FWI cycle skipping mitigation with an intermediate data approach based on dynamic time warping

Claus Naves Eikmeier*[1,2], Ernani Vitillo Volpe[1,2] and Carlos Alberto Moreno Chaves[2], [1]Research Centre for Gas Innovation, [2]Universidade de São Paulo

This paper was prepared for presentation during the 17th International Congress of the Brazilian Geophysical Society held in Rio de Janeiro, Brazil, 16-19 August 2021.

Abstract

In this work, we present an intermediate data approach based on the Dynamic Time Warping (DTW) method to avoid the cycle skipping effect in full-waveform inversion. We carried out two acoustic inversions with a portion of the Marmousi 2 as the true model and an extremely smoothed version of it as the original one. This provides a challenging inversion case since it is prone to cycle skipping. Whereas in the first numerical experiment we only use the multiscale approach to mitigate cycle skipping, in the second one we combine the multiscale with intermediate data based on DTW. Based on these preliminary results, we believe that the use of DTW-based intermediate data is a promising strategy to mitigate cycle skipping since our second experiment shows improved results compared to the first one.

Introduction

Among the several challenges that are faced in the Full Waveform Inversion - FWI (Tarantola, 1984), the mitigation of cycle skipping is essential to a successful data inversion. When gradient–based methods are used to minimize the objective function, frequently the minimization process converges to a local minimum instead of reaching the global one. This characterizes the so-called cycle skipping effect. The number of local minima increases significantly in seismic data with higher frequency, which makes the function extremely nonconvex. Without cycle skipping mitigation techniques, the initial inversion model must be close enough to the true model to avoid convergence problems. However, good initial guesses are not always available. Thus, to overcome this problem, we aim to develop a new strategy to mitigate the cycle skipping effect, by means of an intermediate data approach, which is based on the Dynamic Time Warping (DTW) method.

The most successful cycle skipping mitigation method to date has been the multiscale strategy (Bunks et al., 1995). Even other methods have made use of that strategy in their own processing routines, owing to its efficiency and ease of application. Despite its good performance, this strategy has an important limitation when frequencies below a value, usually 5Hz, are not present in the data (Bunks et al., 1995). Another limitation is when the geological environment is complex, and there is no satisfactory initial model (Wang et al., 2016).

The DTW method has proved to be a powerful tool for comparing and aligning nonlinear time series in several areas. The interest of the Geosciences community in the subject is not recent. To the best of our knowledge, the oldest use of the technique in applied seismology dates to the early eighties, with the work of Anderson & Gaby (1983). Anderson & Gaby (1983) present the DTW as a generalization to the cross-correlation method and show several application possibilities, such as for waveform clustering, interpretation of full-wave acoustic well logs and well to well correlation, among others. In the context of seismic inversion, the interest in DTW lies in the cycle skipping effect mitigation (e.g., Ma & Hale, 2012). The DTW method has limitations when the analyzed time series present fast variations in time and space, besides noise contamination. The work by Hale (2013) presents an extension to the traditional DTW to mitigate the mentioned problems. Venstad (2014) demonstrates that the DTW is a powerful tool for measuring time shifts in 4D seismic surveys, where the changes are purely temporal. The author also introduces the use of a stiffness parameter, which helps to reduce spurious warping caused by changes in amplitude and noise.

Our approach is closely related to the studies of Yao et al. (2019) and Wang et al. (2016). Yao et al. (2019) use an intermediate data approach, but without the DTW. The intermediate signals are then obtained only with the first arrival information. However, the authors mention that the use of DTW can be an important alternative to what was presented. The proposal of Wang et al. (2016) is similar to what we present here, that is, they use intermediate data generated by DTW to mitigate the effect of cycle skipping. However, their study presents limited information and, as far as we know, no further studies have been published on the topic. In addition, we are not aware of any available open-source code that is related to this subject. Our work is being developed in Python on the Devito finite-difference platform (Luporini et al., 2018; Louboutin et al., 2019). Therefore, as we proceed with this development, we should be able to collaborate with the Geosciences community in that regard as well.

