TOWARDS AN AUTOMATIC FULL WAVE INVERSION: SYNTHETIC STUDY CASES

Jean Kormann [1,2]: jean.kormann@unileoben.ac.at, Juan Esteban Rodríguez [1]: juan.rodriguez@bsc.es, Natalia Gutierrez [1]: natalia.gutierrez@bsc.es, Miguel Ferrer [1]: miguel.ferrer@bsc.es, Otilio Rojas [1,3]: otilio.rojas@bsc.es, Josep de la Puente [1]: josep.delapuente@bsc.es, Mauricio Hanzich [1]: mauricio.hanzich@bsc.es and José María Cela [1]: josem.cela@bsc.es

[1]: Barcelona Supercomputing Center (BSC), Jordi Girona 29, 08034 Barcelona, Spain.

[2]: Chair of Applied Geophysics, MontanUniversitaet Leoben, Peter-Tunner-Straße 25, 8700 Leoben, Austria.

[3]: School of Computer Science, Faculty of Science, Universidad Central de Venezuela, Caracas, Venezuela.

Original paper date of submission: August 12, 2016

Revised paper date of submission: ----

ABSTRACT

Full Waveform Inversion (FWI) in seismic scenarios continues to be a complex procedure for subsurface imaging that might require extensive human interaction, in terms of model setup, constraints and data preconditioning. The underlying reason is the strong non-linearity of the problem that forces the addition of a priori knowledge (or bias) in order to obtain geologically sound results. In particular, when the use of long offset receiver is not possible or may not favor the reconstruction of the fine structure of the model, one needs to rely on reflection data. As a consequence, the inversion process is more prone to get stuck into local minima. It is then possible to take advantage of the cross-correlation error functional, less subject to starting models error, in order to output a suitable background model for inversion of reflection data. By combining these functionals, high-frequency data content with poor initial models can be successfully inverted. If we can find simple parameterizations for such functionals we can can reduce the amount of uncertainty and manual work related to tuning FWI. Thus FWI might become a semi-automatized imaging tool.

INTRODUCTION

Full Waveform Inversion (FWI) represents a seismic imaging method able to improve Earth structural models up to spatial resolutions beyond the limits of standard Travel Time Tomography (TTT), and more adequate for seismic imaging. TTT only inverts the time residuals of (mostly) P-wave phases picked on the recorded field traces, requiring human interaction.On the other hand, FWI processes the whole waveforms achieving a finer resolution. Nevertheless, given our surface to surface acquisition limitations, noise effects and initial models with poor

low frequency content, convergence to the true model cannot be guaranteed. Among the strongest concerns when using FWI is the matching of synthetic and data phases when they are apart more than half a cycle in time, an effect known as cycle skipping (Luo and Schuster, 1991; Warner and Guasch, 2014; Metivier et al., 2016). Some functional have been developed over the last decades to cope with this issue: e.g the cross-correlation (CC) traveltime functional (Luo and Schuster, 1991), the adaptive FWI from Warner and Guasch (2014), or the optimal transport distance (Métivier et al., 2016). Although less sophisticated, the CC is able to provide good background models as reported by, e.g., Jimenez-Tejero et al. (2015), but lack in resolution. On the other hand, it has been showed that short-offset allow for better reconstruction of the fine-structure of the model (Kormann et al., in press) by means of the Dynamic Offset method. Another benefit of such approach is that it could be easily parametrized with only one variable. Unlikely to the CC, this technique relies on a good starting model in order to avoid cycle skipping effects and misplacement of the reflectors. As theses methods aim at recovering two different but complementary components of the same model, it is natural to set up a workflow that would embedded both approaches.

Thus in this article we have two main goals: 1) developing two procedures for FWI, one aiming at correcting the background model based on the CC functional, and the other one focusing on the reflection data by means of the L2 norm and receiver offset limitation. 2) Automatizing as much as possible these two procedures by taking advantage that both procedures share the same modeling engine, and could therefore be embedded in the same global FWI package. In the following we define two procedures M1 and M2 that we use in our subsequent inversion strategies. The first method could be parameterized with only one variable

(the maximum offset for receiver), and we also propose a criterion that allows for an automatic selection of it as a function of the frequency. The second approach relies on automatic picking of the first arrivals (detailed below), including some constraints over the maximum CC time. As a result, it produces a background model which is used as an input for M1.

We therefore define a minimal set of inversion parameters that can be used to properly select and window first arrivals in the data and focus on reflected waveforms strongly reducing human interaction. We use the FWI package from the Barcelona Subsurface Imaging Tools (BSIT[1]) operating on the Marenostrum supercomputer[2] to exemplify our automated procedures. In the following, we will describe and exemplify the proposed methodology through two synthetic dataset.

