Pre-processing methodology to apply the Marchenko method in conventional marine datasets

L. D. L. Vergne., Diego F. Barrera P., Victor Kohene., Marcelo Santana

SENAI/CIMATEC Supercomputing Center, Salvador, Bahia, Brazil

Abstract

Recently the application of the Marchenko method in marine field data has been of great interest in the oil industry. This method allows us to calculate the up- and downgoing Green's functions at a focal point in depth without artifacts coming from the overburden, and build images with high resolution without the influence of internal multiples. However, to run the Marchenko scheme correctly in field data it is necessary to fulfill the theoretical premises that this method requires in a pre-processing stage. The main premises are elevation statics, regularization, deconvolution, reciprocity, and removal of free-surface multiples. In this work, we present and discuss the steps elevation statics, interpolation, and reciprocity in order to successfully apply the Marchenko scheme in a marine seismic datasets.

Introduction

The Marchenko method is a powerful geophysical tool that allows us to build high-resolution images in depth and remove internal multiples from a seismic dataset (Wapenaar et al., 2014; Zhang and Slob, 2020). The quality of the results depends on the input field data pre-processing. Authors as Peng et al. (2019); Staring and Wapenaar (2019); and others, discuss the sensitivity of the Marchenko method due to deconvolution, sampling, acquisition geometry, and surface related multiples.

Brackenhoff et al. (2019) show in numerical examples the importance of the dense and regular sampling of sources and receivers in the Marchenko method. This is because lack of sampling may cause the method not converge. On the other hand, Barens et al. (1995) and Jia et al. (2018) discuss the possibility to transform end-on datasets to split-spread seismic arrays using other domains on the data, e.g., receiver gathers. This way of changing the partial array to a full coverage array is called a reciprocity method.

Free surface multiples in marine data are always present. These kind of multiples in the conventional Marchenko approach are a problem to guarantee the convergence of the method. It is recommended to apply Surface Related Multiple Elimination (SRME) (Verschuur et al., 1992) to remove the free-surface multiples of the dataset. New developments with the Marchenko equations could take into account the free-surface multiples in the input data (Singh et al., 2017), but this work does not explore this method.

We will show three pre-processing steps to ensure that the Marchenko method works in field data. Namely, elevation statics, interpolation, and reciprocity. Here, we will show the theory behind these methods, and validate this in numerical examples.

Methodology

To apply Marchenko iterative scheme to real seismic data Jia et al. (2018) propose a pre-processing flow of marine seismic data, where the most important steps are: (1) apply Surface Related Multiple Elimination (SRME) to the seismic data for the attenuation of surface multiples, (2) estimate the best signature of the source to deconvolve the data, (3) apply reciprocity to the data to transform it into a split-spread configuration, and (4) calibrate the amplitudes of the data using the scale factor estimated from comparisons with synthetic shots. On the other hand, Singh et al. (2017) propose a new scheme in order to eliminate step (1). Brackenhoff et al. (2019) show that after applying step (3), the traces that correspond to the near offsets are null, being necessary an interpolation. This is the case that the target in depth is shallow. Otherwise, is not necessary to do the near offset interpolation.

Elevation statics

Elevation static is a method that allows us to relocate sources and receivers at the same level. In marine acquisition data, the calculation is done by applying only one shift in the seismic trace to the reference datum. This correction is usually necessary when the sources or receivers are below the surface or datum, the latter being equal to sea level in probably all marine cases.

Two acquisition arrangements for typical marine data can be seen in Figure 2, where sources and receivers are a few meters below sea level. Elevation statics T_A for sources located in A is given by:

$$T_A = \left(\frac{d_A + E_d}{V_w} \right). \tag{1}$$

With d_A being the depth of the source below sea level, V_w the velocity of the water, and E_d is the elevation of the reference datum (usually zero). In equation 1, the sign of E_d is positive above sea level (convention used for elevation), and d_A is positive for depths below sea level. Since the reference datum is equal to sea level, the elevation statics is positive, which means that the times

