Robust estimation of primaries by sparse inversion and Marchenko equation-based workflow for multiple suppression in the case of a shallow water layer and a complex overburden

A 2D case study in the Arabian Gulf

Staring, Myrna; Dukalski, Marcin; Belonosov, Mikhail; Baardman, Rolf H.; Yoo, Jewoo; Hegge, Rob F.; Borselen, Roald van; Wapenaar, Kees

DOI
10.1190/geo2020-0204.1

Publication date
2021

Document Version
Accepted author manuscript

Published in
Geophysics

Robust estimation of primaries by sparse inversion

~~R-EPSI~~ and Marchenko equation-based workflow for multiple suppression in the case of a shallow water layer and a complex overburden: a 2D case study in the Arabian Gulf

Myrna Staring[*], *Marcin Dukalski*[†], *Mikhail Belonosov*[†], *Rolf H. Baardman*[†], *Jewoo Yoo*[†], *Rob F. Hegge*[†], *Roald van Borselen*[†], *and Kees Wapenaar*[‡]

ABSTRACT

Suppression of surface-related and internal multiples is an outstanding challenge in seismic data processing. The former is particularly difficult in shallow water, whereas the latter is problematic for targets buried under complex, highly scattering overburdens. Here we propose a two-step, amplitude- and phase-preserving, inversion-based workflow, which addresses these problems. We apply Robust Estimation of Primaries by Sparse Inversion (R-EPSI) to solve simultaneously for the surface-related primaries Green's function and the source wavelet. A significant advantage of the inversion approach of the R-EPSI method is that it does not rely on an adaptive subtraction step that typically limits other de-multiple methods such as SRME. The resulting Green's function is used as input to a Marchenko equation-based approach to predict the complex interference pattern of all overburden-generated internal multiples at once. In this approach, no a priori information about the subsurface is needed. In theory, the interbed multiples can be predicted with correct amplitude and phase and, again, no adaptive filters are required. We illustrate this workflow by applying it on an Arabian Gulf field data example. It is crucial that all pre-processing steps are performed in an amplitude preserving way to restrict any impact on the accuracy of the multiple prediction. In practice, some minor inaccuracies in the processing flow may end up as prediction errors that need to be corrected for. Hence, we did conclude that the use of conservative adaptive filters were necessary to obtain the best results after interbed multiple removal. The obtained results show promising suppression of both surface-related and interbed multiples.

INTRODUCTION

Marine seismic processing and imaging can be very challenging in settings with shallow water and complex overburden. Typically in these settings, many orders of strong surface-related and internal multiples interfere with the primary reflections in the target area. This interference results in an image that may no longer be suitable for structural or quantitative interpretation. This is a well-recognized issue in various areas around the world, e.g., in Norway (Rønholt et al., 2014), the West Central Graben of the North Sea (Yanez et al., 2018) or offshore Middle East (Belonosov et al., 2019). The multiple problem in the latter area is particularly challenging due to a strongly scattering overburden (El-Emam et al., 2011, 2015). Addressing this issue in a fit for purpose and practically achievable de-multiple workflow poses important technical and practical challenges.

Geophysicists discriminate between surface-related and internal multiples based on the location of sources and receivers relative to the multiple generator (an interface creating downward reflections) as well as on the amount of a priori assumed or available information. Surface-related multiples have a known generator – a stationary and flat water-air interface at $z = 0$ with a reflectivity of about $r_0 = -1$, with some deviations in rough sea conditions (Kryvohuz and Campman, 2019). Internal multiples are generated by any reflector in the subsurface, with unknown position and reflectivity.

These differences make the two kinds of multiples, depending on the geological setting, very different in terms of their strength, character and removal in subsequent processing. Surface-related multiples tend to be much stronger (*the first order de-multiple problem*). However, internal multiples, though always individually weaker, are more numerous. Collectively, they often result in a complex, hazy and, on occasion, strong interference pattern, rather than individual, easy-to-discriminate events. In contrast, individual internal multiples tend to stand out more when two dominant reflectors (e.g. water bottom or a salt body interface) are present in the subsurface. A wide range of potential purpose- or area-specific de-multiple approaches exist. One that is commonly applied to both is prediction and, in the absence of a good wavelet estimate (a source and a receiver response), adaptive subtraction. The latter, however, does not offer any guarantees of successful multiple attenuation (Verschuur, 2013) and should be applied with great care.

Many approaches exist aimed at tackling surface-related multiples. Move-out discrimination between high velocity primaries and lower velocity surface-related multiples is still used frequently, although this approach is only applicable in simple enough geological settings. These techniques result in many artifacts, damage to the primary reflections and fail in shallow water where the water column does not result in a sufficient move-out difference between surface-related multiples and surface-related primaries (primary reflections and internal multiples). Predictive deconvolution approaches, with their many intrinsic assumptions about the medium, have been proposed and tested in a shallow water setting (Biersteker, 2001). Unfortunately, they may be unsuitable in combination with more advanced internal de-multiple schemes, as they tend to artificially whiten (or otherwise alter) the surface-related primaries output amplitude spectrum.

More sophisticated approaches aim to predict the multiples using either model-based methods (Lokshtanov, 1999; Jahdhami et al., 2017) that require a priori knowledge of the subsurface, or data-driven methods such as Surface-Related Multiple Elimination (SRME) (Verschuur et al., 1992). Both of these rely on some implementation of adaptive subtraction, and also may impact primary reflections and hence result in images that may not be suitable for AVO analysis. An improved multiple model can be obtained if a good source wavelet estimate is available. This estimate can be obtained explicitly by modeling airgun arrays, by reading it off a simple and well consolidated water bottom reflection or by simultaneously solving for it (van Groenestijn and Verschuur, 2009; Savels et al., 2011). Unfortunately, these approaches no longer work in the shallow water environment (Hargreaves, 2006; Hung et al., 2010; Barnes et al., 2014) due to interference (and thus inseparability) of primaries and the first (few) order(s) of (short-period) surface-related multiples, and missing near-offset data.

Internal multiple removal in general tends to be more challenging and to this day there are only a few approaches with limited applicability. The use of move-out discrimination between primary reflections and internal multiples is impossible in practice due to relatively small timing differences. Predictive deconvolution or filtering approaches (Lesnikov and Owusu, 2011) can fail in sufficiently complex settings (Griffiths et al., 2011; Brookes, 2011) as the periodicity they assume is not necessarily physical. Applications of post-stack dip filtering tends to affect anticlines and low relief structures (Retailleau et al., 2014) for flat strongly scattering overburdens. Model-based (Lokshtanov, 1999; Jahdhami et al., 2017), or data-driven (e.g. Jakubowicz, 1998; Weglein et al., 1997) wave equation based methods predict kinematically-accurate individual events. However, they rely on understanding the interbed generation mechanisms or on the separability of individual primary reflections (internal multiple generators). Failure to do so results not just in amplitude but more importantly phase prediction errors. Layer stripping can be used to account for more generation mechanisms, however this approach tends to be more computationally expensive (even in 2.5-D media), and can result in amplitude damage to deeper primaries resulting in a downward cascading, and likely cumulative, prediction-and-subtraction error. Ways of partially mitigating this problem have been suggested by Ypma and Verschuur (2013).

Recently, two new de-multiple technologies have been proposed that aim to address many of the aforementioned issues. First, Lin and Herrmann (2013) proposed Robust Estimation of Primaries by Sparse Inversion (R-EPSI), which simultaneously solves for the source wavelet and the surface-related primaries Green's function. This approach had been tested on field data (Czyczula-Rudjord et al., 2015) and was deemed well-suited for shallow water (Belonosov and van Borselen, 2017). Second, recently proposed Marchenko equation-based approaches (henceforth referred to as the *Marchenko methods*) could be seen as *a form* of internal de-multiple, where the entire overburden is treated as a single collective internal multiple generator (Slob et al., 2014; Wapenaar et al., 2014; van der Neut and Wapenaar, 2016). The Marchenko methods, by solving an under-constrained inverse problem, produce an inverse transmission response of the overburden. The latter are then (indirectly) used in an Amundsen et al. (2001)-like multidimensional deconvolution (MDD) to remove the effect of overburden-borne reverberations on the target primary reflections. Thus far, the Marchenko equation-based approach had been tested in settings where source signature estimation was not a limiting factor (Staring et al., 2018; Pereira et al., 2018, 2019). Alternatively, given multicomponent data, Ravasi et al. (2016) suggest deconvolving the upgoing and downgoing wavefields, which results in data free from (a) surface-related multiples and (b) the source signature.

Both de-multiple problems could in principle be handled by merging SRME- and Marchenko-type methods. This requires some modifications to the Marchenko equation (Ware and Aki, 1969; Singh et al., 2016). Unfortunately, these approaches are not likely to work very well in practice, since not only they struggle converging in the presence of strong multiples (Dukalski and de Vos, 2017), but, more importantly, they still need a good wavelet estimate on input. To address this point, Ravasi (2017) proposes an alternative approach that merges the Amundsen et al. (2001)-MDD with the Marchenko equations. This approach, however, requires both the up- and the down-going field measurement and relative hydrophone and geophone amplitude calibration. Additionally, in a shallow water setting, this scheme might be more difficult to implement due to limited (or saturated) dynamic range of the receivers by the direct (down-going) wave. Having exhausted the list of available possibilities, perhaps the combined R-EPSI and Marchenko equation-based approaches could work together in theory and, more importantly, *in practice*?

