Two-grid full-waveform Rayleigh-wave inversion via a genetic algorithm — Part 1: Method and synthetic examples

Zhen Xing[1] and Alfredo Mazzotti[2]

ABSTRACT

When reliable a priori information is not available, it is difficult to correctly predict near-surface S-wave velocity models from Rayleigh waves through existing techniques, especially in the case of complex geology. To tackle this issue, we have developed a new method: two-grid genetic-algorithm Rayleigh-wave full-waveform inversion (FWI). Adopting a two-grid parameterization of the model, the genetic algorithm inverts for unknown velocities and densities at the nodes of a coarse grid, whereas the forward modeling is performed on a fine grid to avoid numerical dispersion. A bilinear interpolation brings the coarse-grid results into the fine-grid models. The coarse inversion grid allows for a significant reduction in the computing time required by the genetic algorithm to converge. With a coarser grid, there are fewer unknowns and less required computing time, at the expense of the model resolution. To further increase efficiency, our inversion code can perform the optimization using an offset-marching strategy and/or a frequency-marching strategy that can make use of different kinds of objective functions and allows for parallel computing. We illustrate the effect of our inversion method using three synthetic examples with rather complex near-surface models. Although no a priori information was used in all three tests, the long-wavelength structures of the reference models were fairly predicted, and satisfactory matches between "observed" and predicted data were achieved. The fair predictions of the reference models suggest that the final models estimated by our genetic-algorithm FWI, which we call macromodels, would be suitable inputs to gradient-based Rayleigh-wave FWI for further refinement. We also explored other issues related to the practical use of the method in different work and explored applications of the method to field data.

INTRODUCTION

Multichannel analysis of surface waves (MASW) (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) is the current standard for Rayleigh-wave inversion, in which the observed data are given as the dispersion curves extracted from frequency-wavenumber (f-k) or frequency-slowness (f-p) spectra. However, over the past few years, a new approach, full-waveform inversion (FWI) of Rayleigh waves, has emerged (Schäfer et al., 2013; Tran et al., 2013; Groos et al., 2014; Masoni et al., 2014, 2016). This new approach seems to be promising because of several advantages over conventional techniques. First, instead of exploiting only the dispersion curves of the fundamental and higher modes, it makes use of the entire information (traveltime, amplitude, and phase) present in the recorded Rayleigh waves. Second, it naturally supports the predictions of multidimensional models. Third, it requires no subjective interpretation, such as the picking of dispersion curves on the f-p spectra of the observed data.

Nevertheless, the application of FWI to Rayleigh-wavefield data is still rare. Schäfer et al. (2013) invert Rayleigh-wave data to study a vertical fault system near Frankfurt, Germany, but they limit the inverted data within 10 Hz because adding higher frequencies would lead to local minima. Tran et al. (2013) use an FWI approach for sinkhole detection in Florida, United States. Groos et al. (2017) obtain reasonable results by applying FWI to a Rayleigh-wave data set acquired at a gliding airfield near Karlsruhe, Germany.

[1]Formerly University of Pisa, Department of Earth Sciences, Pisa, Italy; presently CGG, Crawley, UK. E-mail: zhnxng@outlook.com (corresponding author).
[2]University of Pisa, Department of Earth Sciences, Pisa, Italy. E-mail: alfredo.mazzotti@unipi.it.

However, due to the strong nonlinearity of Rayleigh waves (Forbriger, 2003a, 2003b; Rix, 2004; Brossier et al., 2009; Schäfer et al., 2013), local optimization methods, which require the computation of the gradient of the objective function, need adequate initial models to avoid getting trapped into local minima.

As a general rule, an adequate initial model ought to contain the long-wavelength structures of the investigated near-surface zone. In addition, it should lead to simulated seismograms that limit cycle skipping, particularly in the portions containing the fundamental mode. Unfortunately, in actual data cases such optimal initial models may be difficult to obtain because of the lack of a priori information. To tackle this issue, we propose a new approach of Rayleigh-wave FWI with a global stochastic optimization based on genetic algorithms. Genetic algorithms (Stoffa and Sen, 1991; Sen and Stoffa, 1992; Mallick, 1995, 1999) explore a wide model space to attain inversion outcomes and, thus, they are much less affected than local methods by the presence of local minima. As a result, the importance of finding an adequate initial model is not as crucial as for local, gradient-based optimization methods.

Nonetheless, stochastic methods generally require huge computational resources especially when costly forward modeling is needed, leading to serious limitations in their practical applicability.

To attenuate this problem, we proceed as follows: (1) We parameterize the subsurface by adopting a two-grid strategy (Sajeva et al., 2014, 2016; Aleardi and Mazzotti, 2017; Mazzotti et al., 2017), one coarse grid for the inversion phase and one fine grid for the modeling phase; (2) we make use of an offset-marching scheme and/or a frequency-marching scheme (Bunks et al., 1995) if deemed necessary; and (3) we perform parallel computing (our code has been parallelized with Open MPI — C++). With these strategies adopted, the genetic-algorithm Rayleigh-wave FWI method that we propose can be used to derive near-surface V_S models even in cases in which null a priori information is available.

We illustrate the proposed method starting from a brief description of the finite-difference modeling (FDM) algorithm that we use. Then, we introduce the two-grid strategy and the genetic-algorithm optimization. Finally, we discuss the results of three inversion tests, carried out without making use of any a priori information, on three synthetic examples that reproduce complex near-surface models. All of the inversion tests have been performed assuming elastic wave propagation.

The application of the two-grid genetic-algorithm Rayleigh-wave FWI to two actual data sets, along with additional considerations required for field data inversion, is presented in a companion paper (Xing and Mazzotti, 2019).

METHOD

This section describes the three key parts that construct our method. The first is the reliable simulation of Rayleigh waves. The second illustrates the two-grid scheme, which is quite important for reducing the computational time. The third is the genetic-algorithm workflow.

Rayleigh-wave modeling

The engine that we use for Rayleigh-wave modeling is a time-domain 2D elastic FDM code developed by Thorbecke and Draganov (2011) and further modified by Xing and Mazzotti (2016). We use the second-order approximation of derivatives in time and the fourth-order approximation in space.

A convolutional perfectly matched layer (Roden and Gedney, 2000; Collino and Tsogka, 2001; Festa and Vilotte, 2005; Komatitsch and Martin, 2007; Gedney, 2011) is implemented in the modeling code to attenuate wave energy in the absorbing boundaries. The free-surface condition recommended by Robertsson (1996) is used to simulate wave propagation in the presence of irregular topographic surfaces. To achieve a reliable simulation of Rayleigh waves, we suggest that the number of points (n) per minimum wavelength in the elastic FDM is set to 20 instead of to 5, which is the standard value for modeling body waves (Alford et al., 1974). The suggestion is based on our comparisons among results generated from various modeling codes and on many tests in which n has been increased up to 100. This indication also coincides with that given by Nagai et al. (2005).

