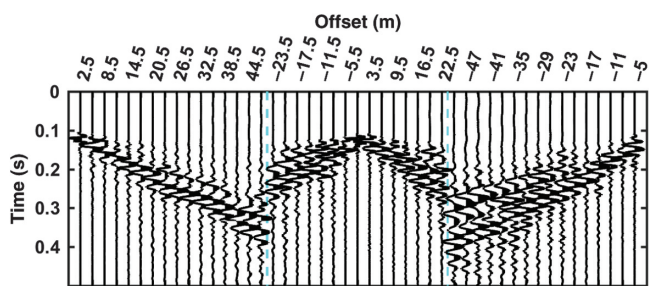


Figure 16. Raw seismic data acquired in Grenoble, France. One of every three traces is plotted. The dashed cyan lines delimit the left, middle, and right shot gathers. The seismograms have been trace-by-trace normalized.

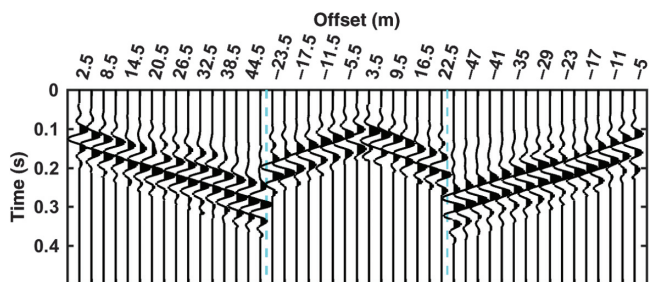


Figure 17. The Grenoble field data of Figure 16 after processing and the 3D to 2D correction. The maximum frequency is 30 Hz. The dashed cyan lines delimit the left, middle, and right shot gathers. The seismograms have been trace-by-trace normalized.

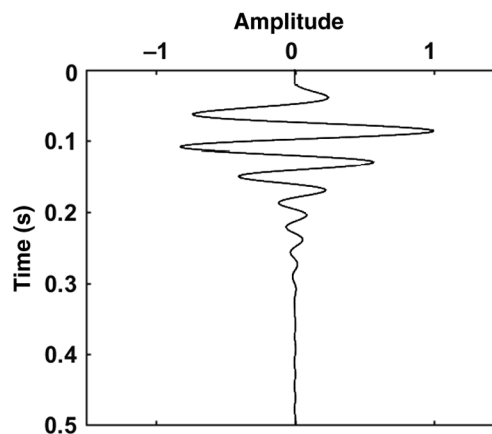


Figure 18. The wavelet estimated from the Grenoble field data using the Wiener method.



variety and any further evolution allowing more generations would not be efficient.

The predicted  $V_S$  model is displayed in Figure 20b. It shows a profile mostly characterized by vertical velocity variations and

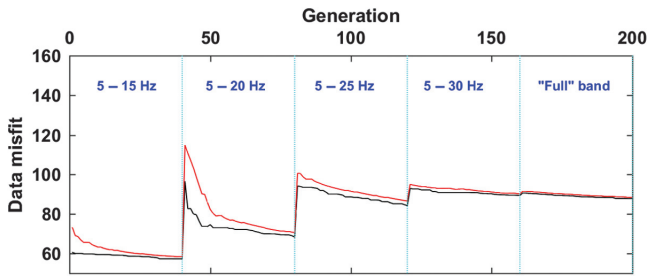


Figure 19. The evolution of the minimum (black) and mean (red) data misfit in the Grenoble data inversion. The annotations in blue bounded by the dashed cyan lines indicate the frequency bands of the inverted data in the frequency-marching scheme. In this case, the data misfits also decrease in each frequency-marching step.