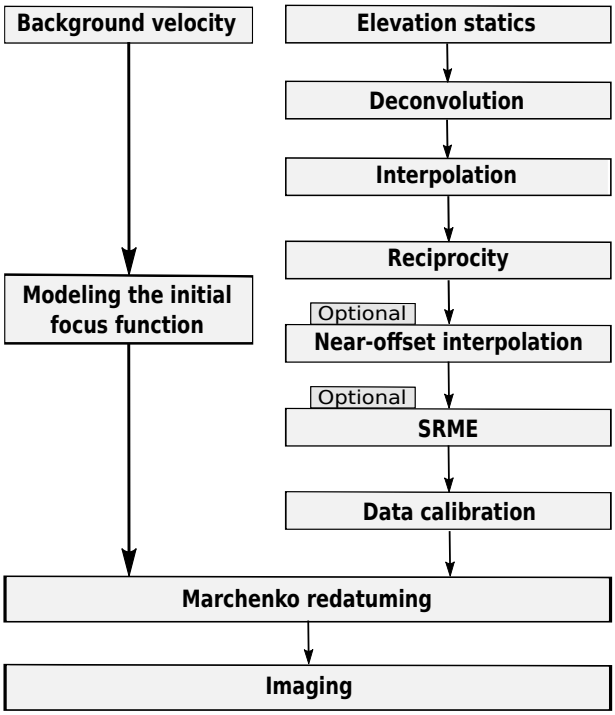

Figure 1: Flow of the application of the Marchenko iterative scheme (Modified from Jia et al. (2018)).

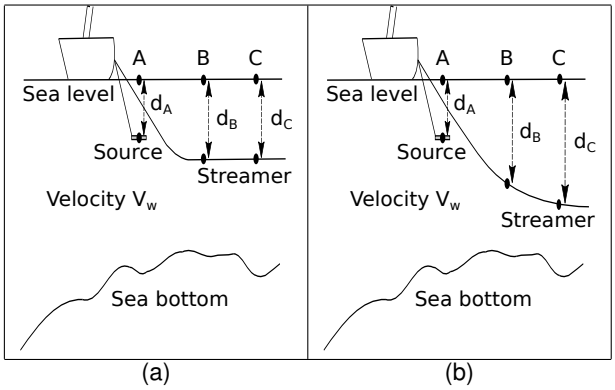

Figure 2: a) Conventional and b) broadband marine seismic acquisition.

of seismic reflections from the datum are longer than the observed.

Interpolation

The interpolation methods proposed in the literature seek to overcome spatial sampling obstacles in the seismic data. This problem can affect lateral resolution making a severe constrain in the Marchenko application.

A method to perform seismic trace interpolation that handles spatially aliased events was proposed by Spitz (1991). This method states that linear events constructed by equally spaced traces can be interpolated exactly, regardless of the original spatial interval (Porsani, 1999). It works in the frequency-space $(f-x)$ domain and was originally developed for 2D interpolation.

Porsani (1999) developed an algorithm based on Spitz's method. In this new methodology, applied in this work, a half-step prediction filter is used. It works on even data components predicting odd ones and vice-versa. By using this filter only one system of linear equations must be solved, making the algorithm faster and easier for implementation.

According to Porsani (1999), the Spitz's method may be implemented following the steps: (1) Use FFT algorithm, with $2nft$ points, (nft = number of points used to perform the Fourier transform) to transform the input data section from the $t-x$ domain to the $f-x$ domain; (2) Compute the one-step prediction filter for each frequency in the bandwidth of the data; (3) Squeeze the frequency axis preserving the odd data components; (4) Solve the equation presented by Porsani (1999) for the missing traces; (5) Back to the $t-x$ domain using FFT with nft data points.

The main differences of the Porsani (1999) method are in steps (2) and (4). In step (2) a different system of normal equations must be solved to obtain the half-step prediction filter. In the step (4) the interpolated missing traces are estimated directly as output of the convolution of the half-step prediction filter obtained at frequency f and the input data components at frequency $f/2$.