The value of 20 assigned to the number of points per minimum wavelength and the fact that V_S at the near surface is usually low, dictate that very fine grids should be used for modeling Rayleigh waves, thus making Rayleigh-wave modeling quite costly.

We checked the reliability of the used FDM code by comparing its simulations on various reference models with the results of other modeling codes on the same models. The modeling results generated by the adopted 2D FDM code match well with seismograms from reflectivity modeling (Fuchs and Muller, 1971), seismograms from spectral element modeling (Komatitsch et al., 2001), and seismograms provided by Eni S.p.A. using another FDM code, confirming the reliability of the used forward-modeling code.

The two-grid scheme

To reach satisfying results, a stochastic (global optimization) approach roughly requires a computing time that is exponentially proportional to the number of unknowns (Bellman, 1957); so does a genetic algorithm. In the FWI method that we propose, the unknowns are the V_S, V_P, and rho at the nodes of the model grid. In the case of actual data (Xing and Mazzotti, 2018), the number of nodes in the modeling grid is usually tens of thousands. Such a huge number of grid nodes, triplicated, determines several unknowns that would require unacceptable computing time for a direct application of the genetic-algorithm Rayleigh-wave FWI.

The solution that we propose to render the number of unknowns workable with the computing resources of standard computers is the adoption of a two-grid scheme. Figure 1 introduces the two-grid strategy. It shows a fine grid (the black net) and the nodes (the magenta dots) of a much coarser grid superimposed on the V_S model pertaining to the third synthetic example that we will discuss in detail later on. Only the portion of interest of the model is shown, but we consider that absorbing boundaries are also present at the borders of the shown model.

The fine grid is used in forward modeling to guarantee the reliable computation of Rayleigh waves; therefore, it follows the spacing criterion in FDM. Instead, the V_S, V_P, and rho at each node of the coarse grid constitute the unknowns, i.e., the total number of unknowns is three times the number of the nodes of the coarse grid. Therefore, it is on earth models parameterized with coarse grids that genetic-algorithm optimization is performed. With coarser grids, the model resolution attained by the inversion is lower and the required computing time is less. Although in Figure 1 the shown coarse grid is regular, it can practically be irregular or even random.

The choice of the coarse-grid dimensions depends on the available a priori information that we may have, particularly on the minimum velocities of the subsurface and on the recorded maximum frequencies that could be recorded. With such information, we could devise a grid whose spacing among the nodes is governed by the expected resolution, as can be roughly estimated making use of rules of thumb, such as that indicated by Park et al. (1999) for dispersion curve inversion. The grid node spacing can even be variable as a function of the (supposed) velocity variations. When no a priori information is available, as is the case that we discuss here, the issue of how to parameterize the coarse grid model is inevitable: Unless a transdimensional approach (Bodin et al., 2012) to the inversion is adopted, i.e., an approach in which the number of unknowns is itself an unknown, the common practical choice is to design a regular grid with a reasonable number of nodes to maintain the computing times acceptable and, possibly, perform a second run with a more sophisticated grid designed on the basis of the provisional results.

We realize the conversion from the coarse-grid model to the correspondent fine-grid model by means of bilinear interpolation. Due to the long spacing among the nodes of the coarse grid and the smoothing effect of the interpolation, we usually obtain smooth predicted models reproducing only the long-wavelength structures of the subsurface. We name such predicted models "macro" models.

Genetic-algorithm FWI with frequency and offset marching

Compared with local optimization methods, global optimization approaches are much less vulnerable to local minima and this is why we choose the genetic algorithm in the context of Rayleigh-wave FWI, which is notoriously a strong nonlinear problem. In addition, different objective functions can be more easily implemented in global optimization methods.

The reason why we choose the genetic algorithm over other global optimization approaches is discussed by Sajeva et al. (2017). They compare the genetic algorithm with adaptive simulated annealing, particle swarm optimization, and the neighborhood algorithm. They find that in the context of analytical objective functions and 1D elastic FWI, the genetic algorithm outperformed the others. Moreover, the genetic algorithm is naturally parallelizable, which is particularly important in our case for speeding up computation.

Genetic algorithms imitate the biological evolution process to realize optimization. Figure 2 displays the workflow of our genetic-algorithm FWI. The mechanism of the inversion is as follows.

The first step is the creation of the initial population, which in our case is an ensemble of V_S, V_P, and density (rho) models. The randomly created models (individuals) are uniformly distributed within predefined search ranges that limit the model space to be explored by the algorithm. The user can establish such search ranges on the basis of deductions on the observed data (e.g., on dispersion spectra) or on some kind of a priori information on the investigated zone, when available.

Next, a synthetic seismogram is computed by forward modeling for each model of the initial

population. Based on a user defined objective function, data misfits between observed data and simulated data are calculated for the evaluation of the model fitness to sift out the promising models that are then paired to produce offspring by combination and mutation. "Combination" here means exchanging part of the values (i.e., velocities and densities at some of the grid nodes) of paired models, whereas "mutation" is randomly changing within the predefined search ranges a small portion of values of selected models.

After that, the models leading to minor data misfits in the offspring are inserted back to the original population to replace the models associated with larger data misfits. At each generation, selection, recombination, mutation, and reinsertion are performed to obtain models with ever-decreasing data misfits. This indicates that, theoretically, the data misfits can always be improved until the ideal one is found. In practice, considering efficiency, a stopping criterion, such as a predefined maximum generation or a threshold on the data misfit, is set to terminate the inversion.

The forward modeling, which is performed between the creation (or reinsertion) of models and the evaluation of the objective function, requires most of the computational time.

As shown in Figure 2, frequency (Bunks et al., 1995) and offset (Masoni et al., 2016) marching are embedded in the inversion
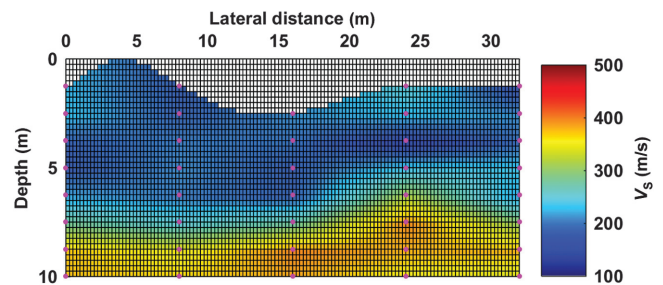