FORMULATION AND PROCEDURES

We use a standard FWI formulation as described in recent literature (Fichtner, 2011). Forward modeling of 2D and 3D seismic wavefields is carried out by a high-order finite-difference scheme in time domain, with sponge zones or perfectly matched layer as absorbing boundaries. The application of the adjoint method in time domain results in the correlation of a forward and backward wavefield (Tarantola, 1984). The efficient retrieval of forward wavefields is accomplished by means of lossy compressed wavefield storage and temporal decimation (Boehm et al., 2016), while, adjoint sources for the backward wavefield are obtained by the derivation the misfit functional respect to time, and are back propagated to obtain

---

[1] www.bsc.es/bsit
[2] www.bsc.es/marenostrum-support-services

the adjoint wavefield. Finally, the correlation between the stored forward wavefield and the backward propagated wavefield allows the computation of the model gradients. In our FWI, the optimization problem is solved by means of the Polak-Ribière non-linear conjugate gradient method (see e.g. Nocedal and Wright, 2006).

FWI Procedures

At the core of FWI lies the adjoint method and the minimization of an error functional E that measures the difference between the seismic response of two structural models, the current **m** and the real unknown model **m***. The evaluation of E quantifies the differences between the modeled seismic response u of model **m** and the real records u^* recorded in a survey obtained from **m***. The choice of E, together with data conditioning and selection, define what we call an FWI *procedure*, i.e. a recipe to apply FWI to a certain dataset may include the choice of E, regularization, preconditioning, data selection. In the following we introduce two complementary FWI procedures that are the backbone of our nearly-automatic FWI methodology.

M1 Procedure: Normalized L2 with Dynamic Offset

Typical FWI implementations rely on the L2 norm between data and synthetic traces (Tarantola, 1984). Such norm is very sensitive to small perturbations in the measurement u^* but also prone to suffer from cycle skipping whenever the waveforms $u(x_r,x_s,t,\mathbf{m})$ and $u^*(x_r,x_s,t)$ are too far apart. In addition the classical L2 norm is strongly affected by amplitude changes which might not always describe properly model changes, for example when inverting velocity.

In order to mitigate the problem of amplitude dependence, a normalization can be used that gives equal weight to a data trace and its equivalent synthetic trace. Thus, only relative amplitude differences matter. Combining a box-like data selection and difference evaluation results in the following misfit function

$$E_{M1}(\mathbf{m}) = \sum_s \sum_{|x_r - x_s| < x_D} \sum_{t < T} \left[\frac{u(x_r, x_s, t)}{max_t(|u(x_r, x_s, t)|)} - \frac{u^*(x_r, x_s, t)}{max_t(|u^*(x_r, x_s, t)|)} \right]^2 ,$$

where x_s and x_r are the source and receiver locations, respectively, t represents time samples, T is the maximum time used for the inversion and x_D is the maximum data offset coming what we call Dynamic Offset (DO) (Kormann et al., 2016), defined as

$$x_D = \frac{D}{f_0} ,$$

being f_0 the frequency to be inverted (e.g the frequency of the low-pass filter applied) and D a proportionality constant, typically related to a characteristic velocity of the model. Although the choice of D is rather empirical and based on numerical experimentation (Kormann et al., 2016), we find $D = 2 \cdot max(V_p)$ a safe choice in our applications, being V_p the compressional velocity. Hence x_D can be interpreted as twice the maximum wavelength of our model. Consequently, this procedure results in data dominated by reflections instead of refractions, as we increase frequency. We notice that when $D \to \infty$ we recover a classical, yet normalized, full-offset L2 norm. Finally, using DO also results in some computational savings, as the domain modelled for each shot is smaller than that for the full offset case (i.e. $D \to \infty$).

Fundamental for our interests, the M1 procedure is fully parameterized with parameters *D* and *T* (the same for all shots and receivers) which act as a data selection criterion, whereas the normalization acts as data conditioning. Both effects are thus included in our misfit function E_{M1} automatically, which results in a very automatized FWI procedure.

M2 Procedure: L1 cross-correlation, full offset

By maximizing the cross-correlated observed-to-synthetic differences (in some norm), FWI can attain a better reconstruction of low-frequency features (Jimenez-Tejero et al., 2015). The dependence of first-arrival traveltime shifts with respect to velocity perturbations is more linear than other waveform measures, which explains the higher robustness of CC-based metrics compared to original L2 waveform misfits when inverting velocity models (Luo and Schuster, 1991; Jimenez-Tejero et al., 2015). In this way, cross-correlation functional relies on the assumption that data and synthetics are close enough.

CC time is usually defined in terms of the L2 norm of the travel-time shifts; however, we aim at reducing the method's dependence on noise by using the L1 norm here (Brossier et al., 2010). This implies that the adjoint source derived from the error functional is only dependant on the sign of the maximum delay τ. We define the L1 cross-correlation (L1 CC) as follows

$$E_{M2}(\mathbf{m}) = \sum_{s,r} |\tau(x_s, x_r)|,$$

where

$$\tau(x_s, x_r) = max_{\widehat{\tau}} \int_{-\infty}^{\infty} u(x_r, x_s, t - \widehat{\tau}) \, u^*(x_r, x_s) \, dt \quad .$$

Notice that, for implementational purposes, the CC required for τ calculation is performed in the frequency domain.