Method

Dynamic time warping (DTW)

This method is mainly used for comparison and non–linear alignment of two time series. Among its various areas of application, we may cite speech recognition, musical synchronization and seismology. To show the rationale

behind the method, given two time series, $A = (a_1, a_2, \dots, a_N)$ and $B = (b_1, b_2, \dots, b_M)$, the DTW aims to obtain the best alignment (best warping path) between them, with the alignment being expressed through a series $D = (d_1, d_2, \dots, d_L)$, with $d_l = (n_l, m_l) \in [1{:}N] \times [1{:}M]$ for $l \in [1{:}L]$ (Figure 1).

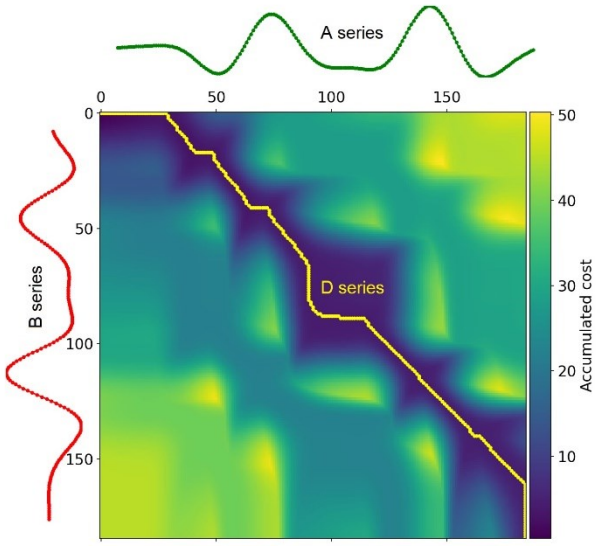

Figure 1 - *Accumulated cost matrix with the best warping path (D series, in yellow) between A (green) and B (red).*

The algorithm behind the DTW is based on dynamic programming. The best warping path is computed through an accumulated cost matrix in which the path with minimal cost is achieved (Müller, 2015).

Generation of intermediate data based on DTW for cycle skipping mitigation

This procedure generates intermediate signals between the initial and the target data, which are less than half a cycle to the former, to avoid cycle skipping. At first, the inversion will occur from the initial to the intermediate data. When convergence is achieved, the current data is then inverted to a new intermediate signal, which now sits between the current and the target data. The process is repeated until the true data is finally recovered.

As for the generation of intermediate signals, it is of interest to impart them with some degree of similarity to the observed and to the current ones. That can be achieved by an interpolation method, which is developed in this study on the basis of DTW. With the optimal warping path provided by the DTW (Figure 2), one knows that, for each element a_n of a given series A, there is a corresponding b_m from another series B, which has the highest degree of similarity to the former. With that information, one can compute the differences in amplitude and time between the two elements, so as to interpolate an intermediate series $z_l \to Z$ between them (Figure 3). For the purpose of cycle skipping mitigation, such procedure should be based on a measure of proximity to the current signal (B series in our example).

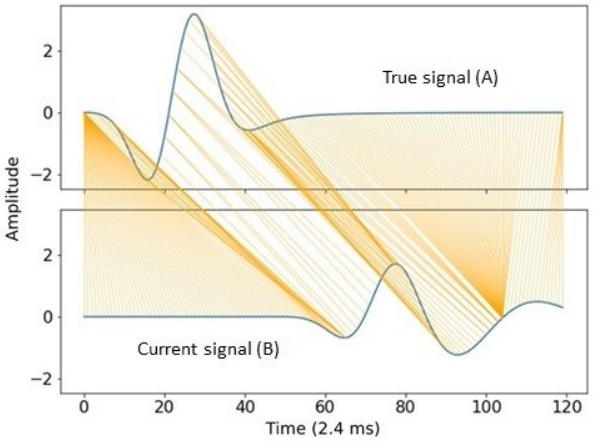

Figure 2 - *The DTW point-to-point alignment between the time series A (true signal) and B (current signal).*