This work is structured as follows. The theory section following this introduction briefly discusses a Marchenko equation-based method and studies its strengths and limitations. For more detail, readers can refer to existing literature on the topic. We propose a practical shortcut to predict (and later adaptively subtract) the internal multiples and avoid the post-Marchenko equation MDD step. Next, we explain why R-EPSI output could be a very good fit to be the input to the Marchenko equation. We then show how the proposed workflow can be applied to field data from the Arabian Gulf. Lastly, we discuss the results and some open questions.

THEORY

In a traditional processing workflow, surface-related multiples are treated first (to our knowledge the only notable exception are deep water settings with sub-salt targets) and dominant internal multiples are typically handled afterwards. This sequence is also often assumed by most internal de-multiple schemes, the Marchenko method being no exception. At the cost of stringent input data quality (i.e., the output of the surface-related de-multiple), the Marchenko method can simultaneously handle internal multiples due to the entire overburden. To better understand this process, we take an anti-chronological approach and first explain the internal de-multiple before we show how the surface-related de-multiple of our choice fits the Marchenko method.

Marchenko equation-based multiple elimination

The method proposed by van der Neut and Wapenaar (2016) removes the internal multiples generated by the overburden from the surface reflection data. By staying at the surface, this method is independent of background velocity information and helps in gaining a better understanding of the imprint of internal multiples as compared to Marchenko redatuming (Wapenaar et al., 2014). Moreover, much like Jakubowicz IME, it only uses convolutions, correlations, and temporal windowing of the reflection data. We propose to combine the method of van der Neut and Wapenaar (2016) with the double focusing approach proposed by Staring et al. (2018). Moreover, we suggest a viable quality control (QC) step to increase confidence in the data preprocessing.

The reflection data R_0 contain the target reflection response R_{tar} past a specified source- and receiver-dependent onset time $t_2(x_r, x_s)$. Due to the overlying overburden, the latter response is not only covered by overburden-only multiple reflections, but it is also dressed with overburden-borne internal-multiple-coda reverberations, which we denote with $(v^+)^{-1}$. Van der Neut and Wapenaar (2016) and Elison et al. (2020) show that v^+, a Marchenko focusing function f_1^+ which is “propagated” or “projected” to the surface, is a solution to the (here already reduced) coupled set of Marchenko equations

$$\Theta_\epsilon^{t_2}\left[v^+\right] - \Theta_\epsilon^{t_2}\left[R_0^*\Theta_\epsilon^{t_2}\left[R_0 v^+\right]\right] = 0\,. \tag{1}$$

Above, we use a detail hiding notation, where ”multiplications” denote Rayleigh integrals (Berkhout, 1982; van der Neut et al., 2015). Moreover, $\Theta_\epsilon^{t_2}$ is a source and receiver dependent mute which preserves the signal inside the $[\epsilon, t_2(x_r, x_s)]$ time window, i.e., after a small time interval ϵ (discussed later) and before t_2 i.e. the onset of target primaries from R_{tar}.

Given $v_e^+ = v^+ - \Theta_\epsilon^{t_2}[v^+]$, equation 1 is easily solved using the nested series

$$v^+ = \sum_{k=0}^{\infty} \Omega_k\left[v_e^+\right], \tag{2}$$

where $\Omega_k[A] = \Theta_\epsilon^{t_2}\left[R_0^* \Theta_\epsilon^{t_2}\left[R_0 \Omega_{k-1}[A]\right]\right]$, $\Omega_0[A] = A$, and *in theory* the first term in the series v_e^+ is given by an identity (in practice convolved with a user-defined wavelet – see discussion on practical limitations below).

The aforementioned authors then proceed to derive auxiliary fields $U^\pm$, where U^- is a reflection response – one without overburden-only multiples and source-side reverberations. The MDD inversion of $U^- = R_{\mathrm{tar}} U^+$ recovers a target-only reflection response (as if the overburden was transparent). Van der Neut and Wapenaar (2016) calculate a low-order series expansion to obtain R_{tar}. However, Elison et al. (2020) show that in the presence of more complex internal multiple scattering, more terms need to be computed in equation 2 and hence in $U^\pm$. Note that in general each term provides amplitude and phase corrections instead of predictions of individual events or their specific order. Individual terms may correspond to predictions of individual events in cases where multiple scattering is dominated by only a few reflectors. In such a case, only the leading order terms will be most relevant, and then one could attempt to adaptively subtract them.

Here we propose an approach combining the strongest points of the methods by van der Neut and Wapenaar (2016) with insights from the double-focusing method by Staring et al. (2018). We propose a method called "double de-reverberation", which evaluates

$$v^{+T} U^- = v^{+T} \Theta_{t_2}^{\infty}\left[R_0 v^+\right], \tag{3}$$

where left multiplication with v^{+T} (source-receiver swapped version of v^+) amounts to removing receiver-side reverberations, such that the result contains the target reflection response R_{tar} and the later arriving internal multiples generated by scattering between the overburden and the target. We can also isolate the multiples predicted by this scheme via

$$v^{+T} \Theta_{t_2}^{\infty}\left[R_0 v^+\right] - \Theta_{t_2}^{\infty}\left[R_0\right], \tag{4}$$

which we later adaptively subtract, bearing in mind that the term on the right (left) needs to be convolved (deconvolved) with the user-defined wavelet used inside the Marchenko equation solver (equation 2). We wish to note that collecting the $k \leq 1$ terms in equation 2, plugging them into equation 3, and ignoring terms higher than third order in R_0, yields two terms, where one conforms with that one suggested by van der Neut and Wapenaar (2016).

There are two major, somewhat hidden, assumptions in the scheme presented above, which are very likely to be violated by real complex finely layered overburdens. Firstly, equation 1 (and hence its solution equation 2) only holds if the time pick t_2 does not cut into either a target or an overburden primary. In the event that this assumption is not satisfied, the magnitude of the resultant artifact is expected to be proportional to the magnitude of the cut-into primary event and will be spread over the entire multiple prediction rather than localized. Secondly, the (infinitesimal in theory) ϵ parameter delineating the boundary between the initial guess v_e^+ and the rest of the v^+ solution, needs to be set to about half-wavelet length when the scheme is applied to band-limited data. This means that in the presence of short-period internal multiples (SPIMs-those generated by reflectors spaced by less than the dominant wavelength anywhere in the overburden), the estimate v_e^+ also needs to contain part of the solution responsible for attenuating them (Slob et al., 2014). Although the scheme is somewhat robust to these challenges (Elison et al., 2020) and currently no suggestions have been proposed to address the first problem, the SPIMs challenge however can be mitigated in cases where the overburden is sufficiently flat (Dukalski et al., 2019; Elison et al., 2020), but this will not be considered here. Lastly, our Marchenko method is acoustic and its extensions to include elastodynamic effects are even more sensitive to the aforementioned effects (Reinicke et al., 2019). Even though we may express the extent of the expected artifacts qualitatively and quantitatively (Reinicke et al., 2020a), they still require further development and further tests on sufficiently complex synthetics.

Marchenko equation-based methods are made up of the same fundamental building blocks as many other de-multiple methods, however the inverse-problem-type approach involving summation of many terms (see equations 1- 3) requires a specific reflection data format for the equation to hold. Ravasi et al. (2016) show that a reflection response R_0 satisfying the Marchenko equation requirements can be extracted from multicomponent up- $P^{\uparrow}$ and down- $P^{\downarrow}$ separated measurements by inverting the MDD relation $P^{\uparrow} = RP^{\downarrow}$. In other words, the data have to be defined on a grid of co-located sources and receivers (with some spacing Δx) at the same depth level. In addition, the data need to be wavelet-free, dipole-source, monopole-receiver and multiplied by $2\Delta x$ for equation 1 to be satisfied. In the absence of reliable multicomponent data that had been accurately up- and down-decomposed, this can be challenging. The main challenge therefore is finding the correct wavelet. Its shape can be well approximated by modeling or by any of the methods mentioned in the introduction, however getting the correct scale tends to be much harder. We know however that the series (equation 2) is unconditionally convergent (Dukalski and de Vos, 2017), thus accidentally over-scaling the data might result in a divergent behavior (run-away solutions). This helps impose an upper, but not a lower limit on a typically unknown scaling. As first suggested in Elison et al. (2020), in the presence of additional information (e.g., regional geology or well-log information), one can build a (well-log-based) 1-D synthetic, determine the expected convergence rate, and compare it to one obtained using field data to approximately verify whether the data were appropriately scaled. This however only helps to determine a single (for the whole data set) frequency-independent scalar. In the next section, we show how the reflection response obtained on output from R-EPSI fits the above requirements. In the case-study section, we show how this works in practice and compare the Marchenko equation convergence curve against that obtained from a synthetic showing a good match.