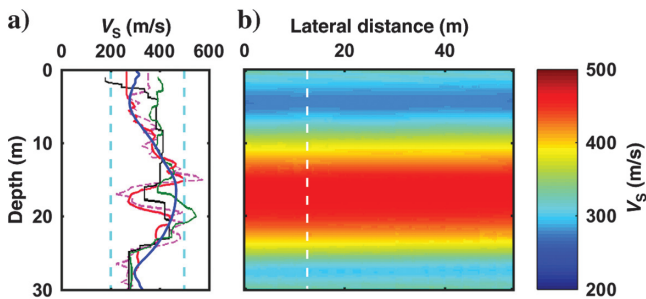


Figure 20. (a) The predicted  $V_S$  profile (the blue curve) whose position is the nearest to the boreholes (the dashed white line in [b]).  $V_S$  search ranges in the two-grid genetic-algorithm FWI (the dashed cyan lines). The other curves were taken from Garofalo et al. (2016b) and show  $V_S$  values derived from borehole log measurements. The black curve represents downhole recordings. The green curve represents crosshole recordings. The red curve represents P-S suspension logging through the analysis of the direct traveltime between the source and the first geophone. The dashed magenta curve represents P-S suspension logging through the analysis of the differential traveltime between the first and the second geophone. (b) The 2D  $V_S$  model predicted from the two-grid genetic-algorithm FWI.

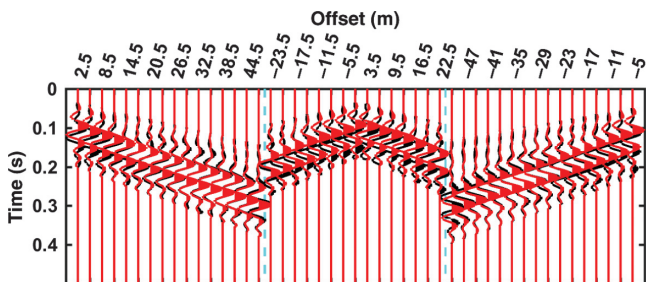


Figure 21. The observed (black) and predicted (red) Grenoble data. The dashed cyan lines delimit the left, middle, and right shot gathers. The seismograms have been trace-by-trace normalized. We obtained a general fair match, and where mismatches occur, the differences are within half-periods.

some velocity inversions. Extracting the  $V_S$  profile nearest to the boreholes gives the blue curve presented in Figure 20a. In the same figure, the dashed cyan lines indicate the  $V_S$  search range in our inversion. The other curves exhibit the  $V_S$  values derived from the borehole log measurements (refer to the caption of Figure 20a). As is clearly shown in the figure, the  $V_S$  curve from our final model fairly matches the long-wavelength profiles of the  $V_S$  logs, especially the P-S suspension logging curves. A layer at approximately 16 m, which the logs indicate as being characterized by a decrease in the velocity, was not detected by our inversion because of the limited resolution due to the adopted coarse inversion grid.

Figure 21 exhibits the predicted data (red traces) overlaid on the observed seismograms (black traces). There is an overall fair match between the predicted and observed waveforms. When mismatches occur, the phase differences are always within half-periods. This indicates that the  $V_S$  model predicted by our two-grid genetic-algorithm FWI should be an adequate initial model for gradient-based Rayleigh-wave FWI for further refinement.

## CONCLUSION

The two-grid genetic-algorithm FWI of Rayleigh waves that we developed in this and in the companion paper is aimed at retrieving a reliable macromodel of the subsurface  $V_S$  in absence of any a priori information. The model to be inverted for is parameterized by a coarse grid, and the resolution of the estimated model mainly depends on the spacing of its nodes, which also determines the number of unknowns to be inverted for: The shorter the spacing and the higher the resolution, the higher the computing time that is needed to achieve convergence. In general, using conventional computer resources, to maintain acceptable computing time, we have designed grids with quite large spacings resulting in numbers of unknowns up to approximately 200; thus, the retrieved models necessarily contain the long-wavelength structures only of the velocity field. More powerful computing resources and, we may add, an optimized coding of the computer program, should allow for the inversion of a greater number of unknowns and thus for the achievement of a higher resolution of the estimated velocity model.

The approach we propose can consider laterally varying velocities, nonflat topography, and it does not require any prior knowledge of the subsurface. In the inversion of the two actual data cases that we have presented, the boundaries of genetic-algorithm search ranges were constant with depth and with the horizontal direction, a situation that corresponds to no available a priori information on the specific survey area.

In both cases, we obtained fair model and data predictions using a quite coarse inversion grid and limiting the frequencies to 30 Hz.