Reciprocity

The application of reciprocity in marine data seeks to overcome limitations related to the acquisition geometry. According to Wapenaar et al. (2014); Jia et al. (2018) the Marchenko method requires a split-spread geometry, which we understand to be the best geometry covering the most stationary points corresponding to subsurface events to be dealt with by the Marchenko scheme. Given a standard marine acquisition, we can construct the split-spread from a end-on or end-off shot $\mathbf{D}^i$ represented by:

$$\mathbf{D}^i = \begin{bmatrix} \mathrm{d}^i_1(1) & \cdots & \mathrm{d}^i_{nr}(1) \\ \vdots & \ddots & \vdots \\ \mathrm{d}^i_1(nt) & \cdots & \mathrm{d}^i_{nr}(nt) \end{bmatrix} \tag{2}$$

where $\mathbf{D}$ is the matrix that represents a single shot, i the index that represents the number of shots, $\mathrm{d}^i_1(1)$ the first sample of trace 1 of shot i, nt the total number of samples, and nr the total number of traces.

A set of seismic data can be organized into different common gathers if they are arranged in a diagonal configuration. It is then possible, from several shots in the end-on configuration, such as the shot represented by the equation 2, to generate the matrix with shifted shots $\overline{\mathbf{D}}$, and its transpose $\overline{\mathbf{D}}^T$ respectively:

$$\overline{\mathbf{D}} = \begin{bmatrix} \mathbf{D}^1_1 & 0 & \cdots & 0^{ns} \\ \vdots & \mathbf{D}^2_1 & \ddots & \vdots \\ \mathbf{D}^1_{nr} & \vdots & \ddots & 0 \\ 0 & \mathbf{D}^2_{nr} & \vdots & \mathbf{D}^{ns}_1 \\ \vdots & \vdots & \ddots & \vdots \\ 0_{ns+nr-1} & 0 & \cdots & \mathbf{D}^{ns}_{nr} \end{bmatrix} \tag{3}$$

$$\overline{\mathbf{D}}^T = \begin{bmatrix} \mathbf{D}_1^1 & \cdots & \mathbf{D}_{nr}^1 & 0 & \cdots & 0_{ns+nr-1} \\ 0 & \mathbf{D}_1^2 & \cdots & \mathbf{D}_{nr}^2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \cdots & \ddots & \vdots \\ 0^{ns} & \cdots & 0 & \mathbf{D}_1^{ns} & \cdots & \mathbf{D}_{nr}^{ns} \end{bmatrix} el \quad (4)$$

where $\mathbf{D}_1^1$ represents the first trace of shot 1, $\mathbf{D}_{nr}^1$ the last trace of shot 1, ns the total number of shots, and nr the total number of traces.

The data in split-spread geometry can be calculated from the sum of $\overline{\mathbf{D}}$ and $\overline{\mathbf{D}}^T$:

$$\overline{\mathbf{D}}_f = \mathbf{B}^\mathbf{T}\overline{\mathbf{D}} + \overline{\mathbf{D}}^\mathbf{T}\mathbf{B} \quad (5)$$

Since the dimension of $\overline{\mathbf{D}}$ is $(nx \times ns)$, where $nx = (ns + nr - 1)$, and $\overline{\mathbf{D}}^T$ is $(ns \times nx)$, it becomes necessary to use auxiliary matrices filled with zeros of size $(nx \times nx)$ to perform the sum (given by the matrix $\mathbf{B}$). As a result we have the resulting split-spread $\overline{\mathbf{D}}_f$, an array $(nx \times nx)$:

$$\overline{\mathbf{D}}_f = \begin{bmatrix} 0 & \cdots & \mathbf{D}_1^1 & 0 & \cdots & 0^{nx} \\ \vdots & \ddots & \vdots & \mathbf{D}_1^2 & \ddots & \vdots \\ \mathbf{D}_1^1 & \cdots & \mathbf{D}_{nr}^1 & \vdots & \cdots & 0 \\ 0 & \mathbf{D}_1^2 & \cdots & \mathbf{D}_{nr}^2 & \cdots & \mathbf{D}_1^{ns} \\ \vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\ 0_{nx} & \cdots & 0 & \mathbf{D}_1^{ns} & \cdots & \mathbf{D}_{nr}^{ns} \end{bmatrix} \quad (6)$$

In equation 6 we can see that there is a zeroed region in the final result, and as it was said, zeros are not related to any geological configuration. Due to this problem, it is necessary to cut this information, in order to eliminate this empty space generated after the sum of the auxiliary matrices filled with zeros.