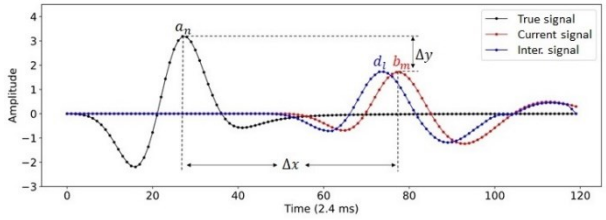

Figure 3 - *The series A (true signal) and B (current signal) with the interpolated D (intermediate signal) with a proximity constraint concerning B.*

Results

Here we present the results of an acoustic seismic inversion with homogeneous density. Two strategies are compared for mitigating the cycle skipping effect. Whereas the first numerical experiment only makes use of the well-known multiscale approach, the second one combines that approach with intermediate data based on the DTW.

The true model is a portion of Marmousi 2 (Figure 4), while the original (initial) one corresponds to an extremely smoothed version of the true model, which was filtered with a Gaussian filter of $\sigma = 75$ (Figure 5). The idea is to create a difficult inversion case, which is clearly prone to cycle skipping.

Ten shots are taken and recorded by 160 receivers, both evenly spaced at the top of the models at a depth of 5m. A Ricker wavelet with a peak frequency of 10Hz represents the source. Although the data has no noise, the true data and the source were both bandpass filtered with cut-off frequencies of 4 and 25Hz to simulate a situation with significant noise outside that band.

Both numerical experiments make use of a mesh with 15m uniform discretization. To avoid numerical dispersion, we

use a 2nd order in time and 30th order in space finite-differences scheme. Moreover, a 600m damping layer is added to the edges of the whole domain to mitigate spurious reflections.

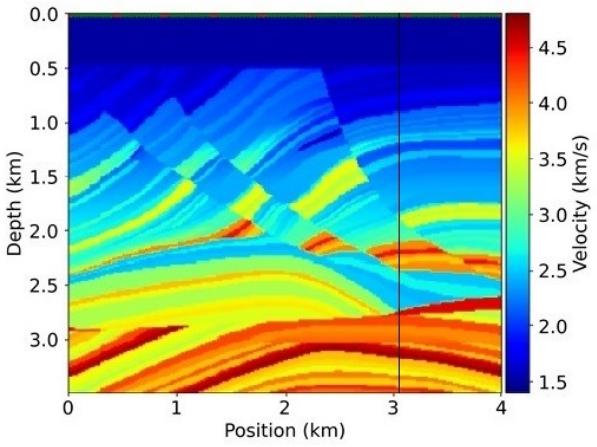

Figure 4 - *A section of the P wave Marmousi 2 velocity model, used as the true model. The black line is the position of the velocity profiles shown in Figures 8 and 9.*

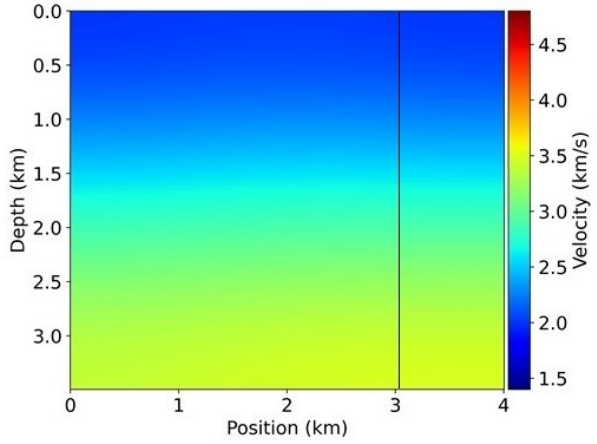

Figure 5 - *The P wave velocity model in Figure 4 subjected to a Gaussian filter with $\sigma = 75$. This was used as the original (initial) model. The black line is the position of the velocity profiles shown in Figures 8 and 9.*

In the first experiment, using the multiscale technique only, the inversion was divided into 4 steps, where the final model of one step was used as the initial model of the next. The low-pass filters correspond to the sequence of 5, 7.5 and 10Hz, whereas the last (4th) step does not involve any filter.