Basics of robust EPSI

R-EPSI estimates a surface-related primaries Green's function directly from the data, without any additional knowledge of the subsurface properties. In a very shallow water setting, the method produces a high fidelity result and outperforms common practice approaches (Belonosov et al., 2019). Hence, its output is a good candidate for input to the aforementioned Marchenko approach. In this section, we briefly outline the R-EPSI method.

The method assumes a grid of seismic dipole sources co-located with monopole receivers at the sea surface, a water surface reflection coefficient of -1 and a source signature which is the same for all sources. Then, a monochromatic component of the upgoing data R_{-1}

(the subscript denotes the reflection coefficient at the free surface) can be explained by (Berkhout, 1982)

$$R_{-1} = R_0 S - R_0 R_{-1}, \tag{5}$$

where R_0 is the unknown surface-related primaries Green's function i.e. the input for the Marchenko method, while S is the unknown effective source wavelet. The first term on the right hand side, $R_0 S$, represents primaries and internal multiples, and the second term the surface-related multiples.

R-EPSI solves equation 5 in the time domain to find R_0 and S given R_{-1}. This problem is unconstrained, and to tackle this issue, R-EPSI imposes conditions on the unknowns. Sparsity is often used as an assumption in geophysical field data applications (van Groenestijn and Verschuur, 2009) and, as a good approximation of this assumption for the Green's function, the candidates for R_0 are restricted to solutions with a L_1-norm time-domain minimum. The source wavelet is estimated as a function that has a limited temporal support i.e. the length of the wavelet.

To solve equation 5, R-EPSI applies an automatic iterative process with only one input parameter which is the level of noise in the data. For the details of this iterative process, we refer to Lin and Herrmann (2013). As mentioned, R-EPSI produces highly accurate surface-related primaries in a very shallow water situation where conventional methods like SRME, based on prediction plus subtraction, may fail (Hung et al., 2010). As a disadvantage, R-EPSI is much more computationally expensive. For each iteration, the method takes the computational runtime of several SRME processes (i.e., several convolutions). The output Green's function R_0 is used directly as input for the Marchenko equation-based approach, explained in the previous section.

CASE STUDY IN THE ARABIAN GULF

Data from the Arabian Gulf provide an excellent illustration of the challenges stemming from a shallow water environment combined with a complex overburden. The water bottom itself shows little variation in depth and contrast (Alá'i et al., 2002). The subsurface consists of a fairly strong sea-bottom reflector at between 25 - 35m depth followed by a thick package of thinly layered deposits of anhydrites, carbonates, and various clastic sediments and subsequent high impedance contrasts (El-Emam et al., 2011), with just minor undulations across larger distances. Hence, the velocities vary primarily vertically.

Acquisition of towed marine streamer data is difficult in shallow water and the associated large gap in the near offsets would have presented an additional challenge for our de-multiple workflow. However, in the area, a 3D wide-azimuth OBC dataset had been acquired with minimum inline and crossline offsets of only 25 m. Extracting a single 2D line of shots and inline receivers makes an ideal dataset for this test case. Further details of the acquisition geometry are provided in Figure 1 where the red oval indicates the extracted 2D line. A raw split-spread shot gather with the ground roll typical for the survey area and various kinds of other noise is shown in Figure 3a (gained and clipped for display). Despite the noise, the effects of shallow water depth and thin layering are already evident in the multitude of similar reflection events in the background. Related images of shot gathers will be shown for some key points of the processing workflow. Later, some RTM images will be shown.

Figure 2 shows a workflow of the processing steps performed in this section. Preprocessing stage 1 started with a standard debubble. The data were converted from 3D to 2D by applying a $\sqrt{t}$ gain, followed by the application of water-column-related static corrections. High-amplitude (linear) noise and mud-roll were filtered out in the frequency domain. The data were decomposed in up- and down-going wavefields by a data-driven method similar to that described in Soubaras (1996). The upgoing pressure wavefield is selected for further processing to make this workflow also suitable for streamer data (in theory). An example of the resulting upgoing pressure wavefield is shown in Figure 3b. The noise introduced above the first break and the effects of one or more dead channels, as indicated by the arrow, needs to be addressed. Nevertheless, the overall quality and continuity of the reflection events has improved. Source ghosts were left in the data to mimic the effect of dipole sources required by R-EPSI and Marchenko, which is reasonable given a source depth of 7 m and mild weather conditions. However, the initial 100 m shot and 50 m receiver spacing along the line did not fulfill their requirements. Hence, the data were regularized and interpolated to a 25 m grid and fixed spread of 239 co-located sources and receivers, thereby removing any acquisition related gaps (see Figure 3c), and muted. This step further reduces random noise. During the full preprocessing sequence, particular attention was given to each step to ensure amplitude and phase fidelity was maintained.

Stage 2 starts by applying a bandpass filter (2-6-38-42 Hz) to avoid aliasing related effects during subsequent processing. The data are now suitable as input for surface-related multiple suppression with R-EPSI. Before its application, and in order to suppress the artefacts related to the limited aperture available, we also applied cosine-squared tapers for 15 traces at each edge. We tested this taper width and chose the smallest value for which the artefacts became negligible. We used 80 ms as the wavelet length. This estimate is based on interpretation of the first arrivals in the data as well as basic dominant frequency analysis. In all our tests, R-EPSI was insensitive to small deviations in the range of about 20 ms to this parameter. As for the noise level in the data, we tried different values starting from 20% and smaller. We noticed that the residual did not change below a noise level of 15%. Hence, we set this parameter to this value. The method converged after about 250 iterations of the SPGL1 solver (van den Berg and Friedlander, 2008, 2019), i.e. mostly updates of the Green's function as the wavelet was updated only 30 times.

The effect of the method on the field data is illustrated by Figures 4 - 6. The shot gather in Figure 4a shows the difficulty of identifying primaries when there is heavy contamination by shallow water reverberations, which are of comparable strength. The result of R-EPSI in Figure 4b is much improved in that respect. The surface-related multiples are generated by the shallow water bottom and are thus present mostly in the near offsets before spreading out as the difference panel in Figure 4c shows. The multiples also appear to be concentrated in a certain frequency band, which is clearly visible on the corresponding FK spectra shown in Figure 5.

The RTM images in Figure 6 show that the local geology is almost horizontally layered. The result after R-EPSI, Figure 6b, reveals likely primary reflection events which stand out at later times. The upper shallow part appears to show a package of many thin layers. Finally, note that in Arabian Gulf-like synthetics with a water layer depth of 25 m (water depth varied between 25 m and 35 m in the field data), the R-EPSI primaries are almost identical to the true primaries (Belonosov et al., 2019). This fact gives us a strong confidence in the R-EPSI result obtained.

Following the above preprocessing and correction steps, the output of stage 2 (the R-EPSI method) satisfies the input data requirements of stage 3 (our Marchenko equation-based method): the Green's function is wavelet-free, its dipole sources and monopole receivers are co-located at the same datum, and the wavefield is properly scaled. We use the effective wavelet output by R-EPSI such that the data can be compared with the other stages of the processed workflow. Alternatively, for instance, a Ricker wavelet that fits the frequency spectrum or bandwidth of the data can be used.

We were provided with a rough estimate of the top of the target area based on knowledge of the regional geology and an initial stack. Subsequently, we identified a strong reflection event on the shot gathers, which allows us to determine the traveltime curve t_2 separating overburden and target. This event was hand-picked and auto-tracked on each gather, whereby smoothing was applied to remove any remaining jumps in the traveltime curve between shots. If a velocity model were available then this curve could also be modeled instead of being picked. Finally, the curve was shifted up to exclude the event itself from the truncation and to avoid cross-cutting into a target primary. Figure 8b shows the effect of using the t_2 curve to mute a shot gather in the middle of our 2D line. We also define the time horizon ϵ, which is simply a hyperbola with an apex at half the R-EPSI wavelet length with a moveout approximately corresponding to the first event in the data.

Combining R_0 with the mutes the series in equation 2 can be evaluated to obtain the v^+. The value of and change in the L_2 norm of the individual terms, $||\Omega_k[v_e^+]||_2$, is used to determine convergence. Provided the Green's function derived by R-EPSI was scaled correctly, this value should decrease continuously. At some point, it will become negligible and the corresponding term unlikely to be still relevant. In addition, the total required number of terms or iterations will depend on the geological setting (Elison et al., 2020). For example, Staring et al. (2018) only use two iterations to predict internal multiples for a target area in deep water Brazil. In contrast, 20 iterations were needed in our case as Figure 7 shows. This (perhaps slow) convergence is supported by a well-log based synthetic test and other synthetic studies from the Middle East (Elison et al., 2020). Since the subsurface of the Arabian Gulf generates highly complex interference patterns, convergence is of the utmost importance to avoid missing important amplitude and phase updates and thus to ensure the internal multiple predictions match the data.