Thus, the number of shots resulting in the split-spread configuration is equal to the number of end-on shots used as input for the application of reciprocity, since the matrix with the empty space of the equation 6 is square due to the performed manipulations.

Numerical examples

In this section, we will show results related to the pre-processing methods explained above. This is in order to discuss in a simple model the effects that Marchenko could have if is not taking into account the theoretical premises that the method requires. The idea is to show the numerical examples in a didactic way, allowing an easy reproducibility for the new users of the Marchenko method.

Elevation statics

In this numerical example we will observe the effects of elevation statics using synthetic data generated from the model in Figure 3, a four-layer model with velocities of 1800, 2300, 2000 and $2500m/s$, from first to last layer. Also, was considered variation in the density field with values 1000, 3000, 1100 and $4000g/cm^3$ with the same geometry that velocity field.

The generated synthetic data has 901 shots, 901 receivers for each shot, and 1024 samples each trace. The data was modeled over a horizontal window of the original model covering -2250 to $2250m$, with spacing between sources and between receivers of 5 meters. This model will be used for the tests that will follow.

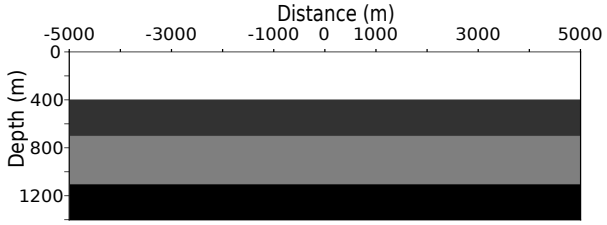

Figure 3: P-wave velocity model in plane-parallel layers.

To guide the numerical examples, we located the sources on the suface $(z = 0m)$, while varying the receivers on depth. With that configuration we can obtain synthetic data sets on different datums, which will be useful when validating the results of the Marchenko scheme. In order to calculate the total Green's functions with Marchenko, we located the focal point with a source at $900m$ in depth and receivers over the surface. These numerical examples were generated with Thorbecke's open source package (Thorbecke et al., 2017).

We build a reference result modeling the exact Green's function with the seismic array at the same surface (Figure 4a). When the sources are on the surface, and receivers are at $z = 50m$ in depth the total Green's function retrieved with Marchenko have all of the events with a vertical shift upwards (Figure 4b) when is compared with Figure 4a. This was expected, regarding that one of the premises in the Marchenko equations is that all of the integrals are applied to the same surface for both source and receivers (Wapenaar et al., 2014). Then, in order to move the receivers to the same datum as the shots we use the equation 1. After elevation statics we retrieve the Green's function with Marchenko and after eight iterations we obtain the result in Figure 4c, that is similar kinematically to all events in Figure 4a.

For the case of the marine acquisition of Figure 2b when sources are located on the surface $(z = 0m)$ and receivers are positioned in depth exponentially, we will have results similar to tests related to Figure 2a. In this case equation 1 could be applied trace by trace at each shot of the seismic array.

We can confirm these results by analyzing the central traces of the Figure 4. Then, in Figure 5 we can see that after application of elevation statics the events are located in the original position, when compared with the exact model. The loss of amplitude in Figures 4b and c are in detail in Figure 5 corresponds to normal effects related to the Marchenko scheme.

Interpolation

In order to satisfy the prerequisite of the Marchenko method, related to dense sampling of equally spaced sources and receivers, the interpolation method was applied in a synthetic data to reconstruct its zeroed information, simulating a process that often happens in real marine data acquisitions.