Not only does the second experiment keep the multiscale steps of the first, but it further subdivides each of them into 10 other sub-steps. At each sub-step, intermediate data is interpolated based on the DTW. The idea is to mitigate cycle skipping effects that may still appear when frequencies below 4Hz are unavailable. Furthermore, the DTW is only applied to pairs of seismic signals (observed and initial) within a window, which singles out their relevant portion. For that purpose, we have devised a semi-automatic algorithm that identifies such a window. It scans the signal in both directions, searching for the points where it exceeds a user-specified amplitude threshold. This procedure is repeated at each new multiscale step. For the first experiment, the inversion stopping criterion at each stage is based on a measure of the evolution of the objective function termed as f_tol. We set $f\_tol = 10^{-3}$ so that the inversion stops, and the process moves on to the next step when

$$\frac{f_k - f_{k+1}}{max(|f_k|, |f_{k+1}|, 1)} \leq f\_tol \qquad (1)$$

where f_k is the value of the objective function at iteration k. The same criterion is used in the second experiment, with the difference that now we have two of these parameters. The first is called $inter\_ftol$, which is set 10^{-2}, and the other, $inter\_ftol\_ending$, 10^{-3}. Both work in the same way as f_tol, except that $inter\_ftol\_ending$ is for the last intermediate data sub-step and $inter\_ftol$ is for all the others. This procedure was implemented so that we do not have an unnecessary refinement in some sub-steps, saving computational effort. As a result of the stop criterion, the experiments do not have the same number of iterations. The first reaches a total of 159 iterations while the second reaches a total of 422 iterations.

Figures 6 and 7 present the velocity models obtained after the inversions of the first and second experiment, respectively. We notice that both final velocity models present different features between them. The first model does not resemble the true model at all, while in the second some features were recovered. Yet, despite some convergence of the second model to the true one, the deeper structures remain poorly resolved. We believe that this can be ascribed to an illumination problem, which we plan to tackle in the next experiments.

The same conclusions one draws from Figures 6 and 7 can be reached from Figures 8 (first experiment) and 9 (second experiment), which shows the velocity profiles along a vertical cut at $x = 3.015km$ of the true, original and the final velocity models. Although neither experiment recovers the deeper structures of the true model with acceptable accuracy, the second experiment presents a model that correlates better with the true model than the first.

As one compares the evolution of seismic signals across several iterations of both experiments (Figures 10, 11 and 12), it becomes clear how cycle skipping mitigation affects the final data, and why it presents a better correlation with ground truth in the second experiment, as compared to the first.

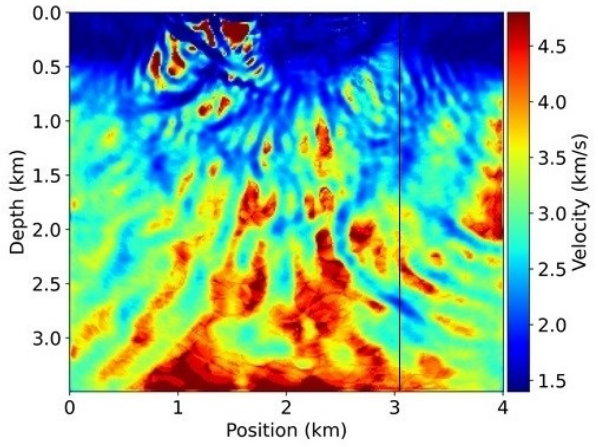

Figure 6 - *Final P wave velocity model of the first experiment. The black line is the position of the velocity profiles shown in Figure 8.*