An important issue to address for CC is how to define an efficient windowing for our dataset. We use an automatic picking function to detect first arrivals. For each trace in u^* we first find the maximum amplitude $u^*_{max}=max_t[|u^*(x_r,x_s,t)|]$, $t\in[0\ T]$. To that goal, we compute the starting time for our gaussian-tapered rectangular window as follows

$$t_{start}= \{min(t) \mid |u^*(x_r,x_s,t)| \geq u^*_{max}\,\varepsilon\,\},$$

where ε is a detection threshold. In many real datasets SNR decreases at low frequencies which results in our detection algorithm working best at high frequency ranges. We suggest choosing ε proportional to the length of the time window δ, which we take equal to the inverse of the low-pass filter cutoff frequency. As last step, traces are weighted as follows: the CC time is constrained such that it can not exceed the predicted first arrival time of the synthetic. Whenever this condition is not fulfilled, the trace is muted. This also mitigates the influence of the shadow zone, where data could be dominated by noise. It is worth noticing that this workflow uses the full offset range ($D=\infty$) and all temporal samples.

Hence, M2 is fully parametrized with ε and the window length δ, both of which are global parameters. This results in a very automated procedure, similar to M1. Contrary to M1,

M2 is a model builder, i.e. it allows recovering long-wavelength model features with little spatial resolution.

Line search and frequency stopping criterion

Our line search is based on the assumption that the misfit surface should be mostly convex. Once the gradient and the search direction have been computed, we start a line search procedure by updating the model as follows

$$\mathbf{m}_n^{i+1} = \mathbf{m}_n^{i} + \alpha_0 2^{\,i \times sign[E(m_n^i) - E(m_n^{i-1})]} \cdot \mathbf{S}_n, \; i=0,1,2,$$

where $\mathbf{S}_n$, $\mathbf{m}_n^{i}$ and $E(\mathbf{m}_n^{-1}) = E(\mathbf{m}_n)$ are the search direction, the model at the ith line search iteration, and the misfit value at the nth gradient iteration, respectively. α_0 value should be compatible with the model parametrization used. If the second order polynomial obtained from the set

$$\{E(\mathbf{m}_n^i), \; i=-1,0,1\} \; \text{or} \; \{E(\mathbf{m}_n^i), \; i=0,1,2\}$$

is not convex and $E(\mathbf{m}_n^i) \geq E(\mathbf{m}_n)$, the gradient procedure is stopped. Whenever the Polak-Ribiere coefficient is positive we reset the search direction, thus resulting in a steepest-descent method at the next gradient iteration. Furthermore, in case that the line search with the steepest-descent gradient also fails, we stop the gradient loop and jump to the next frequency.

RESULTS

We will evaluate the behavior of the proposed procedures M2 and M1 by means of two case studies based on 2-D acoustic media: The Marmousi model with a starting frequency above 6 Hz, and the challenging 2014 SEG Chevron dataset.

Case study: Marmousi

Our first case study is based on the Marmousi model (Bourgeois et al., 1991). Although this is a well-known model, which can be inverted with FWI easily given a good initial model, we propose a more challenging experiment, where we remove low-frequencies from the data and use a very poor initial model, in order to highlight the advantages of our FWI strategy under suboptimal conditions. For such model, the section is 10 km long and 3 km deep. The acquisition geometry consists in 64 sources equally spaced along the *x* direction starting at range 1 km, with 192 fixed receivers separated 50 m from each other. Sources and receivers are at depth 20 m and recording time is T=8 s. We use a Ricker wavelet with central frequency 20 Hz for the sources and then high-pass filter both source wavelet and data at 4 Hz, in order to remove as much as possible information from the lowest frequency. Next, we add white noise to the data ensuring that no useful signal remains below 4 Hz. In Fig. 1 (Top), we present the original acoustic velocity model used to generate the synthetic data set, and the starting model for inversion in Fig. 1 (Bottom). Figure 2 presents one shot gather (namely shotgather 32) low-pass filtered at 4, 6.2, and 8.2 Hz, respectively and without any filter applied. We observe that the trace signal becomes cleaner as frequency increases. In Figure 3, we present the Signal to Noise Ratio (SNR) for traces located at 100, 1000, and 10000 m respectively. The measurement indicates that for all three traces SNR~1 below 3 Hz, and SNR<5 below 4 Hz. Thus we can conclude that our

synthetic dataset is dominated by noise below 3 Hz and also that FWI inversion could hardly be started for frequencies below 4 Hz. Notice, that the starting model is a velocity gradient, ranging from 1500 m/s at the surface to 4500 m/s at bottom, which is a rather poor initial model.

We start the first case study by inverting the model using only M1, that is a classic FWI approach. This strategy consists in low-pass filtering the data, following a multi-scale approach (Bunks et al., 1995). At each selected frequency, we apply an 8th order minimum-phase Butterworth filter to both source and data. In this case, the inversion starts at 6.2 Hz. Also, we resample the parameter models and temporal traces according to the smaller wavelength resulting from the filtered wavelet used for simulation; therefore, sources and receivers are not coincident with the mesh grid and are interpolated.

Figure 4a shows the inversión results using $D = \infty$ for the DO strategy, while Figure 4b shows the inverted model when using the DO strategy with $D=2 \cdot \max(V_p)$ as the offset limiter. After 70 iterations, the algorithm is not able to converge towards an acceptable model neither for FO or DO hence we stop it. Clearly, there are gaps between the wavelengths present in the initial model and in the data directly inverted with classical FWI, the latter appearing as superposed reflectors in Fig. 4a and 4b.