Once the operator v^+ was obtained, we evaluate the right hand side of equation 3 to obtain the target response without source- and receiver-side overburden-borne internal multiples. Next, those multiples are isolated according to equation 4 thereby providing us with the prediction (similar to Figure 8d). Clearly, we observe a complex interference pattern, instead of a multitude of distinct events. This matches the pattern in the data very well, thus, only a very conservative or mild adaptive filter could be used to subtract the predictions from the muted reflection response $\Theta_{t_2}^{\infty} R_0$. The parameterization used a filter length of just five samples, windows of 100 samples by 25 traces, and L_2 minimization. Figure 8c shows the actual result after adaptive subtraction of the internal multiple predictions given this parameterization. A clear difference is visible when compared to Figure 8b. Some events that were not visible before have emerged and the continuity of the reflections has improved. By studying the FK spectra in Figure 9, the effect of the subtraction of the internal multiples is also noticeable. The FK spectrum after adaptive subtraction (Figure 9c) does look more balanced, indicating that perhaps this complex overburden generates mainly low-frequency internal multiples. Note that the internal multiple predictions seem to be even more concentrated in a shorter range of low frequencies when compared to the surface-related multiples (Figure 9d vs 5c).

Figure 10 contains the RTM images corresponding to Figures 8a, 8c, and 8d. The yellow boxes indicate the target, where a difference due to the removal of internal multiples generated by the overburden can be observed. Note that a difference is also seen below the area of interest. Internal multiples generated by the overburden itself are removed, however those generated within the target or by scattering between the target and the overburden remain in the data. Figures 11a-11c show a zoom of the target area to study some of its details. A significant reduction of internal multiple energy is visible in the yellow circles. In addition, conflicting seismic events have been resolved, leading to improved visibility of certain known horizons (the green circles). Furthermore, the images now have a better definition of faulty structures (the red circles). We studied the amplitude spectra of the RTM images before and after internal multiple removal (Figure 12). These spectra confirm that the removed interference pattern mainly had low frequencies, but that also some higher frequencies were attenuated (which are perhaps not easily observed in the RTM images). This observation is supported by detailed synthetic controlled experiments (Reinicke et al., 2020b).

As final remarks, we did test two different target boundaries, but the effect was rather insignificant suggesting that the separability issue is not too essential. We also tested a variety of adaptive filter parameter settings, such as window sizes and filter lengths, which had hardly any impact on the result; thus, we settled on the values considered the least invasive or mildest.

DISCUSSION

We presented a case study, which introduced a two-step approach to handle both surface-related and internal multiples in a shallow water environment with a complex overburden. The R-EPSI method not only offers a solution to the shallow water surface-related de-multiple problem, it also provides the correct Green's function R_0 required for the Marchenko IME approach.

The results presented in this paper are encouraging for both the surface-related multiple as well as for the interbed multiple attenuation. However, proper QC of the results on this complex shallow-water field data example is very difficult since we did not have well logs to which to tie our results. Part of ongoing work is to apply the proposed workflow on data where well-log information is readily available. In parallel, additional controlled experiments are being conducted to verify if and to what level the observations we gathered from this initial field data experiment are valid. This will help us to understand if the amount and complexity of the predicted multiple patterns are realistic, or if the observed dominant low-frequency content of the interbed multiples can be explained better. On the one hand, the amplitude- and phase-preserving characteristics of the approach are one of its main advantages. These characteristics allow the removal of the complex interference pattern of all overburden-generated internal multiples in a fully data-driven way and without (too much) dependence on any a priori information (for instance generation mechanism identification for multiple prediction). On the other hand, they also put stringent conditions on all (pre-)processing steps to maintain amplitude and phase validity. Therefore, it makes sense to revisit the processing flow used for the 2D field data exercise and discuss where more attention might be required for future projects.

First point of attention is the 3D-to-2D conversion, which could influence the correct amplitude balance between primaries and multiples. In the future work, other approaches can be used, which are expected to predict generally more reliable amplitudes, and produce exact results for horizontally layered 1.5D media (Wapenaar et al., 1992). Sensitivity tests were not part of the scope of work, but given the convergence of both the R-EPSI and Marchenko methods, and a relatively flat vertical effective velocity profile, the impact is believed to be minimal.

Another aspect that could affect the proposed amplitude- and phase-sensitive workflow is the quality of the up-down decomposition, which is not always straightforward for complex field-data situations as used in this paper. More synthetic and well-controlled field data exercises are recommended to test how sensitive this will be on the final de-multiple results. Also, instead of applying the up-down decomposition to the P and V_z data, alternative methods could be investigated. For instance, van Groenestijn and Ross (2011) discuss how the upgoing wavefield can be obtained by applying EPSI to the hydrophone-only data instead.

For this case study, a 2D subset of the data was selected. This subset was reasonably regular (not many gaps in the acquisition), had only limited cross-line offsets, and was bandpass filtered to avoid any aliasing effects. All this minimized the dependence of the interpolation/regularization on our de-multiple results. Further research is needed to determine how sensitive both the R-EPSI and Marchenko methods are to data regularization in case of realistic scenarios where sparse crossline spacing, irregular sampling, larger and more data gaps and higher frequencies will play a role of importance.

Finally, a few topics that were not addressed in the current paper are worth mentioning. These topics were discussed already in other articles or are being addressed in ongoing research projects. A known problem with the R-EPSI method is that the approximation of the Green's function sparsity with a L1-norm may lead to some scaling ambiguity between the estimated wavelet and the Green's function (Esser et al., 2015). Nevertheless, so far this problem was not encountered in other R-EPSI applications (Belonosov and van Borselen, 2017). For the purpose of this paper, we expect that the potential impact of this scaling ambiguity is limited since we saw that the Marchenko Neumann series expansion converged at nearly identical rate to the synthetic based on the well log from the region.

The proposed Marchenko approach only attenuates internal multiples generated while passing through the overburden and ignores the overburden-target reverberations. In many cases, this will be sufficient to sweep the target area of most internal multiples thereby leading to improved interpretation. It would therefore be worthwhile to compare the numerical efficacy of this method on a larger section of field data using the MDD step as in van der Neut and Wapenaar (2016) or by using an extended method proposed by Dukalski and de Vos (2020).

In this paper, the short-period internal multiples have not been considered nor have they been addressed by the here proposed method. Dukalski et al. (2019) and Elison et al. (2020) discuss a method to handle these short-period multiples in 1D and for media with flat overburdens, respectively. Part of ongoing work is on validating this method on synthetic data sets. The relatively horizontal layered character of the Arabian Gulf area makes the data example used in this paper also a good candidate for testing this method.

CONCLUSION

We demonstrated an amplitude- and phase-preserving seismic processing workflow for shallow water environments with a complex overburden on an OBC data set acquired in the Arabian Gulf. The workflow addresses the challenge of removing surface-related multiples without using any a priori information about the subsurface when the water bottom and shallow subsurface boundaries are not represented in the recorded data due to shallow water depth. In addition, the proposed workflow provides a solution for predicting interbed multiples at target level that have been generated by a complex, thin-layered overburden. Robust Estimation of Primaries using Sparse Inversion (R-EPSI) proved to be successful in simultaneously removing the surface-related multiples and predicting the source wavelet. The double de-reverberation Marchenko method is subsequently applied to the resulting surface-related primaries Green's functions to suppress the interbed multiples generated by the thin-layered overburden. The predictions contain a complex interference pattern as opposed to a set of separate events that are typically observed when dealing with interbed multiples for less complex, and generally deeper water, environments. Given the number of closely spaced internal multiple generators in the complex overburden of the Arabian Gulf and the interference of various dominant orders of these multiples, this predicted pattern is very complex, demonstrating that this problem can only be fully addressed with truly physics-based approaches such as the double de-reverberation Marchenko method.

In ongoing research we aim to work towards a better understanding of the robustness of this method. One issue to address is the current use of, although very conservative, adaptive filters to better match the predicted interbed multiples with the data. Theoretically, no adaptive filters are required. Sensitivity studies for the individual (pre-) processing steps will help to obtain further understanding of the proposed workflow when applied to data representing similar challenges as the Arabian Gulf data utilized in this paper. These studies should include controlled experiments on synthetic data from realistic complex models and the use of various QC tools, like well-log matching, to verify whether the obtained predictions on field data are accurate.

ACKNOWLEDGEMENTS

We would like to thank Matteo Ravasi, Brahim Abbad, Jeffrey Shragge and an anonymous reviewer for their comments that helped improving this manuscript. Most of this research was performed in the framework of the project 'Marchenko imaging and monitoring of geophysical reflection data', which is part of the Dutch Open Technology Programme with project number 13939 and is financially supported by NWO Domain Applied and Engineering Sciences. The research of K. Wapenaar has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 742703).

REFERENCES

Alá'i, R., D. J. Verschuur, J. Drummond, S. Morris, and G. Haughey, 2002, Shallow water multiple prediction and attenuation, case study on data from the Arabian Gulf, *in* SEG Technical Program Expanded Abstracts 2002, 2229–2232.