Figure 1. An example of the two-grid scheme. The black grid indicates the fine modeling grid that guarantees the reliable modeling of Rayleigh waves. The magenta dots denote the coarse inversion grid. A bilinear interpolation converts the coarse-grid model into its correspondent fine-grid model. The background colors refer to the V_S model of the third example. In the figure, only the portion of the model that is of interest is exhibited; that is, the absorbing boundaries to be added in modeling are not shown here.
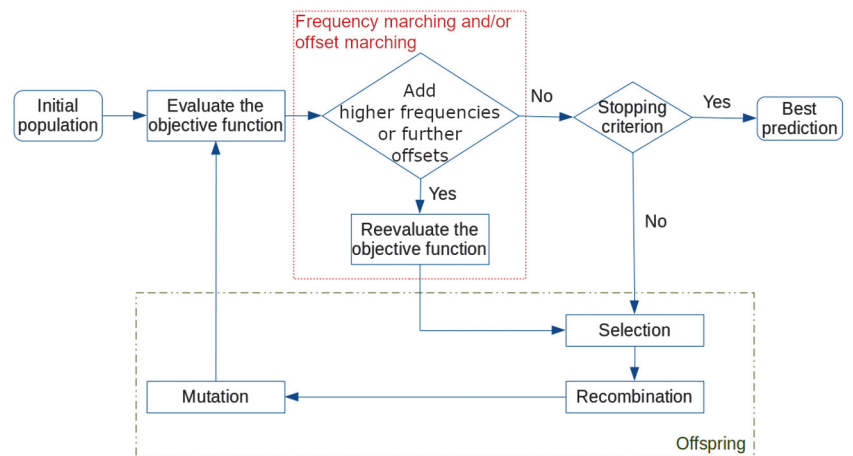

Figure 2. The workflow of our genetic-algorithm FWI. Frequency and offset marching have been embedded in the workflow to further avoid cycle skipping.

workflow for the further avoidance of cycle skipping. Below, we list the main controlling parameters (Pohlheim, 2006) of the genetic algorithm we use. The correct setting of the parameters is empirical, and it is strongly influenced by the number of unknowns. In the list, we also indicate the settings adopted in all the inversion tests shown in this and in the accompanying paper. Our choices were based on several tests and previous works (Sajeva et al., 2014, 2016; Aleardi and Mazzotti, 2017; Mazzotti et al., 2017; Xing and Mazzotti, 2017a, 2017b).

- Number of individuals: It is the number of subsurface models that form the population. The higher this number, the more thorough the exploration of the model space. With approximately 200 unknowns (V_S, V_P, and rho at the nodes of the coarse model grid), the number of individuals that we set was 2000.
- Number of subpopulations: It is the number of groups that population (models) is divided into so as to simulate biological migration. We adopted five subpopulations. In our case, migration means moving a percentage of models from one to other subpopulations. Recombination in Figure 2 occurs for interpopulation.
- Maximum generation: It is the number of generations evolved before the inversion outcome is obtained, and we set it to 200.

- Selection rate: It is the percentage of the models (individuals) that are chosen to be paired for generating models of the next generation (offspring). We set this parameter to 0.8.
- Selection pressure: It is the ratio between the probability that more promising models, namely, the models associated with relatively smaller data misfits, are selected and the probability that any model is chosen. In our tests, it changed linearly from 1 to 2 along generations.
- Mutation rate: It is the probability that a model is mutated, and we set it to 0.1.
- Reinsertion rate: It is the percentage of the offspring that replace the less promising individuals (models) in the original population to form the new population. Here, 0.6 was the setting that we chose.
- The first generation of migration: It is the number of the generation at which migration occurs at the first time. We set it to 30.
- Migration interval: It is the number of the generations between two successive migrations. In our tests, it was set to 20.
- Migration rate: It is the percentage of models that are allowed to migrate. Here, 0.2 was the value that we adopted.

The values of the controlling parameters were chosen considering the balance between the inversion results and the computational time. For instance, if we increase the number of individuals, the exploration of the model space will be more thorough at the expense of higher computing costs. Conversely, if we decrease the selection rate, we may attain a faster convergence and save computing time, but we increase the risk of reaching a premature convergence and being stuck in a local minimum.

According to our experience, with the proposed genetic-algorithm FWI, depending on the selected coarse-grid spacing, we are able to derive final model predictions or provide adequate initial models for Rayleigh-wave FWI with local optimization techniques for further refinements (Xing et al., 2018).

In what follows, we show the application of our FWI to three synthetic examples. Instead, for checking the application of the method to actual Rayleigh-wave data, along with the additional considerations demanded for the practical use of the method, readers can refer to our companion paper.

SYNTHETIC EXAMPLES

We aim at demonstrating that the proposed method is able to fairly predict near-surface V_S models from Rayleigh waves, even in the case in which no a priori information is given. In this section, we first present the reference models. Then, we show the inversion specifications. Finally, we illustrate the predicted V_S models and seismograms together with the evolution of data and model misfits.

Reference models

The reference models that we used in the three tests are displayed in Figure 3. Only V_S models are shown due to the well-known fact that Rayleigh waves are mostly sensitive to V_S; consequently, V_S is the most important parameter that can be retrieved by Rayleigh-wave inversion.

The first model (Figure 3a) is a 1D model with strong velocity contrasts and velocity inversions in the second and fourth layers. This kind of layering is generally considered as difficult to invert
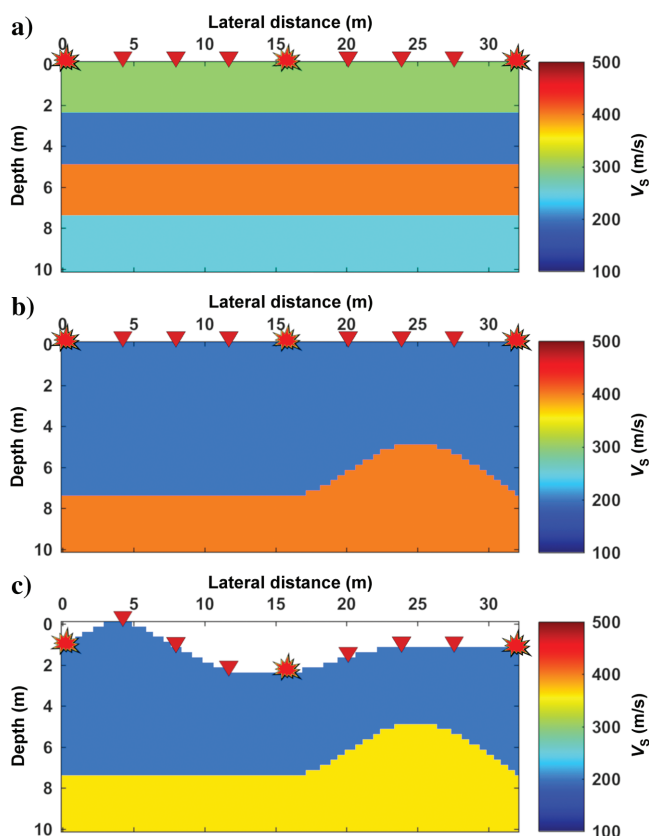
Figure 3. Reference near-surface V_S models: (a) 1D model with sharp velocity contrasts and inversions, (b) 2D model with strong lateral velocity variations, and (c) 2D model with an irregular topographic surface and lateral velocity variations. The blasts and reverse triangles on the topographic surfaces illustrate the locations of the sources and receivers used in the inversions.

for by existing Rayleigh-wave inversion techniques (among others, Cercato, 2009). This seems to be confirmed by some test that we carried out using dispersion curve inversion with no or with scarce a priori information. One of the reasons is that the velocity inversions prevent the generation of head waves and consequently make it difficult to build suitable initial models by approaches such as first-break tomography.