In order to improve the results obtained in the first two tests, we try to bridge the gap in wavelengths between the initial and real model by using the M2 procedure instead of M1 as a first stage for FWI. After that, we fall back to a more classical approach to add the small

wavelength components, by means of M1. For this test two realizations were performed, starting at different frequency, summarized in Tables 1 and 2.

Figure 4c shows the results of applying the strategy shown in Table 1, starting at 6.2 Hz and combining M2 and M1. The algorithm is now able to converge towards the true model of Marmousi. The reservoir at the center is correctly imaged, as are the reflectors from the smooth velocity gradient on the left, while fractures are clearly delineated. We can also see the imprint of the tomographic mode resolved by M2 in the bottom corners of the model, that are not covered by the high-frequency refraction sensitivity kernels or first Fresnel zone. In Figure 4d we find that our methodology also gives an acceptable result even when using only data above 8.2 Hz and the same poor initial model, as defined in the strategy shown in Table 2. Nevertheless there is a quality loss with respect to the 6.2 Hz result, mainly due to the narrower sensitivity zone of M2 at higher frequencies. However, the result is a clear improvement with respect to those obtained applying only M1 since the beginning of the realization, i.e. Figures 4a and 4b (top left and right).

These results show the potential of our strategy and improvements with respect to classical FWI (i.e. only M1 with $D = \infty$) for noisy reflection data cases. Comparisons and Quality Control (QC) are simple in this case as the real model is well known (see Fig. 1, Top). Therefore, in the following case study we apply the same methodology of M2+M1, but to a more complex model and dataset.

Case study 2: Chevron SEG 2014

Our second case study applies our M2+M1 strategy to the Chevron SEG 2014[3] blind test. This test is a marine streamer acquisition 2D synthetic dataset, resulting from an isotropic elastic model that remains unknown. The objective is to offer a realistic (albeit 2D) seismic exploration scenario for both academy and industry. For this purpose, Chevron also provided a starting model (Fig. 5 top, invariant in range), together with a velocity log that can be used for quality control (Fig. 6). The total recording time is $T=6$ s. The dataset is characterized by the presence of a frequency dependant noise, with a very low SNR for frequencies below 3 Hz. Furthermore, diving waves are rapidly refracted to the surface, thus limiting the benefits of using long offsets for background model reconstruction (such as travel-time tomography). As a consequence reflected phases are the most reliable part of the dataset. Notice, that only the starting V_p model is part of the benchmark, while V_s and ρ fields remain unknown. In this study case we ignore the density for which we use the constant value 1000 kg/m^3 and perform an acoustic FWI on 400 shots. A quick inspection of the shotgathers shows that the lowest possible starting frequency is 3 Hz; below that the SNR is too low for inversion, somehow similar to the situation in case study 1.

As already seen in the previous case study, whenever relying on reflections for FWI we can take advantage of DO to build the reflectivity model. Nevertheless, we know from the borehole interpretation that the starting model is off and probably its velocities have been overestimated (see Fig. 6). Thus, we apply M2 to correct the background model and then apply M1 with $D=2 \cdot \max(V_p)$, performing 20 iterations per frequency.

---

[3] http://s3.amazonaws.com/open.source.geoscience/open_data/seg_workshop_fwi/seg_workshop_fwi.html

Figure 5 (Bottom) presents the final inverted model following the inversion strategy shown in Table 3. We can observe that FWI performs well, by recovering both fine and large scale structures. Especially, the top and the bottom of the salt body are now clearly defined, with a small dip on its top. Similarly, three gas pockets appear on the shallowest part of the model as well as three clear inclusions at the top part of the salt body.

A crucial point here, is that we did not pre-process the data except for low-pass filtering and windowing when we use L1 CC misfit. We converge to the final model after only 140 iterations. Figure 6 presents the comparison with a borehole located at 39.735 km and depth 1000 m, showing a good agreement between our inverted model and the real model.

At this point we wish to remark the suggested M1 and M2 procedures' parameters, which have worked equally well in both the Marmousi and Chevron test cases. These are

$$D=2\max(V_P), \ T=\max(t) \text{ for M1,}$$

$$\delta=1/f_0, \ \varepsilon=\delta \qquad \text{for M2.}$$

Hence the task of inverting a dataset with FWI is strongly simplified, only requiring a suitable set of frequency bands and selecting the procedures used at each band.

CONCLUSIONS

We have developed two FWI procedures that allow the inversion of seismic data starting from frequencies significantly higher than those associated to the longest wavelengths in the

initial model. The first procedure (M1) applies the classical L2 norm of normalized traces with an optional Dynamic Offset for receiver selection, while the second (M2) uses an L1 misfit of cross-correlated (L1 CC) observed and synthetic traces with full offset. The M2 procedure is well suited to improve very poor initial models using high-frequency data content. It replicates the benefits of traveltime tomography with finite frequency content. The M1 procedure, on the other hand, performs well at retrieving smaller wavelength information. By combining these FWI procedures, we have successfully inverted two benchmark acoustic tests: The Marmousi model for two starting frequencies above 6 Hz, and the SEG 2014 Chevron dataset. In both tests, initial M2 applications reconstruct well the coarsest structures, and then the M1 procedure progressively retrieves most short-wavelength features. According to these results, our FWI combined methodology is very independent on the quality of the initial model and robust for data with poor SNR at low frequencies, with the added benefit of saving of computer resources by a DO reduction of the shot-domain size.