The first step to proceed with the interpolation was to generate the data that simulates a standard marine

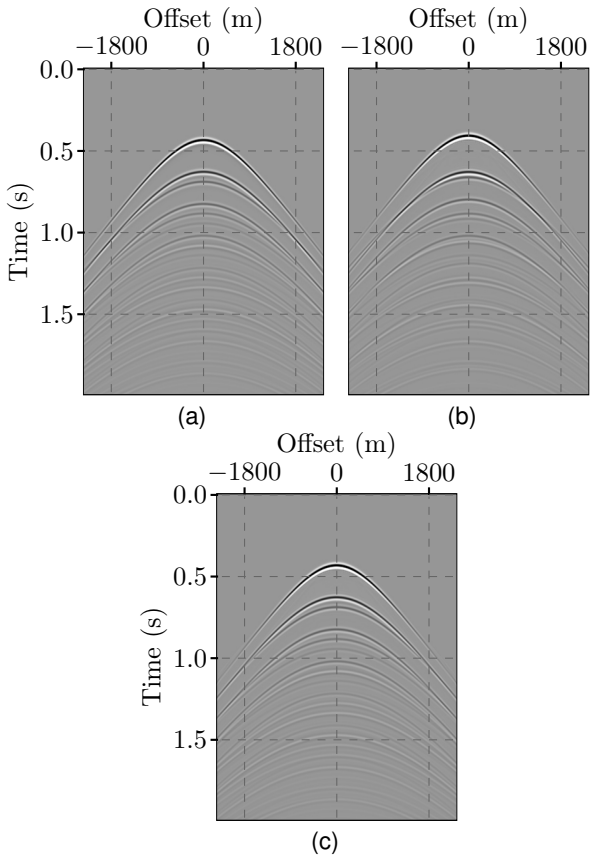

Figure 4: Total Green's functions with source at $900m$ at depth and receivers on surface (Figure 3) considering different levels of source an receivers. (a) is the exact model, (b) is retrieved considering sources on surface and receivers at $50m$ at depth, and (c) after elevation statics.

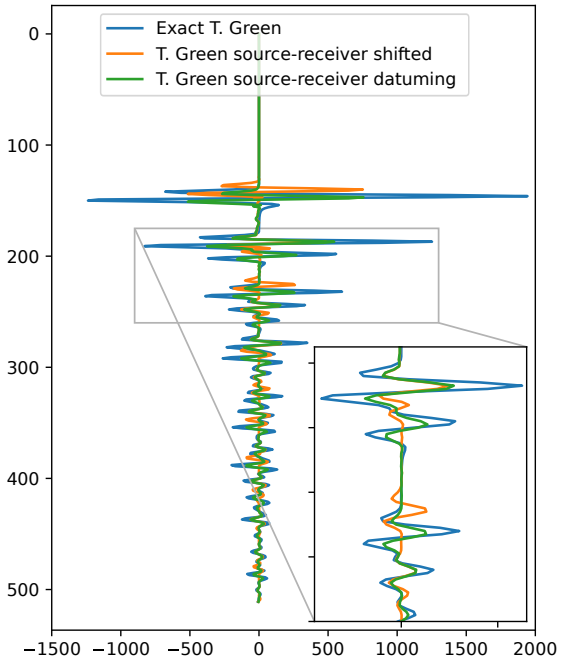

Figure 5: Central traces of results in Figure 4.

acquisition. For the development of the interpolation method, we will zero two shots (Figure 8a) that are positioned between two other neighboring shots (similar to actual acquisitions due to the distance between sources being different from that of receivers). The objective in this case is to reconstruct all zeroed traces belonging to the shots that were fully zeroed.

To model the end-off data used in the Figure 8a, we will use the velocity model in Figure 6 with velocity variation from the first layer up to the last layer down as $2000m/s$, $2300m/s$, $2600m/s$, and $2000m/s$. Also, was considered variation in the density field with values 1000, 5000, 1000 and $5000g/cm^3$ with the same geometry that velocity field. The fixed-spread synthetic data was first generated with 201 shots, 501 receivers for each shot, and 751 samples each trace (Figure 7a). The shots cover the area from $-1000m$ to $1000m$, and the receivers from $-2500m$ to $2500m$. The spacing between sources and between receivers is 10 meters. This dataset will be useful to visualize the result of the interpolation.

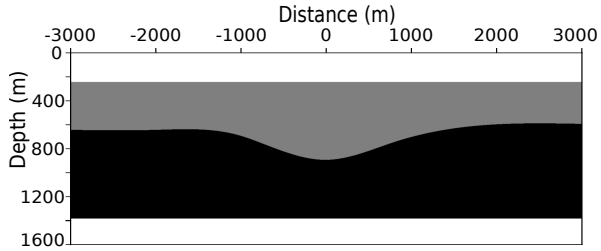

Figure 6: Model of P-wave velocities in a syncline.