The second model (Figure 3b) is a 2D model with strong lateral variations: At the depth of, say, 6 m, at the lateral coordinate 20 m, the lateral velocity contrast amounts to 200 m/s. This "anticlinal" structure violates the 1D assumption that is common in dispersion curve inversion, and makes it difficult to be inverted for via that method.

The third model (Figure 3c), besides the lateral velocity variation at depth, shows an irregular topographic surface that further distantiates the model from 1D geometry. Moreover, the different elevations of sources and receivers will result in the distortion of the observed data and will likely produce diffractions. Intuitively, the increased complexity will make the data more vulnerable to the local-minimum problem (Bunks et al., 1995; Cercato, 2011).

We carried out the synthetic tests committing an inversion "crime," that is, the observed data were computed by using the same forward-modeling engine used in our inversion code and source wavelets were known.

The acquisition geometry was the same for all the three tests. Following the lead of engineering studies in which, upon applying techniques such as MASW, only a few shots are generally used, three shot gathers, of which one is split-spread and two are off-ends with evenly spaced receivers, were the input data for the inversion. The blasts and reverse triangles in Figure 3 symbolically illustrate the positions of the sources and receivers that were placed on the topographic surface. The source-to-receiver offsets ranged from 4.25 to 32 m.

To test the impact of background noise, weak random noise was added to the simulated observed data in the first example.

Inversion specifications

The same inversion specifications were set for the three tests.

Because we simulated the case in which no a priori information was available, the genetic-algorithm search ranges were constant with depth and with lateral position. Search ranges limit the model space that can be explored by the genetic algorithm and, thus, it is recommendable to use a relatively large search range when no a priori information can suggest a restriction. Therefore, we set the search ranges for V_S, V_P, and rho as 100–500 m/s, 500–900 m/s, and 1350–1750 kg/m³, respectively.

Data misfits χ were calculated by the $L1$ norm, which is less vulnerable (Kwak, 2008) to outliers compared with the $L2$ norm, between the observed and predicted data

$$\chi = \frac{\sum_{n_x=1}^{N_x}\sum_{n_t=1}^{N_t}|\mathbf{D}_{n_t,n_x}-\mathbf{P}_{n_t,n_x}|}{\sum_{n_x=1}^{N_x}\sum_{n_t=1}^{N_t}|\mathbf{D}_{n_t,n_x}|}, \qquad (1)$$

where $\mathbf{D}$ and $\mathbf{P}$ are the observed and predicted data, respectively, n_t and n_x are the time and trace sampling number, respectively, N_t and N_x are the total number of time samples and traces, respectively. In the tests, the total misfit was a summation over all three shots.

We made use of the offset-marching strategy, and we started the inversion with the near-offset data until the offset of 9 m. Then, we gradually included in the inversion the data at larger offsets until 12,

15, 20, 26, and 32 m. Given the minimum velocity defined in the search ranges and the inverted maximum frequency that were the same for all three examples, the modeling grid was composed of 5289 nodes to guarantee an appropriate grid spacing. The coarse grids for the inversion were defined simulating the case in which no a priori information could suggest a particular geometry of the grids; thus, they were set as regularly spaced. In the two flat-topography examples, we defined the same coarse inversion grid of 45 nodes, which brought the total to 135 unknowns to be inverted for. Instead, in the test with the 2D irregular topography model, due to the plain fact that the part above the surface was not to be inverted for, the coarse grid was built with 39 nodes, which led to 117 unknowns.

The other parameters related to the genetic algorithm, such as the number of individuals and the maximum generation, were set as the empirical values listed in the "Method" section.

The tests were carried out on a distributed computing system made of computers with two cores. Each core had 20 threads. Using 1 thread, each forward simulation took approximately 2 s. Using 545 threads, each inversion took approximately 1 h.

Inversion outcomes

In this part, we show that fair inversion outcomes can be obtained by applying the two-grid genetic-algorithm FWI method, even though in the tests (1) complex near-surface models have been
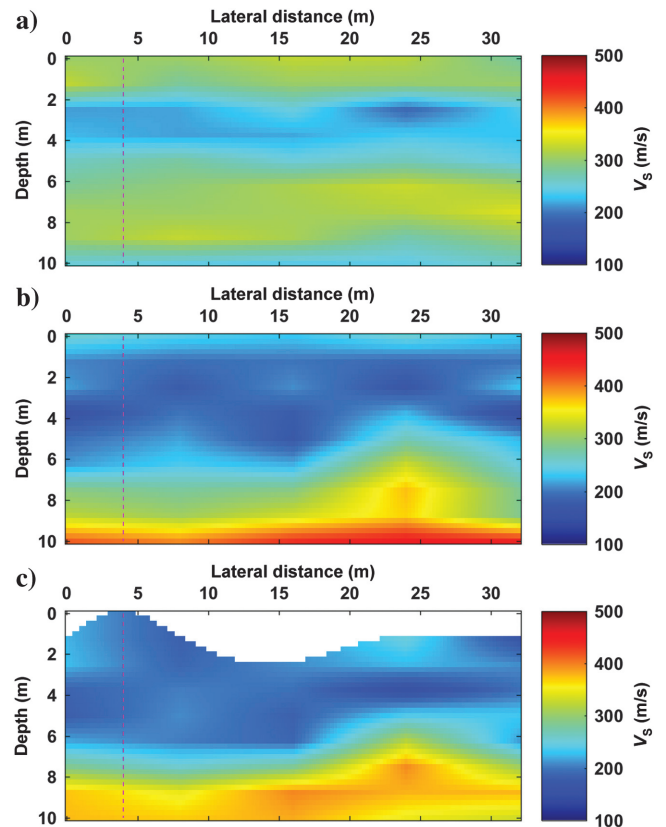

Figure 4. The predicted V_S models corresponding to the best seismograms (minimum data misfits): (a) 1D model, (b) 2D model, and (c) 2D irregular topography model. The dashed magenta lines indicate the positions of the 1D V_S profiles that will be presented in Figure 5. Note that the long-wavelength structures of the reference models (Figure 3) have been fairly reproduced.

considered, (2) coarse inversion grids have been used, and (3) no a priori information has been exploited. The predicted V_S models are displayed in Figure 4, whereas the 1D V_S vertical profiles at the lateral distance of 4 m are presented in Figure 5 to allow for a more detailed comparison. The best predicted seismograms, along with the observed data, are shown in Figure 6. An example of the evolution of the data and model misfits is given in Figure 7.

Concerning the first test on the 1D model (Figure 3a), we can see that the main subsurface features have been recovered (although with moderately different velocities) by the inversion (Figure 4a). Despite the fact that the search ranges had been set constant with depth, the two velocity inversions located at the second and fourth (bottom) layer have been detected and the prediction of their depth positions is acceptable. Obviously, sharp interfaces cannot be reproduced owing to the adopted parameterization of the subsurface. It appears that the random noise added to the observed data has not influenced the results.