The first test with the Luni data has been the one in which we spent the largest efforts because we tried to reach a good data matching not only of the waveforms pertaining to the fundamental mode but also of the higher velocity wavetrain likely related to the higher modes. Thus, many tests were carried out using different inversion schemes and different wavelets, of which some have been discussed here. In all of these trials, we have never been able to simultaneously and satisfactorily match the whole observed seismograms: The lower velocity sectors of the seismograms were generally better matched than the high-velocity parts, in which the predicted waveforms showed distinctly lower amplitudes than the observed waveforms. According to our analyses of the dispersion spectra, the better matched waveforms correspond to the fundamental mode and

to a small portion of, likely, the first higher mode. However, the predicted velocity models were always very similar, indicating nearly the same velocity layering. The stability of the retrieved velocity models induced us to attribute to the coarseness of the inversion grid, and the consequent inability to retrieve rapid velocity contrasts, the likely cause of the partial mismatch of the higher mode components of the seismograms. It would be then left to a local FWI algorithm, which launches the inversion from our predicted model and uses a much finer parameterization of the model space, the task to improve the match of the whole seismogram. A task that, based on other experiences with synthetic data, seems quite feasible for gradient-based inversion.

The second example with the Grenoble data benefited from the availability of borehole logs, which measured the  $V_S$  profile and allowed for a comparison with the model predicted by our inversion. In this case, the degree of dispersion of the recorded Rayleigh waves was less prominent than in the previous example; thus, it was easier to find a fair data matching by our two-grid genetic-algorithm FWI. The predicted model shows mainly vertical velocity variations. The frequency-marching scheme produced results that, frequency step after frequency step, progressively looked more similar to the borehole  $V_S$  stratigraphy, particularly that revealed by the P-S suspension log. A layer characterized by a velocity decrease was not detected by our inversion again due to the inherent low resolution associated with the coarse inversion grid. However, the  $V_S$  profile, extracted from the predicted 2D model, nearest to the boreholes well agreed with the main features of the velocity trend indicated by the logs. In addition, the inversion achieved an acceptable match between predicted and observed traces and when mismatches occurred; they were always within half-periods. As also mentioned in our companion paper, several further developments and improvements of our method may be tried, from the estimation of uncertainties of the solution, to the increase of the efficiency of the inversion code, to the testing on other synthetic and actual data cases. Specifically, concerning the application to actual data, the estimation of the wavelet to give as input to the inversion or whether the search for improved updates of an initial wavelet could be embedded in the inversion are issues worthy of further study.

## ACKNOWLEDGMENTS

The revision process of this paper has seen a particularly positive scientific interaction with knowledgeable reviewers: We wish to thank them for their suggestions and constructive comments. We gratefully acknowledge the support of Landmark/Halliburton for the use of their ProMAX seismic software for the data processing. We also thank F. Garofalo for having made us aware of the availability of the data of the InterPACIFIC project.

## DATA AND MATERIALS AVAILABILITY

Data associated with this research are available and can be obtained by contacting the corresponding author.