Even though we could directly model the shots with an end-off geometry, remove some shots, and then recover them with the interpolation method, we prefer to model a fully covered fixed-spread set of shots (Figure 7a), and then window it to create two new sub-datasets (Figure 7b and c). The reason for creating this full dataset will be better understood in the Reciprocity session.

The first sub-dataset is of split-spread geometry, with 161 receivers, and cover a distance of $-1800m$ to $1800m$. One of its shots is shown in Figure 7b.

The second sub-dataset is obtained by windowing the left side of all shots of the previous sub-dataset, resulting in another sub-dataset with an end-off geometry. One of its shots is shown in Figure 7c. This sub-dataset has 201 shots with 81 receivers each and cover a horizontal area of $-1800m$ to $1000m$. From this sub-dataset, for every 4 shots, we zero out the pair between them, as shown by Figure 8a. We can now test the efficiency of the interpolation method on recovering the missing shots.

The second step is to organize the data in common receiver gathers (Figure 8b). At this stage, it is possible to observe that the non-existent traces are now distributed among the original traces as null traces. In this organization, it becomes easier to apply an interpolation method that reconstructs this zeroed information. In addition, it is necessary to eliminate noise present on the surface of the data that causes errors in the result of the interpolation, such as the direct wave.

In the implementation of the interpolation algorithm used

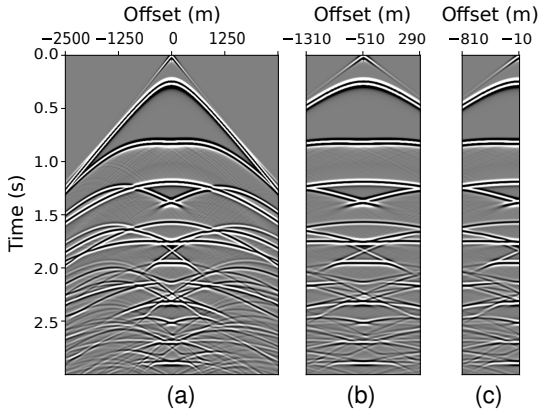

Figure 7: Shot 100 modeled from the model in Figure 6. In (a) we have the first set of shots with 501 receivers for each shot, where sources are limited from $-1000m$ to $1000m$ and receivers from $-2500m$ to $2500m$. In (b) we have the shots in split-spread geometry obtained from the shots in (a), where the range of shots remains the same and the 161 receivers cover the area from $-1800m$ to $1800m$. In (c) we have the shots in the end-off geometry. The 81 receivers are now located from $-1800m$ to $1000m$.

here, it was possible to generate two new traces between two original ones. It should be noted that the two new traces generated in the common receiver domain are corresponding to two neighboring shots that were zeroed in the common shot domain.

After the application of the interpolation is finished, we can see that the result in the common receiver gathers is satisfactory (Figure 8c), by analyzing the continuity of events in the seismogram. From this new interpolated data, we can return to the original common shot gather and proceed with reciprocity.

Reciprocity

The last step shown in this work is directly related to the application of reciprocity. Figure 8c, as said in the previous section, represents the common receiver matrix. We can say that the transpose of the common shot matrix is analogous to the common receiver matrix (Equations 3 and 4). As already described in the methodology section, obtaining the shot with split-spread geometry depends solely on the sum of these two matrices.

Figure 9a shows the result of the shot after applying reciprocity, and after interpolation. This matrix, due to the technique used for the application of reciprocity, it should present a large amount of zeroed information as shown in Equation 6. However, Figure 9a is already the result of removing these zeros found in the matrix.

Finally, a comparison can be made between the original modeled data before zeroing the shots and removing half of the information from each seismogram (Figure 7b), and the result of generating the split-spread after the interpolation method (Figure 9a).

In the reciprocity stage, which occurs only after the elevation statics, and interpolation, there are changes in the position of sources and receivers. When changing the positions, the elevation information is also exchanged.