The model prediction for the 2D model test is shown in Figure 4b and should be compared with the reference model in Figure 3b. The long-wavelength structures hav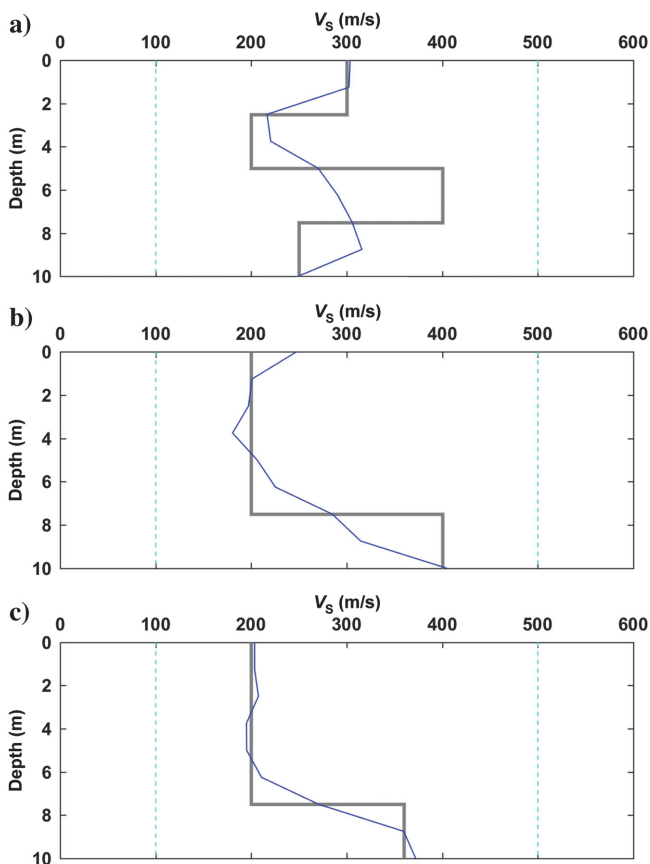e been predicted, particularly the anti-clinal form on the right and the general increase in velocities with depth. The predicted model does not fully coincide with the reference model, but this is expected due to the coarseness of the inversion grid and, likely, illumination problems at the edges of the model.

The inversion result of the third example, the one with an irregular topography model, is shown in Figure 4c. Again, the main structure of the reference model (Figure 3c) has been recovered even in the shallowest part, and the transition from low velocities to deeper and laterally varying higher velocities is quite distinguishable.

A more detailed assessment of the results can be made on the 1D V_S profiles shown in Figure 5. The dashed cyan lines indicate the search ranges set in the inversion. Note that the search ranges are wide and constant; that is, they do not follow the V_S trends of the reference models. Notwithstanding that, all of the predicted velocity trends seem to be a low-frequency reproduction of the actual trends. Although the velocity values of the reference and predicted models do not coincide, the inversion seems to be able to detect the velocity changes (particularly the velocity inversions in the first example).

Coming to the comparison between observed and predicted data, Figure 6 shows the three sets of shot gathers, separated by the dashed cyan lines, with observed traces represented in black and predicted traces represented in red. All of the seismograms have

Figure 5. The V_S profiles picked from the reference (Figure 3) and predicted (Figure 4) models at the lateral distance of 4 m. From (a to c), the profiles sequentially correspond with the 1D model, the 2D model, and the 2D irregular topography model. The profiles picked from the reference models are drawn in gray, whereas those picked from the predicted models are presented in blue. The dashed cyan lines show the V_S search ranges in our genetic-algorithm FWI. Note that the search ranges are wide and constant; that is, they do not follow the trends of the reference V_S profiles, and that the predicted V_S profiles fairly match the reference model profiles.
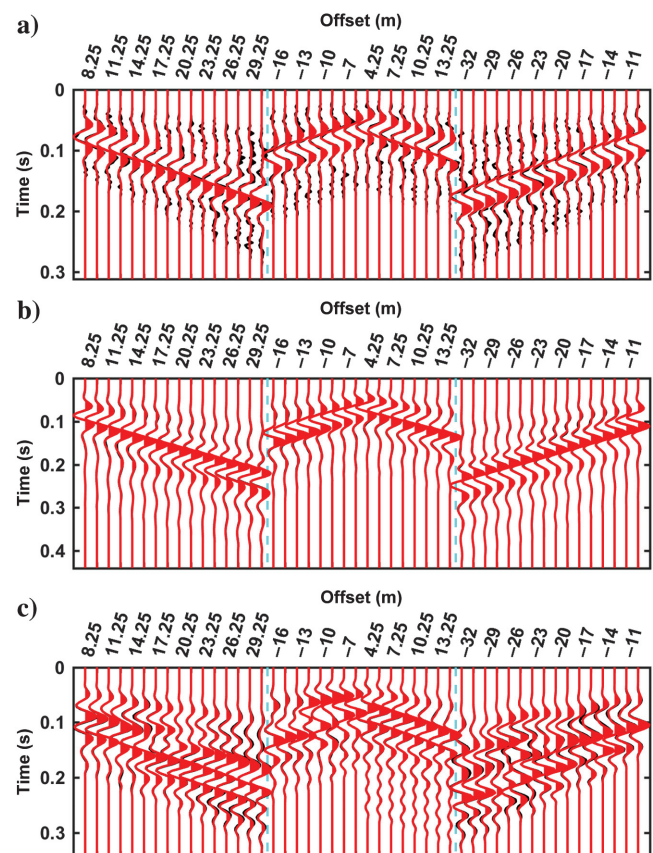

Figure 6. The observed (the black traces) and predicted (the red traces) best seismograms related to the (a) 1D model, (b) 2D model, and (c) 2D irregular topography model. One in every six traces is plotted. The observed seismograms in (a) have been contaminated with weak random noise. The dashed cyan lines delimit the left, middle, and right shot gathers. The seismograms have been normalized trace by trace. The matching between the observed and predicted seismograms is quite good at all offsets.

been normalized trace by trace. Figure 6a illustrates the case of the first example. Although the observed data have been contaminated by weak random noise (S/N equal to 30.7 dB), the predicted data fairly match the observed data along the whole range of offsets. The same conclusion can also be drawn for the other two examples shown in Figure 6b and 6c. Note that in Figure 6c, the complex distortions of the observed wavetrain caused by the irregular topographic surface have been well reproduced in the simulated data.

As a further check, we present in Figure 7 the evolution along the genetic-algorithm generations of the data misfit and of the V_S, V_P, and rho model misfits pertaining to the third example. The evolution of the misfits in the other two examples is not exhibited because they show the same character. The black curves indicate the minimum data misfit at each generation in Figure 7a, whereas in Figure 7b–7d, they indicate the misfit of the model associated with the seismogram giving the minimum data misfit. The red curves in Figure 7 are, sequentially, the mean misfits computed over the entire population of simulated seismograms (Figure 7a) and their respective models of V_S, V_P, and rho (Figure 7b–7d). The dashed cyan lines delineate the offset-marching frame. The blue annotations indicate the offset ranges of the inverted data at each offset-marching phase.