We have shown that successful full waveform inversions can be set up with as few as four parameters. For a given dataset and initial model, we only need to select the frequency bands to be inverted, maximum iterations per band and whether we will apply M1 and/or M2 procedures to each band. Then, the only variables that we need to parameterize are: D and T for M1, and δ and ε for M2. We remark that these are four global parameters and provide simple guidelines to set up their values. Such guidelines have been applied successfully to two completely independent datasets, which pose important challenges in terms of missing low-frequency data, poor initial models or even incorrect physical representation (e.g. elastic

data inverted with constant-density acoustic FWI). We also remark that M2 should be executed prior to M1 in order to reconstruct the background model at least at the earliest stage of the inversion, which seems reasonable. Thus, we conclude that the methodology described here results in an almost automatic geophysical FWI tool that may be readily applied to many datasets with minimal manual interaction.

ACKNOWLEDGMENTS

REFERENCES

Boehm, C., M. Hanzich, J. de la Puente, A. Fichtner, 2016, Wavefield Compression for Adjoint Methods in Full-Waveform Inversion: Geophysics, *81*(6), R385-R397.

Bourgeois, A., M. Bourget, P. Lailly, M. Poulet, P. Ricarte and R. Versteeg, 1991, Marmousi, model and data:, The Marmousi Experience, 5-16.

Brossier, R., S. Operto and J. Virieux, 2010, Which data residual norm for robust elastic frequency-domain full waveform inversion?: Geophysics, *75*(3), 37-46.

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

Fichtner, A., 2010, Full Seismic Waveform Modelling and Inversion. Springer Science & Business Media.

Jiménez Tejero, C.E., D. Dagnino, V. Sallarès, and C.R. Ranero, 2015, Comparative study of objective functions to overcome noise and bandwidth limitations in full waveform inversion: Geophysical Journal International, 203, 632-645.

Kormann, J. A., Rodríguez, M. Ferrer, A. Farrés, N. Gutiérrez, J. de la Puente, M. Hanzich, and J. M. Cela, 2016, Acceleration strategies for elastic full waveform inversion workflows in 2D and 3D: Comput. Geosci, Accepted for publication.

Luo, Y., and G. Schuster, 1991, Wave-equation traveltime inversion: Geophysics, 56, 645-653.

Métivier, L., R. Brossier, Q. Mérigot, E. Oudet and J. Virieux, 2016, Measuring the misfit between seismograms using an optimal transport distance: application to full waveform inversion. Geophysical Journal International, *205*(1), 345-377.

Nocedal, J. and S. Wright, 2006, Numerical optimization: Springer Science & Business Media, 2006.

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

Warner, M. and L. Guasch, 2014, Adaptive waveform inversion: Theory. 2014 SEG Annual Meeting. Society of Exploration Geophysicists.

LIST OF FIGURES

Figure 1: *Velocity map from the Marmousi model used to generate the synthetic data set (Top), and the acoustic starting model for FWI (Bottom).*

Figure 2: *Shot 36 of the Marmousi dataset after a low-pass filtering at frequencies 4 Hz, 6.2 Hz and 8.2 Hz, compared to the unfiltered shotgather. Clearly, the lowest frequencies are dominated by noise.*

Figure 3: *SNR of traces at different offsets for the shot 32 of the Marmousi dataset.*

Figure 4: *Resulting models for the Marmousi dataset obtained with different FWI strategies: (top left) full offset and classical FWI starting at 6.2 Hz, (top right) dynamic offset control and classical FWI starting at 6.2 Hz, (bottom left) L1 CC FWI with full offset starting at 6.2 Hz, (bottom right) L1 CC FWI with full offset starting at 8.2 Hz.*

Figure 5: *Comparison between starting and final inverted Vp model obtained with the strategy of Table 1. The top of the salt body is very well recover, and the low velocity channel is also correctly picked by our FWI scheme. Three gas pocket are also detected on the upper part of the model: at 13, 20 and 35 Km. A point reflector is correctly imaged under the gas pocket at 13 km.*

Figure 6: *Comparison between the borehole data provided by CHEVRON (red), the starting model (dash) and the final inverted model (blue). There is a good agreement between the final inverted model and the borehole measurement, the low-frequency content matching is correct.*

LIST OF TABLES

Table 1: *FWI M2+M1 strategy applied to the Marmousi dataset starting at 6.2 Hz.*

| Frequency	6.2 Hz	6.2 Hz	8.2 Hz	10.0 Hz
Procedure	M2	M1	M1	M1
Window length δ	0.16 s	-	-	-
Window threshold ε	16%	-	-	-