Amundsen, L., L. T. Ikelle, and L. E. Berg, 2001, Multidimensional signature deconvolution and free-surface multiple elimination of marine multicomponent ocean-bottom seismic data: Geophysics, **66**, 1594–1604.

Barnes, S. R., R. F. Hegge, R. van Borselen, and J. Owen, 2014, Shallow reverberation prediction methodology with SRME: Presented at the 76th EAGE Conference and Exhibition 2014.

Belonosov, M., R. Baardman, J. Yoo, and R. van Borselen, 2019, Practical approach to Robust EPSI in shallow water data from the Middle East: Presented at the 81st EAGE Conference and Exhibition 2019.

Belonosov, M. and R. van Borselen, 2017, A comparison study for robust estimation of primaries by sparse inversion, *in* SEG Technical Program Expanded Abstracts 2017, 4833–4837.

Berkhout, A., 1982, Seismic migration, imaging of acoustic wave energy by wave field extrapolation: A: theoretical aspects, 2nd edition: Elsevier.

Biersteker, J., 2001, Magic: Shell's surface multiple attenuation technique: SEG Technical Program Expanded Abstracts 2001, 1301–1304.

Brookes, D., 2011, Case studies in 3D interbed multiple attenuation: The Leading Edge, **30**, 914–918.

Czyczula-Rudjord, Z., K. de Vos, and P. Hoogerbrugge, 2015, Real data comparison of EPSI and Robust EPSI including the near offset reconstruction: Presented at the 77th EAGE Conference and Exhibition 2015.

Dukalski, M. S. and K. de Vos, 2017, Marchenko inversion in a strong scattering regime including surface-related multiples: Geophysical Journal International, **212**, 760–776.

———, 2020, A close formula for true-amplitude overburden generated interbed de-multiple: Presented at the 82nd EAGE Conference and Exhibition 2020.

Dukalski, M. S., E. Mariani, and K. de Vos, 2019, Handling short-period scattering using augmented Marchenko autofocusing: Geophysical Journal International, **216**, 2129–2133.

El-Emam, A., K. S. Al-Deen, A. Zarkhidze, and A. Walz, 2011, Advances in interbed multiples prediction and attenuation: Case study from onshore Kuwait, *in* SEG Technical Program Expanded Abstracts 2011, 3546–3550.

El-Emam, A., C. Kostov, and M. T. Hadidi, 2015, Report on the SEG/KOC joint workshop in Kuwait: Seismic multiples – Are they signal or noise?: The Leading Edge, **34**, 320–324.

Elison, P., M. S. Dukalski, K. de Vos, D. J. van Manen, and J. O. A. Robertsson, 2020, Data-driven control over short-period internal multiples in media with a horizontally layered overburden: Geophysical Journal International, **221**, 769–787.

Esser, E., T. Lin, R. Wang, and F. J. Herrmann, 2015, A lifted l1/l2 constraint for sparse blind deconvolution: Presented at the 77th EAGE Conference and Exhibition 2015.

Griffiths, M., J. Hembd, and H. Prigent, 2011, Applications of interbed multiple attenuation: The Leading Edge, **30**, 906–912.

Hargreaves, N., 2006, Surface multiple attenuation in shallow water and the construction of primaries from multiples, *in* SEG Technical Program Expanded Abstracts 2006, 2689–2693.

Hung, B., K. Yang, J. Zhou, and Q. L. Xia, 2010, Shallow water demultiple: ASEG Extended Abstracts, **2010**, no. 1, 1–4. (doi: 10.1081/22020586.2010.12041899).

Jahdhami, M. A., J. de Maag, A. Mueller, S. Narhari, O. O. Kolawole, V. Kidambi, B. Al-Qadeeri, and Q. Dashti, 2017, 3D surface-related and interbed multiples attenuation on 5D-interpolated OVT data from a single-sensor survey in Kuwait: SEG Technical Program Expanded Abstracts 2017, 4773–4776.

Jakubowicz, H., 1998, Wave equation prediction and removal of interbed multiples, *in* SEG Technical Program Expanded Abstracts 1998, 1527–1530.

Kryvohuz, M. and X. Campman, 2019, Prediction of signatures of airgun arrays using dual near-field hydrophones: Geophysical Prospecting, **67**, 1522–1546.

Lesnikov, V. and J. Owusu, 2011, Understanding the mechanism of interbed multiple generation using VSP data, *in* SEG Technical Program Expanded Abstracts 2011, 4258–4262.

Lin, T. T. and F. J. Herrmann, 2013, Robust estimation of primaries by sparse inversion via one-norm minimization: Geophysics, **78**, R133–R150.

Lokshtanov, D., 1999, Multiple suppression by data-consistent deconvolution: The Leading Edge, **18**, 115–119.

Pereira, R., D. Mondini, and D. Donno, 2018, Efficient 3D internal multiple attenuation in the Santos Basin: Presented at the 80th EAGE Conference and Exhibition 2018.

Pereira, R., M. Ramzy, P. Griscenco, B. Huard, H. Huang, L. Cypriano, and A. Khalil, 2019, Internal multiple attenuation for OBN data with overburden/target separation: SEG Technical Program Expanded Abstracts 2019, 4520–4524.

Ravasi, M., 2017, Rayleigh-Marchenko redatuming for target-oriented, true-amplitude imaging: Geophysics, **82**, no. 6, S439–S452.

Ravasi, M., I. Vasconcelos, A. Kritski, A. Curtis, C. A. d. C. Filho, and G. A. Meles, 2016, Target-oriented Marchenko imaging of a North Sea field: Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society, **205**, 99–104.

Reinicke, C., M. Dukalski, and K. Wapenaar, 2020a, Comparison of monotonicity challenges encountered by the inverse scattering series and the marchenko de-multiple method for elastic waves: Geophysics, **85**, 1–66.

Reinicke, C., M. S. Dukalski, and C. Wapenaar, 2019, Tackling different velocity borne challenges of the elastodynamic Marchenko method: Presented at the 81st EAGE Conference and Exhibition 2019.

Reinicke, C., M. S. Dukalski, and K. Wapenaar, 2020b, Effective media theory consistent multiple elimination with the marchenko equation based methods: Presented at the 82nd EAGE Conference and Exhibition 2020.

Retailleau, M., H. Farran, P. Hugonnet, T. W. Hong, J. Shorter, and B. Michels, 2014, Unveiling the geology below the multiples by filtering planar events: Examples from onshore Oman: Presented at the Third EAGE Exploration Workshop.

Rønholt, G., Ø. Korsmo, S. Brown, A. Valenciano, D. Whitmore, N. Chemingui, S. Brandsberg, V. Dirks, J.-E. Lie, et al., 2014, High-fidelity complete wavefield velocity model building and imaging in shallow water environments – a North Sea case study: First Break, **32**, 127–131.

Savels, T., K. de Vos, and J. de Maag, 2011, Surface-multiple attenuation through sparse inversion: results for complex synthetics and real data: First Break, **29**, 55–64.

Singh, S., R. Snieder, J. van der Neut, J. Thorbecke, E. Slob, and K. Wapenaar, 2016, Accounting for free-surface multiples in Marchenko imaging: Geophysics, **82**, no. 1, R19–R30.

Slob, E., K. Wapenaar, F. Broggini, and R. Snieder, 2014, Seismic reflector imaging using internal multiples with Marchenko-type equations: Geophysics, **79**, no. 2, S63–S76.

Soubaras, R., 1996, Ocean bottom hydrophone and geophone processing, *in* SEG Technical Program Expanded Abstracts 1996, 24–27.

Staring, M., R. Pereira, H. Douma, J. van der Neut, and K. Wapenaar, 2018, Source-receiver Marchenko redatuming on field data using an adaptive double-focusing method: Geophysics, **83**, no. 6, S579–S590.

van den Berg, E. and M. P. Friedlander, 2008, Probing the pareto frontier for basis pursuit solutions: SIAM Journal on Scientific Computing, **31**, 890–912.

———, 2019, SPGL1: A solver for large-scale sparse reconstruction. (https://friedlander.io/spgl1).

van der Neut, J., I. Vasconcelos, and K. Wapenaar, 2015, On green's function retrieval by iterative substitution of the coupled Marchenko equations: Geophysical Journal International, **203**, 792–813.

van der Neut, J. and K. Wapenaar, 2016, Adaptive overburden elimination with the multidimensional Marchenko equation: Geophysics, **81**, no. 5, T265–T284.

van Groenestijn, G. J. A. and W. Ross, 2011, Primary estimation on obc data by sparse inversion: SEG Technical Program Expanded Abstracts 2011, 3531–3535.

van Groenestijn, G. J. A. and D. J. Verschuur, 2009, Estimating primaries by sparse inversion and application to near-offset data reconstruction: Geophysics, **74**, no. 3, A23–A28.

———, 2009, Estimation of primaries and near-offset reconstruction by sparse inversion: Marine data applications: Geophysics, **74**, no. 6, R119–R128.

Verschuur, D. J., 2013, Seismic multiple removal techniques: past, present and future: EAGE publications.