In Figure 7a, we can observe that within each offset-marching phase, the minimum as well as the mean data misfits tend to rapidly decrease, which is encouraging. At generations when an offset-marching transition occurs, the absolute data misfit will likely increase because the data can be very different before and after the addition of an offset range. However, from the plot in Figure 7a, we can assess the evolution of the data misfit only within each offset range, but we cannot make any comparison among the data misfit values within different offset ranges due to the fact that the observed data that are the normalization factor in the objective function of equation 1, change with offset range. At the last (200th) generation of the inversion, the mean data misfit is very close to the minimum data misfit. Combined with the fairly predicted data in Figure 6c, this closeness indicates that the space for a further reduction of the data misfit is limited.

The evolution of the V_S-model misfit (Figure 7b), although generally decreasing, shows significant oscillations that are a consequence of the strong nonlinearity of the Rayleigh-wave inversion. We use the term "best" model misfits to indicate the misfits of the models that give rise to the minimum data misfits. Encouragingly, despite the ambiguity, in Figure 7b, the best V_S-model misfits and the mean V_S-model misfits show a general decrease with generations. At the last (200th) generation, the mean V_S-model misfit is slightly larger than the best V_S-model misfit, but the two misfits are fairly near, indicating that the best and the mean model are likely very similar. This is indeed demonstrated in Figure 8 in which the mean models for all of the previous examples are shown. The reason for the similarity between the best models and the mean models is the continuous loss of the variety in the population along generations as a consequence of selection and recombination.
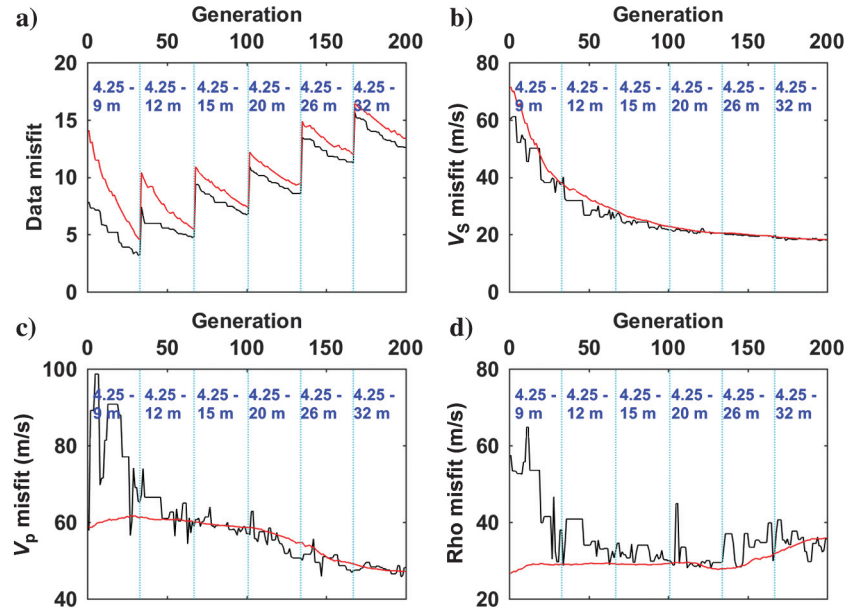

Figure 7. The evolution of the data and model misfits for the third example: (a) data misfits, (b) V_S-model misfits, (c) V_P-model misfits, and (d) rho-model misfits. The black and red curves in (a) indicate the minimum data misfits and the mean data misfits, respectively. In (b-d), the black curve shows the misfits of the V_S, V_P, and rho models associated with the seismograms giving the minimum data misfits (best seismograms), whereas the red curve represents the misfits of the mean models calculated over the entire population. The blue annotations delimited by the dashed cyan lines indicate the offset ranges of the data inverted at each phase of the offset-marching frame.
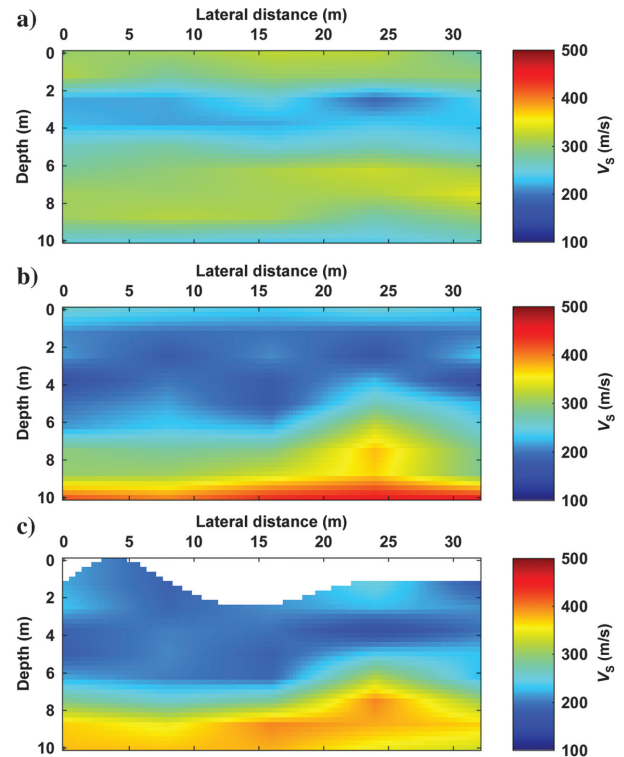

Figure 8. The mean of the V_S models at the last generations of our genetic-algorithm Rayleigh-wave FWI: (a) 1D model, (b) 2D model, and (c) 2D irregular topography model. These mean models are rather similar to the best models presented in Figure 4.

Concerning the estimation of V_P and rho models, Figure 7c and 7d shows their respective evolution. It is immediately evident in the effect of the minor sensitivity of the data to V_P and rho. In fact, although we observe a general improvement in the best V_P-model misfit and the best rho-model misfit (the black curves in Figure 7c and 7d) along generations, the rate of the improvement is significantly lower than that for V_S models. Also, V_P and rho models associated with the best seismograms may show misfits that are much greater than the mean model misfits at the same generation. This is again a consequence of the lower control that V_P and rho exercise on the observed Rayleigh waves: The average of a certain number of — somewhat erratic — V_P and rho models may be more similar to the reference model than the model associated with the best seismogram. This is why, according to our inversion tests, unless a priori information is included, we can rarely obtain V_P or rho models with correct structures. In fact, the V_P and rho models retrieved in the present tests are far from reproducing the true models.

CONCLUSION

We have proposed a two-grid Rayleigh-wave FWI via a genetic-algorithm optimization. A two-grid scheme, which limits the number of unknowns so as to decrease significantly the computational time, has been adopted for the practical applicability of the method. The proposed inversion fairly predicts the long-wavelength components of near-surface V_S models, confirming the ability of genetic algorithms to converge even when no a priori information is available.