Table 2: *FWI M2+M1 strategy applied to the Marmousi dataset starting at 8.2 Hz.*

Frequency	8.2 Hz	8.2 Hz	10.0 Hz
Procedure	M2	M1	M1
Window length δ	0.12 s	-	-
Window threshold ε	12%	-	-

Table 3: *FWI strategy for the SEG 2014 Chevron benchmark.*

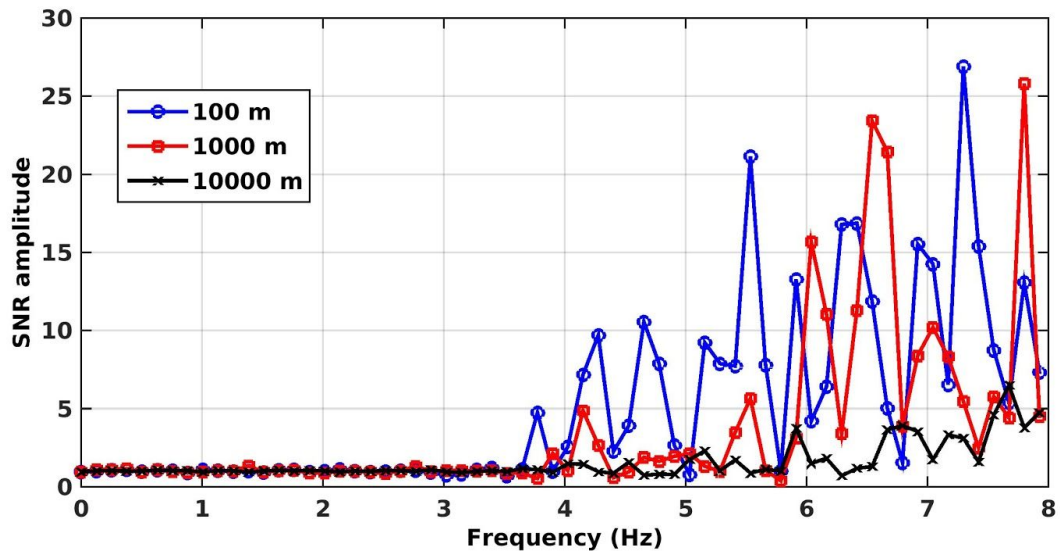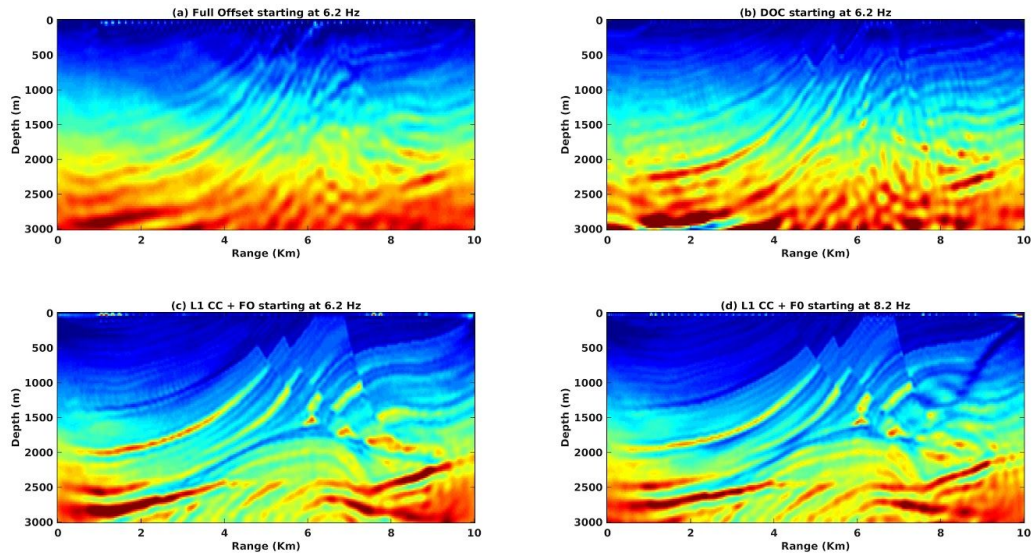|	3 Hz	3 Hz	4Hz	6 Hz	8 Hz	10 Hz	12 Hz
Procedure	M2	M1	M1	M1	M1	M1	M1
Window length δ	0.33	-	-	-	-	-	-
Window threshold ε	33%	-	-	-	-	-	-
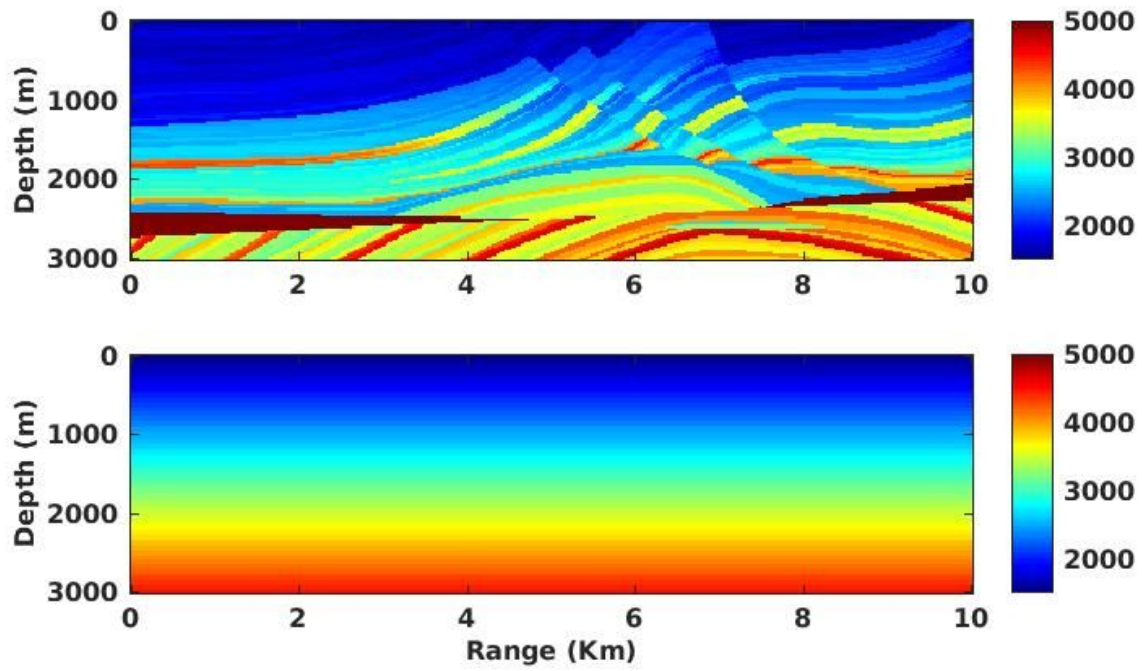
FIGURES