## REFERENCES

- Aleardi, M., and A. Mazzotti, 2017, 1D elastic full-waveform inversion and uncertainty estimation by means of a hybrid genetic algorithm-Gibbs sampler approach: *Geophysical Prospecting*, **65**, 64–85, doi: [10.1111/1365-2478.12397](https://doi.org/10.1111/1365-2478.12397).
- Bellman, R. E., 1957, *Dynamic programming*: Princeton University Press.
- Bohlen, T., 2002, Parallel 3-D viscoelastic finite difference seismic modeling: *Computers and Geosciences*, **28**, 887–899, doi: [10.1016/S0098-3004\(02\)00066-7](https://doi.org/10.1016/S0098-3004(02)00066-7).
- Bohlen, T., S. Kugler, G. Klein, and F. Theilen, 2004, 1.5D inversion of lateral variation of Scholte-wave dispersion: *Geophysics*, **69**, 330–344, doi: [10.1190/1.1707052](https://doi.org/10.1190/1.1707052).
- Brossier, R., S. Operto, and J. Virieux, 2009, Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion: *Geophysics*, **74**, no. 6, WCC105–WCC118, doi: [10.1190/1.3215771](https://doi.org/10.1190/1.3215771).
- Bunks, C., F. M. Saleck, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: *Geophysics*, **60**, 1457–1473, doi: [10.1190/1.1443880](https://doi.org/10.1190/1.1443880).
- Cercato, M., 2009, Addressing non-uniqueness in linearized multichannel surface wave inversion: *Geophysical Prospecting*, **57**, 27–47, doi: [10.1111/j.1365-2478.2007.00719.x](https://doi.org/10.1111/j.1365-2478.2007.00719.x).
- Cercato, M., 2011, Global surface wave inversion with model constraints: *Geophysical Prospecting*, **59**, 210–226, doi: [10.1111/j.1365-2478.2010.00922.x](https://doi.org/10.1111/j.1365-2478.2010.00922.x).
- Chi, B. X., L. G. Dong, and Y. Z. Liu, 2014, Full waveform inversion method using envelope objective function without low frequency data: *Journal of Applied Geophysics*, **109**, 36–46, doi: [10.1016/j.jappgeo.2014.07.010](https://doi.org/10.1016/j.jappgeo.2014.07.010).
- Claerbout, J. F., 1985, *Fundamentals of geophysical data processing*: Blackwell Scientific Publications.
- Dokter, E., D. Köhn, D. Wilken, D. De Nil, and W. Rabbel, 2017, Full waveform inversion of SH- and love-wave data in near-surface prospecting: *Geophysical Prospecting*, **65**, 216–236, doi: [10.1111/1365-2478.12549](https://doi.org/10.1111/1365-2478.12549).
- Forbriger, T., 2003a, Inversion of shallow-seismic wavefields — 1. Wavefield transformation: *Geophysical Journal International*, **153**, 719–734, doi: [10.1046/j.1365-246X.2003.01929.x](https://doi.org/10.1046/j.1365-246X.2003.01929.x).
- Forbriger, T., 2003b, Inversion of shallow-seismic wavefields — 2. Inferring subsurface properties from wavefield transforms: *Geophysical Journal International*, **153**, 735–752, doi: [10.1046/j.1365-246X.2003.01985.x](https://doi.org/10.1046/j.1365-246X.2003.01985.x).
- Forbriger, T., L. Groos, and M. Schäfer, 2014, Line-source simulation for shallow-seismic data — Part 1: Theoretical background: *Geophysical Journal International*, **198**, 1387–1404, doi: [10.1093/gji/ggu199](https://doi.org/10.1093/gji/ggu199).