Among the many tests that had been performed, we have shown here the results pertaining to three synthetic models that are generally considered as being fairly complex to be inverted for by means of dispersion curve inversion or by FWI with local optimization techniques, without a priori information. The three models individually contain (1) strong velocity contrasts and velocity inversions, (2) strong lateral velocity variations, and (3) an irregular topographic surface.

Although fairly coarse inversion grids had been used, so that the reconstructed models can only be low-resolution models, in the first test with the 1D model, the two velocity decreases have been correctly detected, whereas in the second and third examples, the lateral velocity variations and the effects of the irregular topography have been fairly reproduced.

Because of the fair reproduction of the macrostructures of the reference models and the adequate match of predicted and observed seismograms, the models obtained by our stochastic method are supposed to be suitable inputs to Rayleigh-wave FWI through local optimization techniques for further refinement, if needed. Preliminary tests carried out using IFOS2D, a gradient-based FWI code developed by the Toolbox for Applied Seismic Tomography project (Bohlen, 2002; Köhn et al., 2012; Groos et al., 2014), seem to confirm the expectations.

Because all of the models explored by the genetic algorithm can be collected, we can also think of expressing the results not just as the best model or the mean model, but as frequency histograms that could be further elaborated to retrieve probability distributions. This would allow for the parallel estimation of uncertainties associated with the most likely model. However, this approach remains among the works to be done for Rayleigh-wave FWI. Instead, a work that has been done is the application of the proposed method to two field data cases. The results are presented in a companion paper, along with a discussion on practical issues that need to be addressed.

ACKNOWLEDGMENTS

DATA AND MATERIALS AVAILABILITY

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

APPENDIX A

3D TO 2D CORRECTION

In the real world, wavefronts are spherical, whereas in 2D FDM, the source is a line source that gives rise to cylindrical wave propagation. To make field data (or modeled 3D data) and our 2D modeling outcomes comparable, we correct the former via a 3D to 2D correction technique proposed by Forbriger et al. (2014).

Following Forbriger et al. (2014) and Schäfer et al. (2014), we apply the multilayer surface-wave transformation to the 3D data, with the exception of the near-offset traces in which the single-layer transformation is used instead.
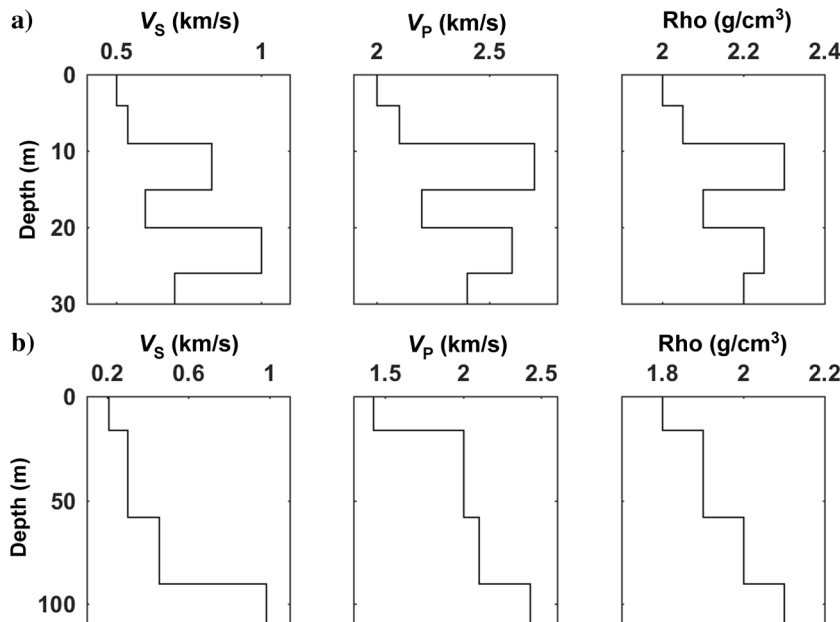

Figure A-1. Models used for checking the effectivity of the adopted 3D-to-2D correction (Forbriger et al., 2014) method. (a) A synthetic model that contains sharp velocity contrasts and velocity inversions. (b) A realistic model derived from 3C down-hole recordings and density log.

The algorithm for multilayer surface-wave transformation is as follows:

1) Convolve the data with $\sqrt{t^{-1}}$ where t is the time.
2) Multiply the data with $r\sqrt{2t^{-1}}$ where r is the source-receiver distance.
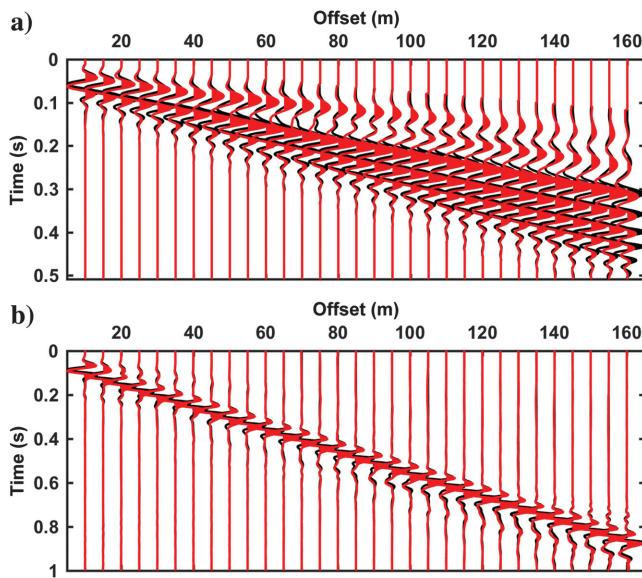

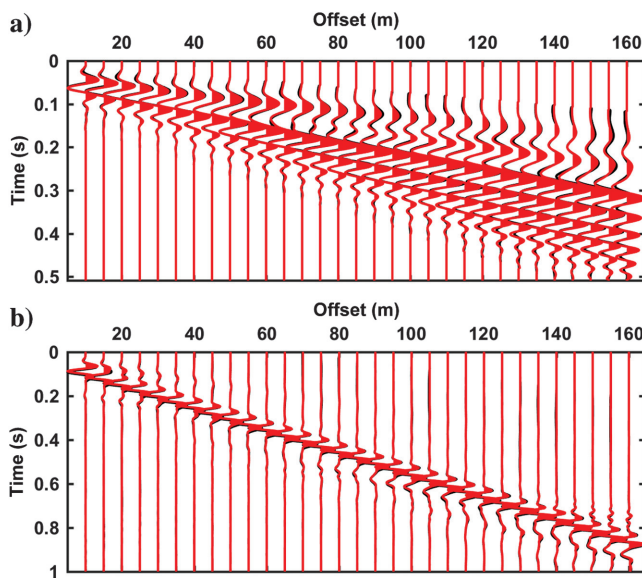Figure A-2. Reflectivity modeling results of (a) the model in Figure A-1a and (b) the model in Figure A-1b. The black and red seismograms are the reflectivity modeling outcomes before and after the application of the 3D-to-2D correction, respectively. The seismograms have been normalized trace by trace. The phase differences are evident at the whole offset range.