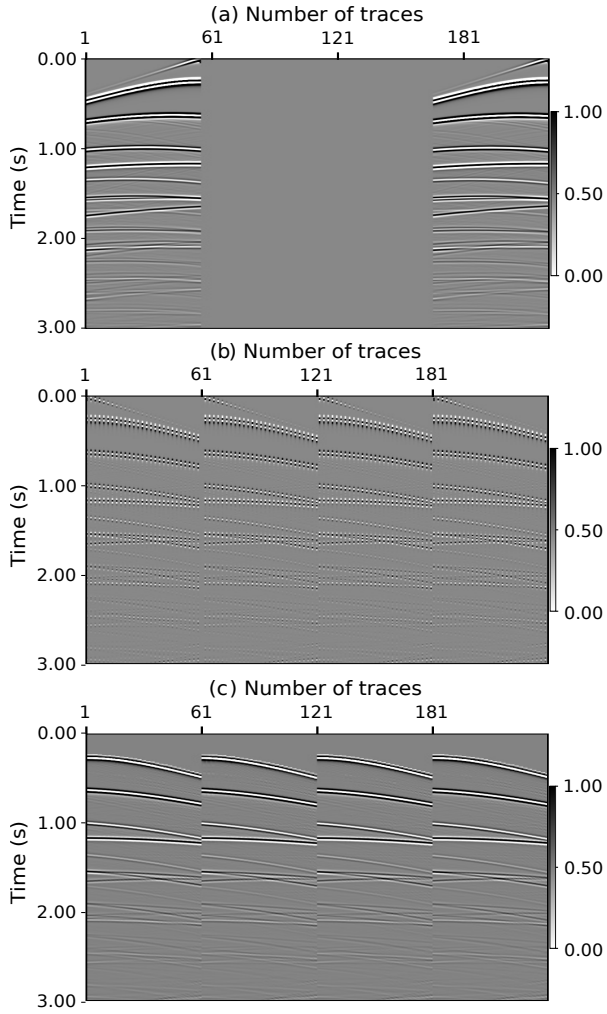

Figure 8: In (a) we see 4 shots neighboring each other, where two central shots are zeroed to be reconstructed by interpolation. This represents a common marine acquisition, where the spacing between sources is 3 times greaters than the spacing between receivers. In (b) the same data in (a), but now, it is organized in common receiver gather. Then we can see the two zeroed traces between the original ones in greater detail. In (c) the result of the interpolation in common receiver gathers, where the two zero traces were reconstructed.

If the sources and receivers are in the same datum, the construction of the data in the split-spread configuration using the reciprocity step works satisfactorily. However, when they are not at the same datum, the unsatisfactory construction of the data in split-spread geometry can lead to errors in the application of the Marchenko method.

For the case of flat and horizontal layers, the error is equal to zero in the application of reciprocity, since exchanging the elevation information of sources and receivers the result is the same data. However, for lateral variations in the layer (Figure 6), the error should not be zero. Because of this, elevation statics is required.

We have a strong influence of the lack of information on the data in the reciprocity stage. In this case, the resulting split-spread shot would have portions of the data where there is no information, traces and entire pieces of the zeroed

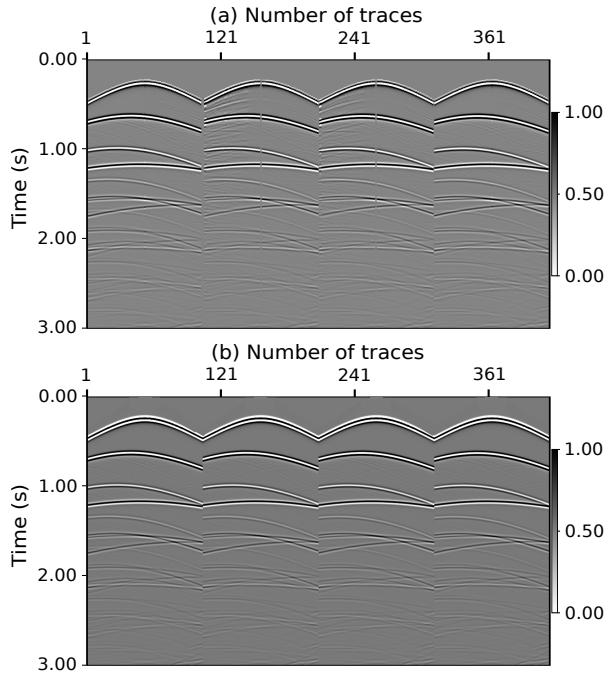

Figure 9: In (a) the result of reciprocity using as input the shots from Figure 7c. In (b) we can compare the result in (a) with the original shot as shown in Figure 7b.

shots, which reinforces the need for interpolation of the original dataset.