Figure 1: *Velocity map from the Marmousi model used to generate the synthetic data set (Top), and the acoustic starting model for FWI (Bottom).*

Figure 2: *Shot 32 of the Marmousi dataset after a low-pass filtering at frequencies 4 Hz, 6.2 Hz and 8.2 Hz, compared to the unfiltered shotgather. Clearly, the lowest frequencies are dominated by noise.*
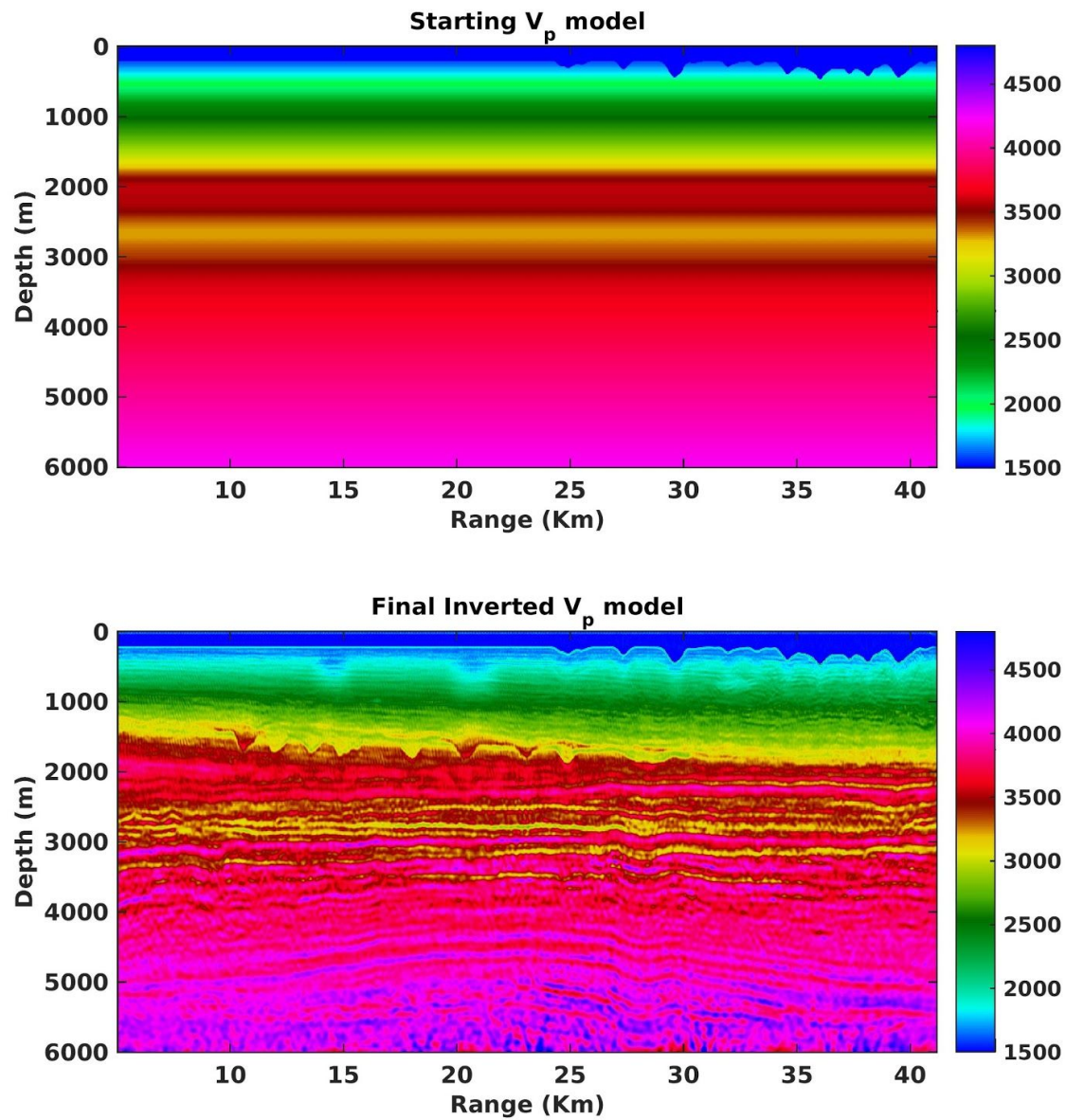
Figure 3: *SNR of traces at different offsets for the shot 32 of the Marmousi dataset.*
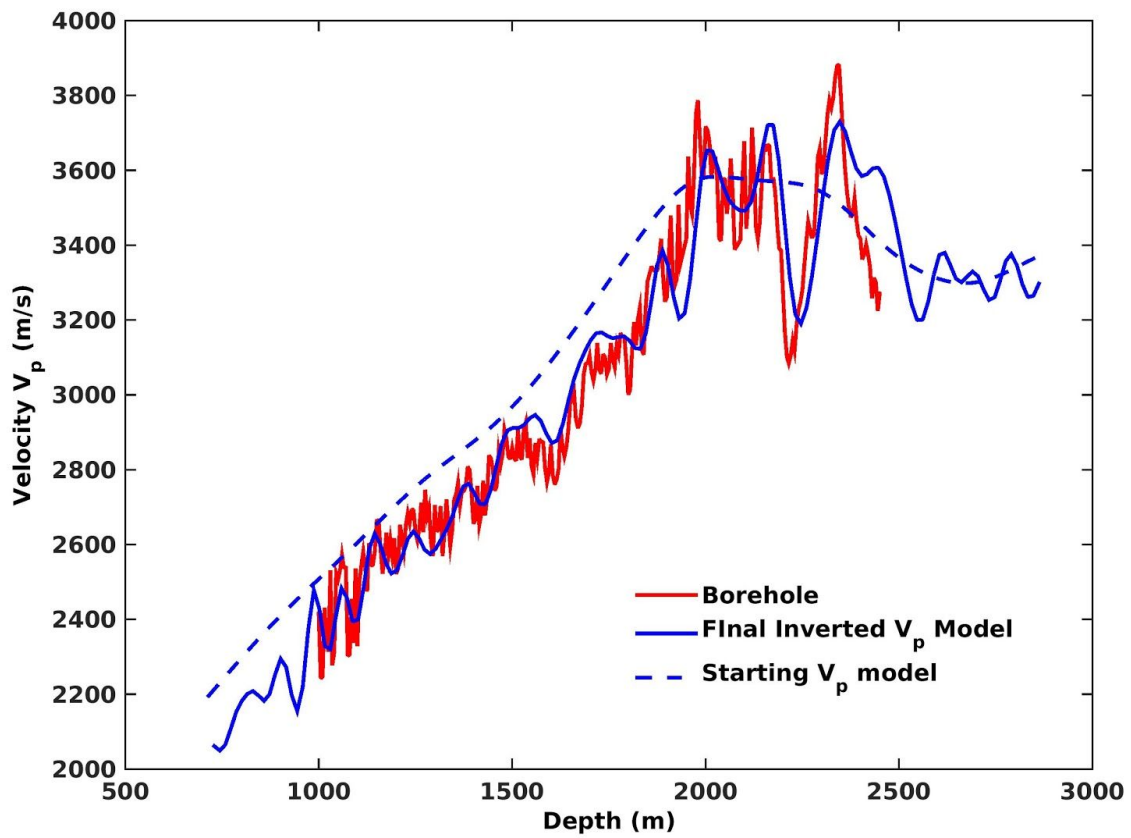
Figure 4: *Resulting models for the Marmousi dataset obtained with different FWI strategies: (top left) full offset and classical FWI starting at 6.2 Hz, (top right) dynamic offset control and classical FWI starting at 6.2 Hz, (bottom left) L1 CC FWI with full offset starting at 6.2 Hz, (bottom right) L1 CC FWI with full offset starting at 8.2 Hz.*

Starting V$_p$ model

Final Inverted V$_p$ model

Figure 5. *Comparison between starting and final inverted Vp model obtained with the strategy of Table 1. The top of the salt body is very well recover, and the low velocity channel is also correctly picked by our FWI scheme. Three gas pocket are also detected on the upper part of the model: at 13, 20 and 35 Km. A point reflector is correctly imaged under the gas pocket at 13 km.*

Figure 6: *Comparison between the borehole data provided by CHEVRON (red), the starting model (dash) and the final inverted model (blue). There is a good agreement between the final inverted model and the borehole measurement, the low-frequency content matching is correct.*