Verschuur, D. J., A. Berkhout, and C. Wapenaar, 1992, Adaptive surface-related multiple elimination: Geophysics, **57**, 1166–1177.

Wapenaar, C., D. J. Verschuur, and P. Herrmann, 1992, Amplitude preprocessing of single and multicomponent seismic data: Geophysics, **57**, 1178–1188.

Wapenaar, K., J. Thorbecke, J. van Der Neut, F. Broggini, E. Slob, and R. Snieder, 2014, Marchenko imaging: Geophysics, **79**, no. 3, WA39–WA57.

Ware, J. A. and K. Aki, 1969, Continuous and discrete inverse-scattering problems in a stratified elastic medium. I. Plane waves at normal incidence: The journal of the Acoustical Society of America, **45**, 911–921.

Weglein, A. B., F. A. Gasparotto, P. M. Carvalho, and R. H. Stolt, 1997, An inverse-scattering series method for attenuating multiples in seismic reflection data: Geophysics, **62**, 1975–1989.

Yanez, M., C. Tyagi, M. Steiger-Jarvis, J. Bailey, M. Shadrina, and S. Dingwall, 2018, Addressing the challenges of heritage surface seismic data in a complex shallow water environment for improved overburden imaging and reservoir characterization, *in* SEG Technical Program Expanded Abstracts 2018, 1123–1127.

Ypma, F. and D. Verschuur, 2013, Estimating primaries by sparse inversion, a generalized approach: Geophysical Prospecting, **61**, 94–108.

LIST OF FIGURES

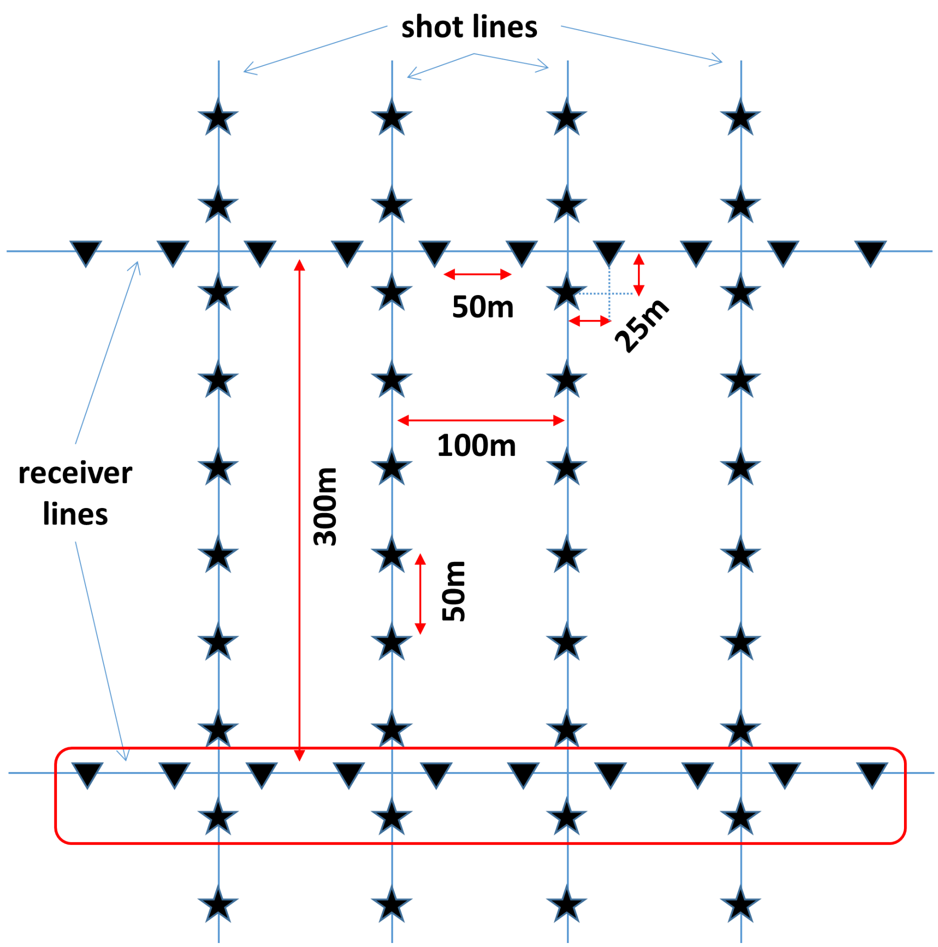


Figure 1: OBC acquisition geometry for the Arabian Gulf data. Receivers are located at the sea-bottom, while sources reside just below the water surface. The oval indicates the 2D line selected out of this 3D survey.

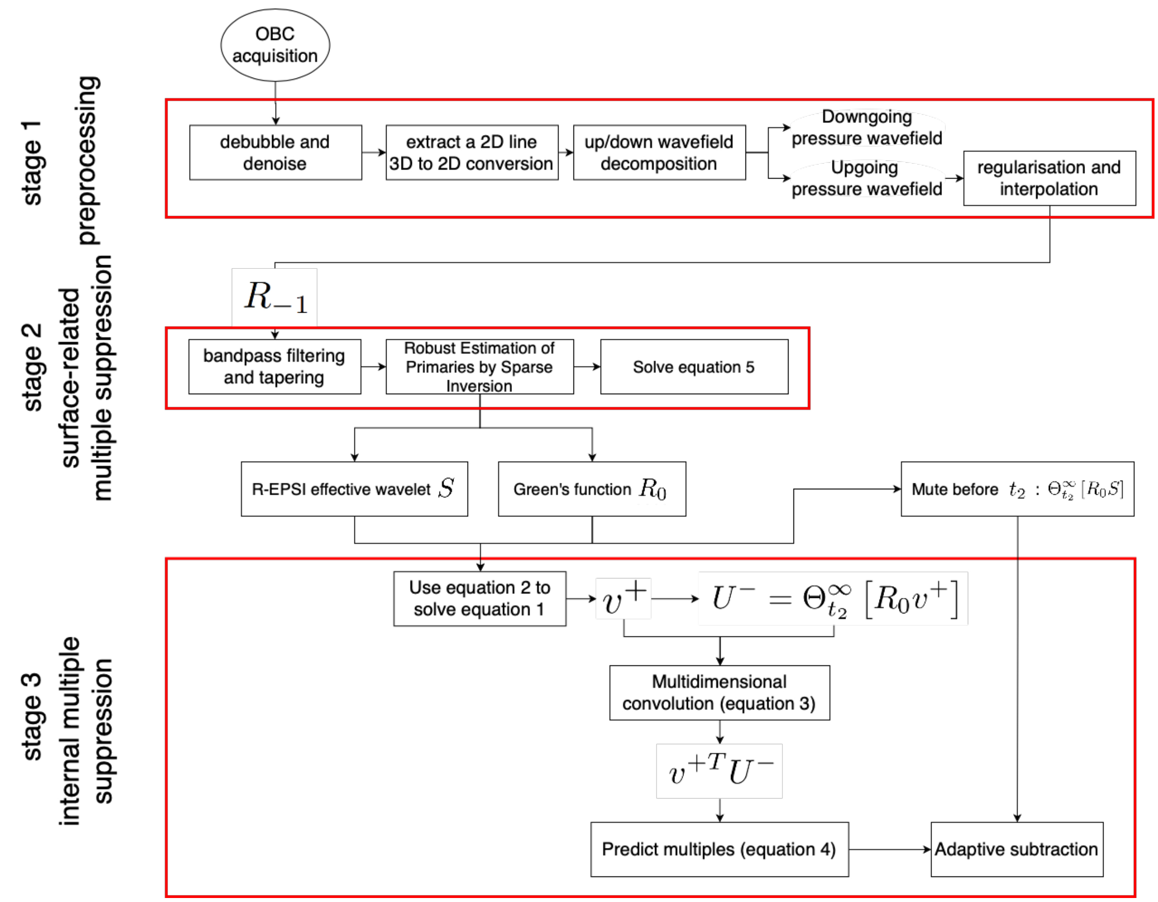
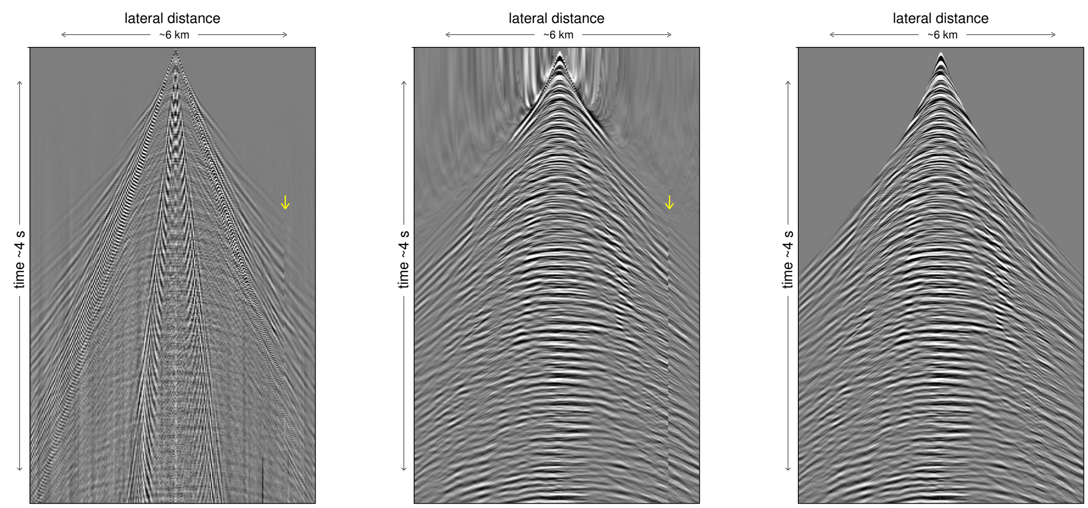
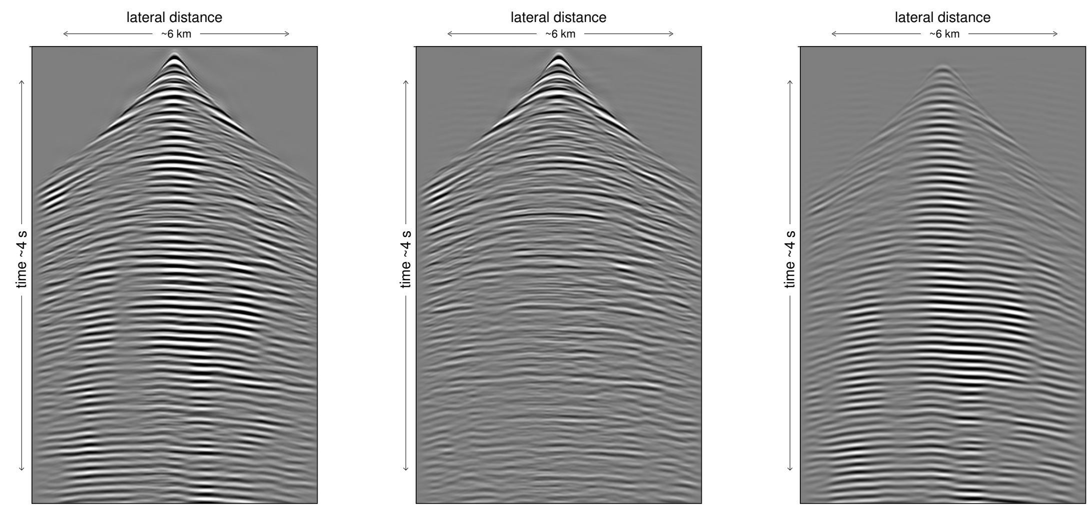