- Garofalo, F., S. Foti, F. Hollender, P. Y. Bard, C. Cornou, B. R. Cox, A. Dechamp, M. Ohrnberger, V. Perron, D. Sicilia, D. Teague, and C. Vergnault, 2016b, InterPACIFIC project: Comparison of invasive and non-invasive methods for seismic site characterization — Part 2: Inter-comparison between surface-wave and borehole methods: *Soil Dynamics and Earthquake Engineering*, **82**, 241–254, doi: [10.1016/j.soildyn.2015.12.009](https://doi.org/10.1016/j.soildyn.2015.12.009).
- Garofalo, F., S. Foti, F. Hollender, P. Y. Bard, C. Cornou, B. R. Cox, M. Ohrnberger, D. Sicilia, M. Asten, G. Di Giulio, T. Forbriger, B. Guillier, K. Hayashi, A. Martin, S. Matsushima, D. Mercerat, V. Poggi, and H. Yamanaka, 2016a, InterPACIFIC project: Comparison of invasive and non-invasive methods for seismic site characterization — Part 1: Intra-comparison of surface wave methods: *Soil Dynamics and Earthquake Engineering*, **82**, 222–240, doi: [10.1016/j.soildyn.2015.12.010](https://doi.org/10.1016/j.soildyn.2015.12.010).
- Groos, L., M. Schäfer, T. Forbriger, and T. Bohlen, 2014, The role of attenuation in 2D full-waveform inversion of shallow-seismic body and Rayleigh waves: *Geophysics*, **79**, no. 6, R247–R261, doi: [10.1190/GEO2013-0462.1](https://doi.org/10.1190/GEO2013-0462.1).
- Groos, L., M. Schäfer, T. Forbriger, and T. Bohlen, 2017, Application of a complete workflow for 2D elastic full-waveform inversion to recorded shallow-seismic Rayleigh waves: *Geophysics*, **82**, no. 2, R109–R117, doi: [10.1190/geo2016-0284.1](https://doi.org/10.1190/geo2016-0284.1).
- Köhn, D., D. De Nil, A. Kurzmann, A. Przebindowska, and T. Bohlen, 2012, On the influence of model parametrization in elastic full waveform tomography: *Geophysical Journal International*, **191**, 325–345, doi: [10.1111/j.1365-246X.2012.05633.x](https://doi.org/10.1111/j.1365-246X.2012.05633.x).
- Mallick, S., 1995, Model-based inversion of amplitude-variations-with-offset data using a genetic algorithm: *Geophysics*, **60**, 939–954, doi: [10.1190/1.1443860](https://doi.org/10.1190/1.1443860).
- Mallick, S., 1999, Some practical aspects of prestack waveform inversion using a genetic algorithm: An example from the East Texas Woodbine gas sand: *Geophysics*, **64**, 326–336, doi: [10.1190/1.1444538](https://doi.org/10.1190/1.1444538).
- Maraschini, M., F. Ernst, S. Foti, and L. V. Socco, 2010, A new misfit function for multimodal inversion of surface waves: *Geophysics*, **75**, no. 4, G31–G43, doi: [10.1190/1.3436539](https://doi.org/10.1190/1.3436539).
- Masoni, I., J. L. Boelle, R. Brossier, and J. Virieux, 2016, Layer stripping FWI for surface waves: 86th Annual International Meeting, SEG, Expanded Abstracts, 1369–1373, doi: [10.1190/segam2016-13859781.1](https://doi.org/10.1190/segam2016-13859781.1).
- Masoni, I., R. Brossier, J. Virieux, and J. L. Boelle, 2014, Robust full waveform inversion of surface waves: 84th Annual International Meeting, SEG, Expanded Abstracts, 1126–1130, doi: [10.1190/segam2014-1077.1](https://doi.org/10.1190/segam2014-1077.1).
- Mazzotti, A., N. Bienati, E. Stucchi, A. Tognarelli, M. Aleardi, and A. Sajeve, 2017, Two-grid genetic algorithm full-waveform inversion: The Leading Edge, **35**, 1068–1075, doi: [10.1190/le35121068.1](https://doi.org/10.1190/le35121068.1).