Figure A-3. Modeling outcomes of (a) the model in Figure A-1a and (b) the model in Figure A-1b. The black seismograms are the results of reflectivity modeling with the 3D-to-2D correction applied, whereas the red seismograms are outcomes of the used 2D FDM code. The seismograms have been normalized trace by trace. In (a and b), the black and red seismograms satisfactorily match, even at the farthest offset.

The algorithm for the single-layer transformation is as follows:

1) Convolve the data with $\sqrt{t^{-1}}$.
2) Multiply the data with $\sqrt{2rV_{ph}}$, where V_{ph} is the single-phase velocity.

We checked the effectiveness of the adopted correction method on the basis of two near-surface models shown in Figure A-1. The first model (Figure A-1a) is a quite complex near-surface model that contains sharp velocity contrasts and strong velocity inversions. The second model (Figure A-1b) is a realistic model derived from an actual borehole.

The reflectivity modeling results of the two models are shown in Figure A-2, with the black traces indicating the seismograms simulated via a 3D wave propagation and the red traces representing the 3D to 2D corrected seismograms. For the seismograms of both models, the phase correction is particularly evident although a significant amplitude correction has also been performed but its effect is less visible due to the trace-by-trace normalization.

In Figure A-3, the reflectivity modeling results with the 3D-to-2D correction applied are displayed as the black seismograms, whereas the 2D elastic FDM outcomes are presented as the red seismograms. The results for both models show very satisfactory matching between FDM traces and 3D-to-2D corrected reflectivity traces at all offsets.

In the perspective of Rayleigh-wave FWI, the application of the 3D-to-2D correction, whose effectiveness is confirmed by the quite-good match between the 2D FDM seismograms and the 3D-to-2D corrected reflectivity seismograms, will assure more correct model predictions.

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.

Alford, R. M., K. R. Kelly, and D. M. Boore, 1974, Accuracy of finite-difference modeling of the acoustic wave equation: Geophysics, **39**, 834–842, doi: 10.1190/1.1440470.

Bellman, R. E., 1957, Dynamic programming: Princeton University Press.

Bodin, T., M. Sambridge, H. Tkalčić, P. Arroucau, K. Gallagher, and N. Rawlinson, 2012, Transdimensional inversion of receiver functions and surface wave dispersion: Journal of Geophysical Research, **117**, 1–24, doi: 10.1029/2011JB008560.

Bohlen, T., 2002, Parallel 3-D viscoelastic finite difference seismic modeling: Computers and Geosciences, **28**, 887–899, doi: 10.1016/S0098-3004(02)00006-7.

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.

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.

Bunks, C., F. M. Saleck, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: Geophysics, **60**, 1457–1473, doi: 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.

Cercato, M., 2011, Global surface wave inversion with model constraints: Geophysical Prospecting, **59**, 210–226, doi: 10.1111/j.1365-2478.2010.00922.x.

Collino, F., and C. Tsogka, 2001, Application of the perfectly matched absorbing layer model to the linear elastodynamic problem in anisotropic heterogeneous media: Geophysics, **66**, 294–307, doi: 10.1190/1.1444908.

Festa, G., and J. Vilotte, 2005, The Newmark scheme as velocity-stress time staggered: An efficient PML implementation for spectral element simulations of electrodynamics: Geophysical Journal International, **161**, 789–812, doi: 10.1111/j.1365-246X.2005.02601.x.

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.

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.

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.

Fuchs, K., and G. Muller, 1971, Computation of synthetic seismograms with the reflectivity method and comparison with observations: Geophysical Journal of the Royal Astronomical Society, **23**, 417–433, doi: 10.1111/j.1365-246X.1971.tb01834.x.

Gedney, S. D., 2011, Introduction to the finite-difference time-domain (FDTD) method for electromagnetics: Morgan and Claypool Publishers.

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.

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.

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.

Komatitsch, D., and R. Martin, 2007, An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation: Geophysics, **72**, no. 5, SM155–SM167, doi: 10.1190/1.2757586.

Komatitsch, D., R. Martin, J. Tromp, M. A. Taylor, and B. A. Wingate, 2001, Wave propagation in 2-D elastic media using a spectral element method with triangles and quadrangles: Journal of Computational Acoustics, **9**, 703–718, doi: 10.1142/S0218396X01000796.

Kwak, N., 2008, Principal component analysis based on L1-norm maximization: IEEE Transactions on Pattern Analysis and Machine Intelligence, **30**, 1672–1680, doi: 10.1109/TPAMI.2008.114.

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.

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.

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.

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.

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.

Mazzotti, A., N. Bienati, E. Stucchi, A. Tognarelli, M. Aleardi, and A. Sajeva, 2017, Two-grid genetic algorithm full-waveform inversion: The Leading Edge, **35**, 1068–1075, doi: 10.1190/tle35121068.1.

Nagai, K., A. O'Neill, Y. Sanada, and Y. Ashida, 2005, Genetic algorithm inversion of Rayleigh wave dispersion from CMPCC gathers over a shallow fault model: JEEG, **10**, 275–286, doi: 10.2113/JEEG10.3.275.

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.

Pohlheim, H., 2006, Evolutionary algorithms: Overview, methods and operators, http://www.geatbx.com/download.html, accessed 24 January 2018.

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.

Robertsson, J. O. A., 1996, A numerical free-surface condition for elastic/viscoelastic finite-difference modeling in the presence of topography: Geophysics, **61**, 1921–1934, doi: 10.1190/1.1444107.

Roden, A., and S. D. Gedney, 2000, Convolutional PML (CPML): An efficient FDTD implementation of the CFS-PML for arbitrary media:

Microwave and Optical Technology Letters, **27**, 334–339, doi: 10.1002/1098-2760(20001205)27:5<334::AID-MOP14>3.0.CO;2-A.

Sajeva, A., M. Aleardi, B. Galuzzi, E. Stucchi, E. Spadavecchia, and A. Mazzotti, 2017, Comparing the performances of four stochastic optimisation methods using analytic objective functions, 1D elastic full-waveform inversion, and residual static computation: Geophysical Prospecting, **65**, 322–346, doi: 10.1111/1365-2478.12532.

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.

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.

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, Extended Abstracts, doi: 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.

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.

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.

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.

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.

Thorbecke, J. W., and D. Draganov, 2011, Finite-difference modeling experiments for seismic interferometry: Geophysics, **76**, no. 6, H1–H18, doi: 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.

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.

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, Extended Abstracts, doi: 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.

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, Extended Abstracts, doi: 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.

Xing, Z., and A. Mazzotti, 2019, Two-grid full-waveform Rayleigh-wave inversion via a genetic algorithm — Part 2: Application to two actual data sets: Geophysics, **84**, this issue, doi: 10.1990/geo2018-0800.1.

Xing, Z., F. Rappisi, and A. Mazzotti, 2018, A new workflow for surface wave FWI combining genetic algorithm and gradient-based optimization algorithms: Near Surface Geoscience Technical Program, Extended Abstracts, doi: 10.3997/2214-4609.201802567.