Figure 2: Workflow showing the different stages involved in the processing of our dataset. Stage 1 deals with the preprocessing, stage 2 performs surface-related multiple suppression and stage 3 uses the proposed Marchenko method for internal multiple suppression.



(a) (b) (c)

Figure 3: Shot gather of (a) raw hydrophone data (with a $t^{1.5}$ gain), (b) upgoing pressure, (c) result after regularization of the wavefield in (b); The yellow arrows point to an aquisition gap that was resolved by regularization.



(a) (b) (c)

Figure 4: Fixed spread common shot gather (a) before R-EPSI, (b) after R-EPSI (i.e. R_0), (c) their difference.

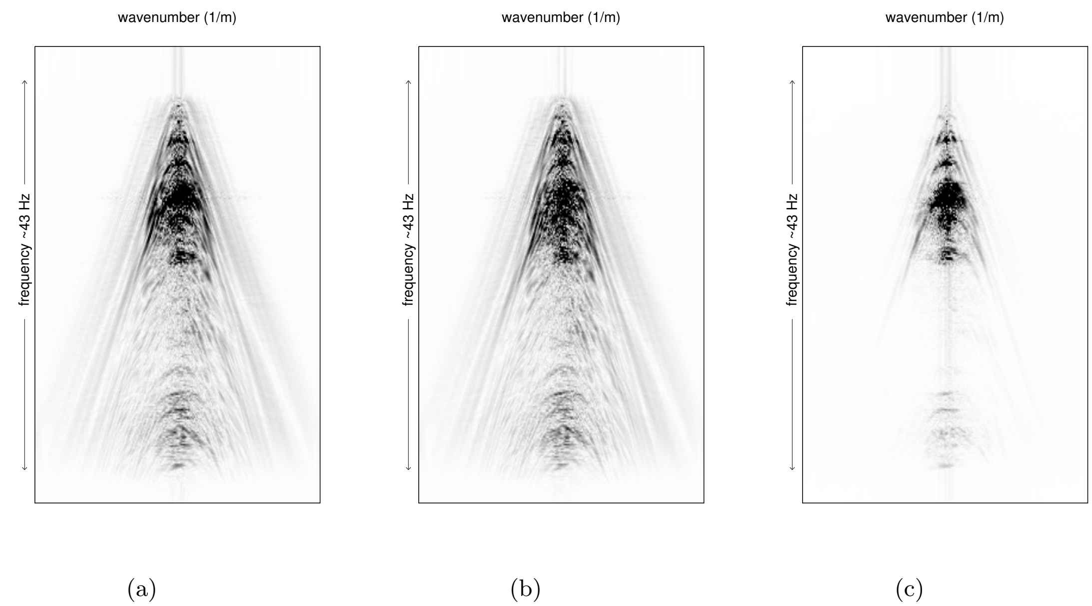
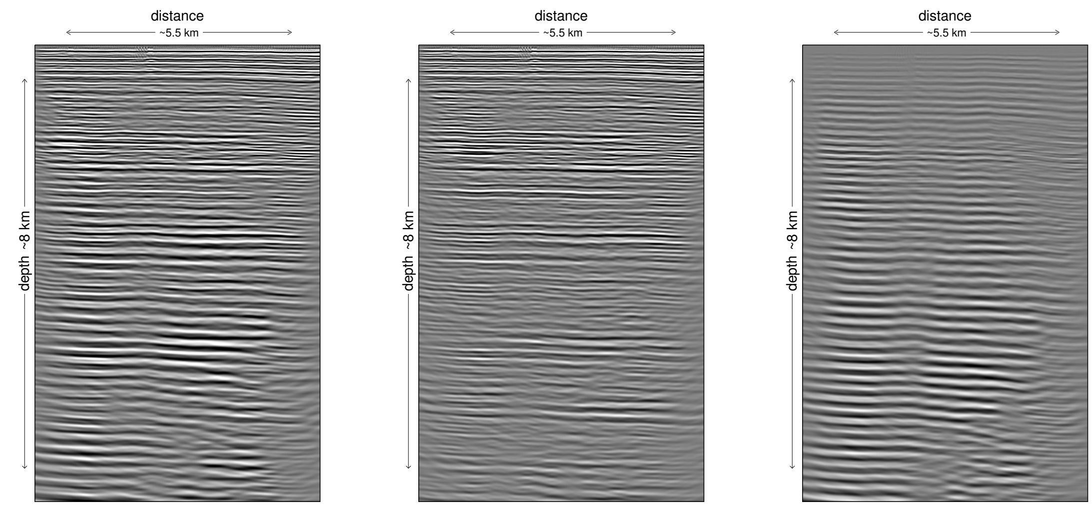


Figure 5: FK spectra for each gather in Figure 4 (a) before R-EPSI, (b) after R-EPSI, (c) their difference.



(a) (b) (c)

Figure 6: RTM of data (a) before R-EPSI, (b) after R-EPSI, (c) their difference.

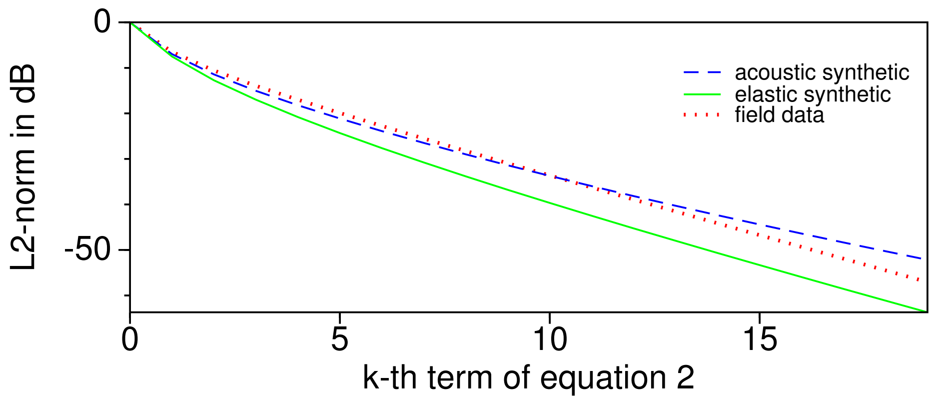


Figure 7: L_2 norm of the individual terms in the series of equation 2. Convergence curve for acoustic synthetic in dashed blue, for elastic synthetic in continuous green and for the Arabian Gulf field data in dotted red.