Conclusions

In this work, we presented three steps in the preprocessing stage before the application of Marchenko. Here we shown with numerical examples the results of elevation statics, interpolation, and reciprocity in synthetic datasets, with the objective of replicating the methods in real marine datasets. We demonstrate that the elevation statics is important to have well-positioned events in the Green's functions after retrieving it with the Marchenko method. The need for interpolation is justified considering that it is common to have a higher density of receivers than sources in real seismic acquisitions. Depending on the input seismic dataset it could be worth increasing the fold of the data, as a step before reciprocity to avoid gaps between sources and receivers. In our numerical examples, the interpolation worked well, where it was shown that the new retrieved shots have coherency with the existent shots. Finally, the reciprocity proved to work for one sided seismic acquisitions, allowing the recovery of the missing side of the shot by using a different organization of the dataset.

Acknowledgments

This work was supported by CENPES/Petrobras through the Marchenko project at SENAI CIMATEC. The authors also wish to thank to the developer team of the Open-Source libraries in TUDelft. A special acknowledgment to the geophysical team at SENAI CIMATEC for discussions and contributions.

References

Barens, L. M., van Borselen, R. G., Fokkema, J. T., and van den Berg, P. M. (1995). A new method to convert unleveled marine seismic data to leveled split-spread data. In *SEG Technical Program Expanded Abstracts 1995*, pages 1393–1396. Society of Exploration Geophysicists.

Brackenhoff, J., Thorbecke, J., and Wapenaar, K. (2019). Virtual sources and receivers in the real earth, a method for induced seismicity monitoring. *arXiv preprint arXiv:1901.03566*.

Jia, X., Guitton, A., and Snieder, R. (2018). A practical implementation of subsalt marchenko imaging with a gulf of mexico data set. *Geophysics*, 83(5):S409–S419.

Peng, H., Vasconcelos, I., Sripanich, Y., and Zhang, L. (2019). On the effects of acquisition sampling on marchenko-based focusing and primary estimation. In *81st EAGE Conference and Exhibition 2019*, volume 2019, pages 1–5. European Association of Geoscientists & Engineers.

Porsani, M. J. (1999). Seismic trace interpolation using half-step prediction filters. *Geophysics*, 64(5):1461–1467.

Singh, S., Snieder, R., van der Neut, J., Thorbecke, J., Slob, E., and Wapenaar, K. (2017). Accounting for free-surface multiples in marchenko imaging. *Geophysics*, 82(1):R19–R30.

Spitz, S. (1991). Seismic trace interpolation in the fx domain. *Geophysics*, 56(6):785–794.

Staring, M. and Wapenaar, C. (2019). Interbed demultiple using marchenko redatuming on 3d field data of the santos basin. In *Sixteenth International Congress of the Brazilian Geophysical Society*. Sociedade Brazileira de Geofísica (SBGf).

Thorbecke, J., Slob, E., Brackenhoff, J., van der Neut, J., and Wapenaar, K. (2017). Implementation of the Marchenko method. *Geophysics*, 82:WB29–WB45.

Verschuur, D. J., Berkhout, A., and Wapenaar, C. (1992). Adaptive surface-related multiple elimination. *Geophysics*, 57(9):1166–1177.

Wapenaar, K., Thorbecke, J., van der Neut, J., Broggini, F., Slob, E., and Snieder, R. (2014). Green's function retrieval from reflection data, in absence of receiver at the virtual source position. *The Journal of the Acoustical Society of America*, 135:2847–2861.

Zhang, L. and Slob, E. (2020). A field data example of marchenko multiple elimination. *Geophysics*, 85(2):S65–S70.