- Park, C. B., R. D. Miller, and J. H. Xia, 1999, Multichannel analysis of surface waves: *Geophysics*, **64**, 800–808, doi: [10.1190/1.1444590](https://doi.org/10.1190/1.1444590).
- Rix, G. J., 2004, Near-surface site characterization using surface waves, in C. G. Lai and K. Wilmański, eds., *Surface waves in geomechanics: Direct and inverse modeling for soils and rocks*: CISM International Centre for Mechanical Sciences 481, 1–46.
- Sajeva, A., M. Aleardi, A. Mazzotti, N. Bienati, and E. Stucchi, 2014, Estimation of velocity macro-models using stochastic full-waveform inversion: 84th Annual International Meeting, SEG, Expanded Abstracts, 1227–1231, doi: [10.1190/segam2014-1088.1](https://doi.org/10.1190/segam2014-1088.1).
- Sajeva, A., M. Aleardi, E. Stucchi, N. Bienati, and A. Mazzotti, 2016, Estimation of acoustic macro models using a genetic full-waveform inversion: Applications to the Marmousi model: *Geophysics*, **81**, no. 4, R173–R184, doi: [10.1190/geo2015-0198.1](https://doi.org/10.1190/geo2015-0198.1).
- Schäfer, M., L. Groos, T. Forbriger, and T. Bohlen, 2013, 2D full waveform inversion of recorded shallow seismic Rayleigh waves on a significantly 2D structure: Near Surface Geoscience Technical Program, Expanded Abstracts, doi: [10.3997/2214-4609.20131338](https://doi.org/10.3997/2214-4609.20131338).
- Schäfer, M., L. Groos, T. Forbriger, and T. Bohlen, 2014, Line-source simulation for shallow-seismic data — Part 2: Full-waveform inversion — A synthetic 2-D case study: *Geophysical Journal International*, **198**, 1405–1418, doi: [10.1093/gji/ggu171](https://doi.org/10.1093/gji/ggu171).
- Sen, M. K., and P. L. Stoffa, 1992, Rapid sampling of model space using genetic algorithms: Examples from seismic waveform inversion: *Geophysical Journal International*, **108**, 281–292, doi: [10.1111/j.1365-246X.1992.tb00857.x](https://doi.org/10.1111/j.1365-246X.1992.tb00857.x).
- Socco, L. V., S. Foti, and D. Boiero, 2010, Surface-wave analysis for building near-surface velocity models — Established approaches and new perspectives: *Geophysics*, **75**, no. 5, 75A83–75A102, doi: [10.1190/1.3479491](https://doi.org/10.1190/1.3479491).
- Socco, L. V., and C. Strobbia, 2004, Surface-wave method for near-surface characterisation: A tutorial: *Near Surface Geophysics*, **2**, 165–185, doi: [10.3997/1873-0604.2004015](https://doi.org/10.3997/1873-0604.2004015).
- Stoffa, P. L., and M. K. Sen, 1991, Nonlinear multiparameter optimization using genetic algorithms: Inversion of plane-wave seismograms: *Geophysics*, **56**, 1794–1810, doi: [10.1190/1.1442992](https://doi.org/10.1190/1.1442992).
- Thorbecke, J. W., and D. Draganov, 2011, Finite-difference modeling experiments for seismic interferometry: *Geophysics*, **76**, no. 6, H1–H8, doi: [10.1190/geo2010-0039.1](https://doi.org/10.1190/geo2010-0039.1).
- Tran, K. T., M. McVay, M. Faraone, and D. Horhota, 2013, Sinkhole detection using 2D full seismic waveform tomography: *Geophysics*, **78**, no. 5, R175–R183, doi: [10.1190/geo2013-0063.1](https://doi.org/10.1190/geo2013-0063.1).
- Ulrych, T. J., S. Freire, and P. Siston, 1988, Eigenimage processing of seismic sections: 58th Annual International Meeting, SEG, Expanded Abstracts, 1261–1265, doi: [10.1190/1.1892508](https://doi.org/10.1190/1.1892508).
- Wu, R. S., J. R. Luo, and B. Y. Wu, 2014, Seismic envelope inversion and modulation signal model: *Geophysics*, **79**, no. 3, WA13–WA24, doi: [10.1190/GEO2013-0294.1](https://doi.org/10.1190/GEO2013-0294.1).
- Xia, J. H., R. D. Miller, and C. B. Park, 1999, Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves: *Geophysics*, **64**, 691–700, doi: [10.1190/1.1444578](https://doi.org/10.1190/1.1444578).
- Xing, Z., and A. Mazzotti, 2016, Rayleigh wave modeling complexities in the perspective of full waveform inversion of surface waves — Synthetic examples: Near Surface Geoscience Technical Program, Expanded Abstracts, doi: [10.3997/2214-4609.201601909](https://doi.org/10.3997/2214-4609.201601909).
- Xing, Z., and A. Mazzotti, 2017a, Two-grid full waveform Rayleigh wave inversion by means of genetic algorithm with frequency marching: 79th Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: [10.3997/2214-4609.201701412](https://doi.org/10.3997/2214-4609.201701412).
- Xing, Z., and A. Mazzotti, 2017b, Surface wave FWI on complex models: The robustness of the inversion to assumptions and forward modeling approximations: Near Surface Geoscience Technical Program, Expanded Abstracts, doi: [10.3997/2214-4609.201702019](https://doi.org/10.3997/2214-4609.201702019).
- Xing, Z., and A. Mazzotti, 2018, Two-grid genetic algorithm full waveform inversion of surface waves: Two actual data examples: 80th Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: [10.3997/2214-4609.201801375](https://doi.org/10.3997/2214-4609.201801375).
- Xing, Z., and A. Mazzotti, 2019, Two-grid full-waveform Rayleigh-wave inversion via a genetic algorithm — Part 1: Method and synthetic examples: *Geophysics*, **84**, this issue, doi: [10.1190/geo2018-0799.1](https://doi.org/10.1190/geo2018-0799.1).
- Yilmaz, Ö., 2015, *Engineering seismology with applications to geotechnical engineering*: SEG.