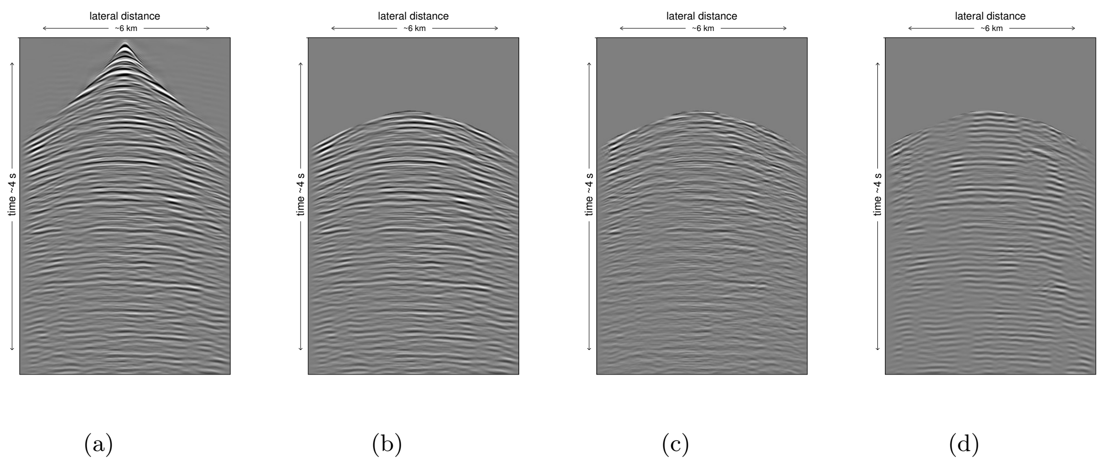
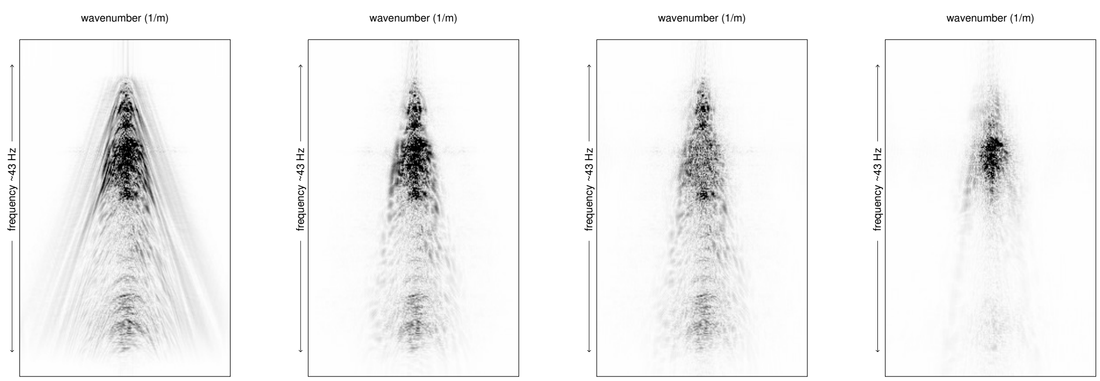
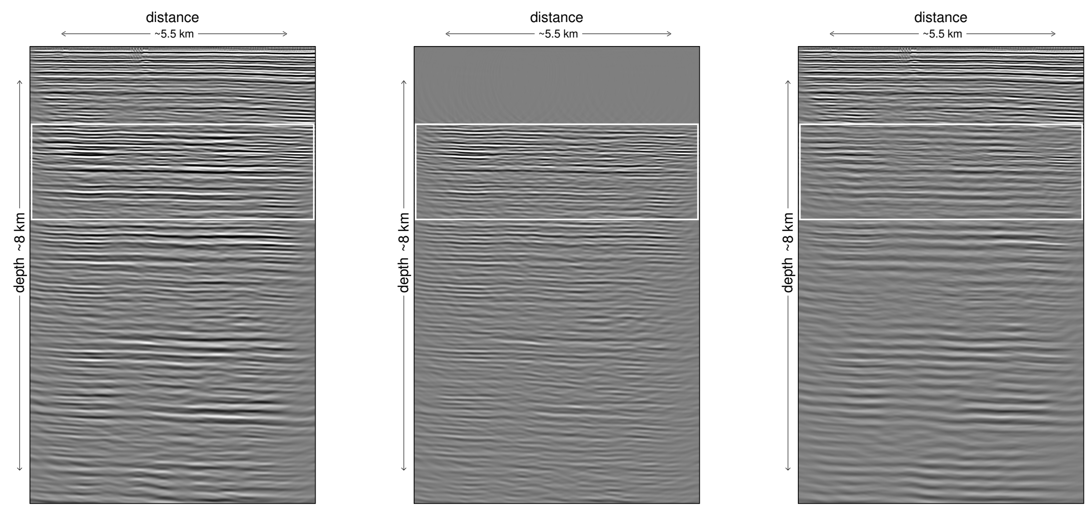


Figure 8: Shot gather of data (a) after R-EPSI (same as Figure 4b), (b) after the t_2 mute, (c) after Marchenko, (d) the difference between (b) and (c).



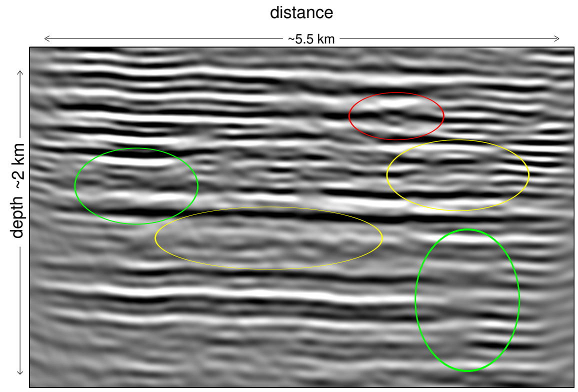
(a) (b) (c) (d)

Figure 9: FK spectra corresponding to Figure 8 of data (a) after R-EPSI, (b) after the t_2 mute, (c) after Marchenko, (d) their difference.

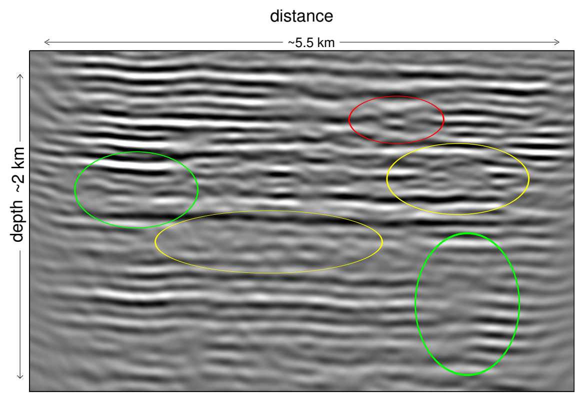


(a) (b) (c)

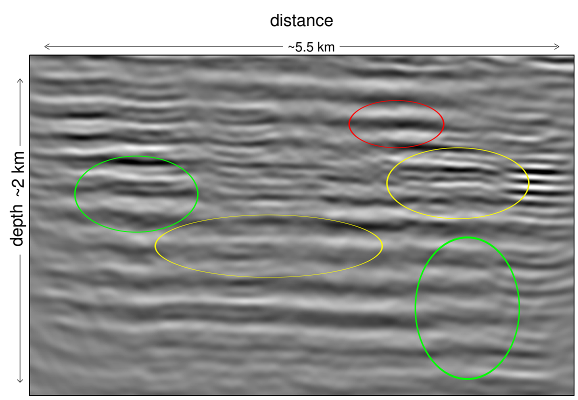
Figure 10: RTM of data (a) after R-EPSI (same as Figure 6b), (b) after Marchenko, (c) their difference.



(a)



(b)



(c)

Figure 11: Zoom of RTM in Figure 10 (a) after R-EPSI, (b) after Marchenko, (c) their difference.

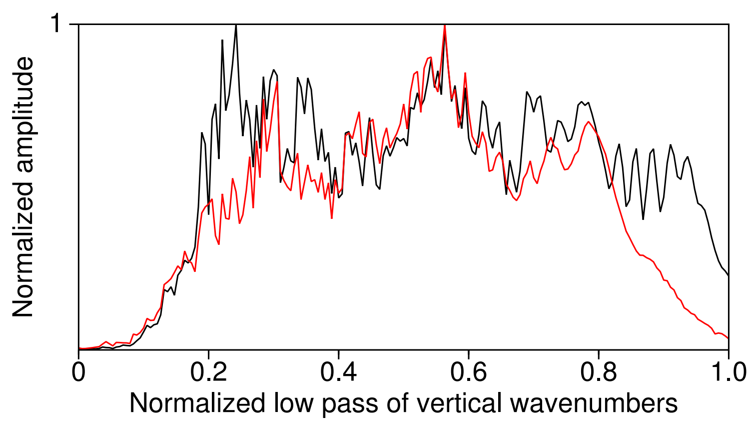


Figure 12: Spectral analysis of vertical wavenumbers for the zoomed area of RTMs of data before and after Marchenko in red and black, respectively.