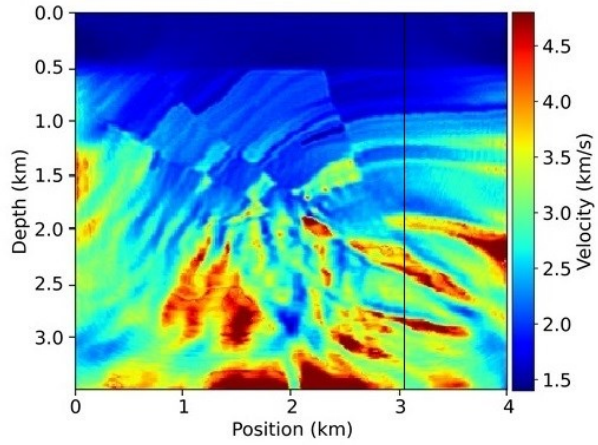

Figure 7 - *Final P wave velocity model of the second experiment. The black line is the position of the velocity profiles shown in Figure 9.*

The signal from the last receiver, at the top right corner, which corresponds to the first shot taken at the top left corner, is shown in Figure 10, 11 and 12. Each image shows the initial signal for a particular step or sub-step, its own target (goal) signal – *i.e.*, the corresponding filtered version of the true signal – and the result, or current signal of the iteration.

Figure 10 shows the signals before the first iteration in the first multiscale step, with the data filtered by a 5Hz low-pass filter. It is not possible to see the initial signal because it is covered by the current one. For the first experiment the observed data is the goal signal and for the second the interpolated.

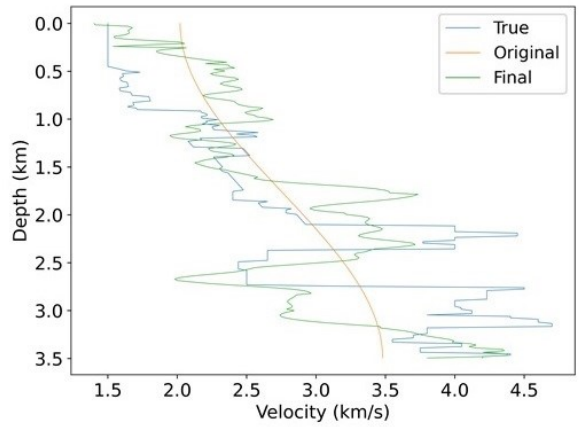

Figure 8 - *Velocity profiles from the experiment 1. The position of the profiles is $x = 3.015km$, that is, the black line in the velocity model figures.*

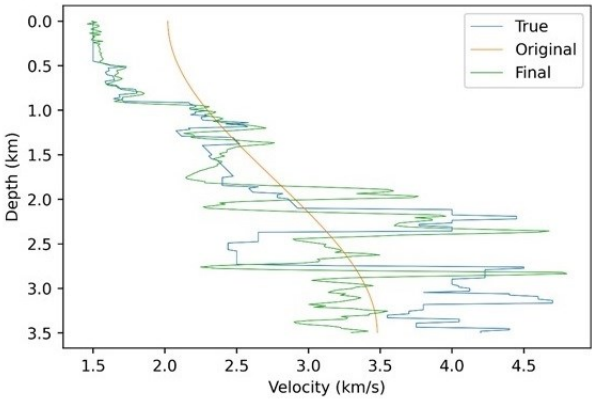

Figure 9 - *Velocity profiles from the experiment 2. The position of the profiles is $x = 3.015km$, that is, the black line in the velocity model figures.*

The signals after the complete inversion are shown in Figures 11 and 12 for experiments 1 and 2, respectively. One already sees a better convergence of the current signal to the goal signal in the second experiment, as compared to the first.

Conclusions

The results presented in this work showed a situation of a seismic data full waveform inversion in which the cycle skipping mitigation strategy, with intermediate data based on dynamic time warping, presented much better results than those obtained by the traditional multiscale technique. Several tests are still needed to confirm that the new strategy can deal with real data, but these preliminary results show potential in this direction.

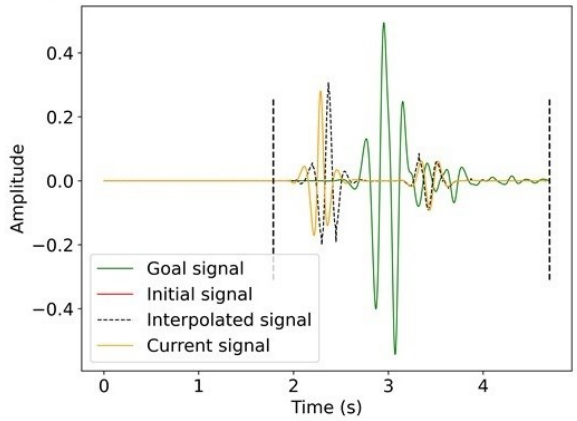

Figure 10 - *Signals from shot position 0 km and receiver position 4 km, before the beginning of the inversion. As it is the first multiscale step the data is filtered with a 5 Hz low-pass frequency filter. It is not possible to see the initial signal because it is overlapped by the current signal. The black dashed signal is the first intermediate signal obtained with the interpolation process based on the DTW. For the first experiment, initially, the observed data is the goal (green) and for the second experiment it is the interpolated (black dashed). The two vertical black dashed lines are the limits of the window in which the DTW algorithm acted.*

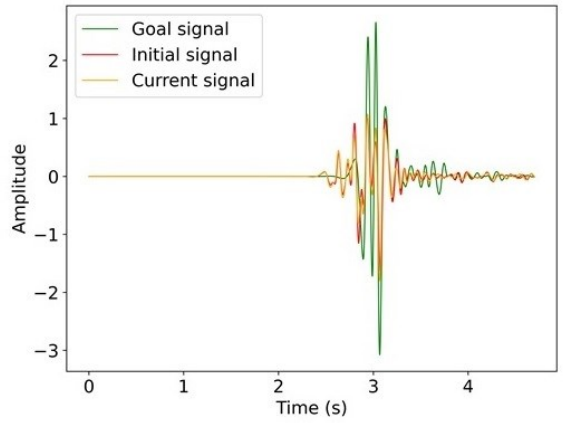

Figure 11 - *Final inversion signals of the first experiment from shot position 0 km and receiver position 4 km. As it is the last multiscale step, that is, without frequency filters, the goal signal (green) is the true, the red the initial for this step and the current (yellow) the final signal.*

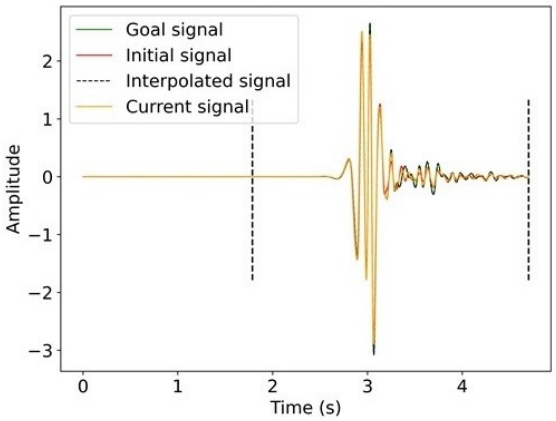

Figure 12 - *Final inversion signals of the second experiment from shot position 0 km and receiver position 4 km. As it is the last multiscale step, that is, without frequency filters, the goal signal (green) is the true, the red the initial for this sub-step and the current (yellow) the final signal. The two vertical black dashed lines are the limits of the window in which the DTW algorithm acted.*

Acknowledgments

We gratefully acknowledge support of the RCGI – Research Centre for Gas Innovation, hosted by the University of São Paulo (USP) and sponsored by FAPESP – São Paulo Research Foundation (2014/50279-4) and Shell Brasil, and the strategic importance of the support given by ANP (Brazil's National Oil, Natural Gas and Biofuels Agency) through the R&D levy regulation.

This research was carried out in association with the ongoing R&D project registered as ANP 20714-2, "Software technologies for modelling and inversion, with applications in seismic imaging" (University of São Paulo / Shell Brasil / ANP) – Desenvolvimento de técnicas numéricas e software para problemas de inversão com aplicações em processamento sísmico, sponsored by Shell Brasil under the ANP R&D levy as "Compromisso de Investimentos com Pesquisa e Desenvolvimento".

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

References

ANDERSON, K. R.; GABY, J. E. 1983. Dynamic waveform matching. Information Sciences, 31(3): 221-242. DOI: https://doi.org/10.1016/0020-0255(83)90054-3

BUNKS, C.; SALECK, F. M.; ZALESKI, S.; CHAVENT, G. 1995. Multiscale seismic waveform inversion. *Geophysics,* 60(5): 1457-1473. DOI: https://doi.org/10.1190/1.1443880

LOUBOUTIN, M.; LANGE, M.; LUPORINI, F.; KUKREJA, N.; WITTE, P. A.; HERRMANN, F. J.; VELESKO, P.;

GORMAN, G. J. 2019. Devito (v3.1.0): an embedded domain-specific language for finite differences and geophysical exploration. Geoscientific Model Development 12(3): 1165-1187. DOI: https://doi.org/10.5194/gmd-12-1165-2019

LUPORINI, F.; LANGE, M.; LOUBOUTIN, M.; KUKREJA, N.; HÜCKELHEIM, J.; YOUNT, C.; WITTE, P.; KELLY, P. H. J.; HERRMANN, F. J.; GORMAN, G. J. 2018. Architecture and performance of devito, a system for automated stencil computation. CoRR abs/1807.03032. DOI: https://doi.org/10.1145/3374916

HALE, D. 2013. Dynamic warping of seismic images. Geophysics, 78(2): S105-S115. DOI: https://doi.org/10.1190/geo2012-0327.1

MA, Y.; HALE, D. 2012. Full waveform inversion with dynamic image warping. Technical Report CWP-707, Center for Wave Phenomena, Colorado School of Mines.

MÜLLER, M. 2015. Fundamentals of Music Processing, Springer International Publishing. DOI: https://doi.org/10.1007/978-3-319-21945-5

TARANTOLA, A. 1984. Inversion of seismic reflection data in the acoustic approximation. Geophysics, 49(8): 1259-1266. DOI: https://doi.org/10.1190/1.1441754

VENSTAD, J. M. 2014. Dynamic time warping - an improved method for 4d and tomography time shift estimation? Geophysics, 79(5): R209-R220. DOI: https://doi.org/10.1190/geo2013-0239.1

WANG, M.; XIE, Y.; XU, W. Q.; LOH, F. C.; XIN, K.; CHUAH, B. L.; MANNING, T.; WOLFARTH, S. 2016. Dynamic-warping full-waveform inversion to overcome cycle skipping. SEG Technical Program Expanded Abstracts 2016, Society of Exploration Geophysicists. DOI: https://doi.org/10.1190/segam2016-13855951.1

YAO, G.; DA SILVA, N. V.; WARNER, M.; WU, D.; YANG, C. 2019. Tackling cycle skipping in full-waveform inversion with intermediate data. Geophysics, 84(3): R411-R427. DOI: https://doi.org/10.1190/geo2018-0096.1