GEOPHYSICS, VOL. 85, NO. 3 (MAY-JUNE 2020); P. V283–V296, 21 FIGS., 1 TABLE.
10.1190/GEO2019-0341.1

Nonlinear beamforming for enhancement of 3D prestack land seismic data

Andrey Bakulin[1], Ilya Silvestrov[1], Maxim Dmitriev[1], Dmitry Neklyudov[2], Maxim Protasov[2], Kirill Gadylshin[2], and Victor Dolgov[3]

ABSTRACT

We have developed nonlinear beamforming (NLBF), a method for enhancing modern 3D prestack seismic data acquired onshore with small field arrays or single sensors in which weak reflected signals are buried beneath the strong scattered noise induced by a complex near surface. The method is based on the ideas of multidimensional stacking techniques, such as the common-reflection-surface stack and multifocusing, but it is designed specifically to improve the prestack signal-to-noise ratio of modern 3D land seismic data. Essentially, NLBF searches for coherent local events in the prestack data and then performs beamforming along the estimated surfaces. Comparing different gathers that can be extracted from modern 3D data acquired with orthogonal acquisition geometries, we determine that the cross-spread domain (CSD) is typically the most convenient and efficient. Conventional noise removal applied to modern data from small arrays or single sensors does not adequately reveal the underlying reflection signal. Instead, NLBF supplements these conventional tools and performs final aggregation of weak and still broken reflection signals, where the strength is controlled by the summation aperture. We have developed the details of the NLBF algorithm in CSD and determined the capabilities of the method on real 3D land data with the focus on enhancing reflections and early arrivals. We expect NLBF to help streamline seismic processing of modern high-channel-count and single-sensor data, leading to improved images as well as better prestack data for estimation of reservoir properties.

INTRODUCTION

Modern 3D land seismic data acquisition is moving toward dense grids of small source/receiver arrays (or ultimately single-sensor systems) instead of the sparser grids of large arrays that have been used previously (Bagaini et al., 2010; Pecholcs et al., 2012; Regone et al., 2015). In theory, better spatial sampling and the wider frequency bandwidth of seismic wavefields should lead to improved imaging after final processing. In real cases with complex near-surface conditions, use of small arrays or single sensors leads to prestack records with weaker reflection signals and stronger contamination by scattered near-surface noise, which previously was suppressed with large arrays. If we could place single sensors as densely as we have geophones in field arrays, then advanced noise-suppression techniques would provide better performance than noise suppression by analog field arrays (Özbek, 2000; Ji et al., 2010). However, such density is not possible in current land acquisition (Figures 1 and 2); as a result, signal levels in the data after noise attenuation often remain very low. Conventional time processing tools, such as surface-consistent scaling and deconvolution, statics estimation, and velocity analysis, require reliable prestack signal in the data. Their application to modern data sets acquired with small field arrays leads to unreliable results because derived parameters are based on noise rather than signal. As such, a significant processing challenge has emerged for high-channel-count seismic acquisition: How do we exploit greater trace density for imaging while mitigating the reduced quality of the raw prestack data? To tackle this challenge, it is imperative to significantly enhance the signal in the prestack data before estimating any of the prestack processing parameters (Gülünay and Benjamin, 2008; Bakulin and Erickson, 2017). Due to economic reasons, sampling of the wavefield

Manuscript received by the Editor 28 May 2019; revised manuscript received 1 January 2020; published ahead of production 31 January 2020; published online 31 March 2020.
[1]Saudi Aramco, EXPEC Advanced Research Center, Dhahran, Saudi Arabia. E-mail: andrey.bakulin@aramco.com; ilya.silvestrov@aramco.com (corresponding author); maxim.dmitriev@aramco.com.
[2]Siberian Branch of the Russian Academy of Sciences, Institute of Petroleum Geology and Geophysics, Novosibirsk, Russia. E-mail: neklyudovda@ipgg.sbras.ru; protasovmi@ipgg.sbras.ru; gadylshinkg@ipgg.sbras.ru.
[3]Saudi Aramco, Geophysical Imaging Department, Dhahran, Saudi Arabia. E-mail: victor.dolgov@aramco.com.

in high-density 3D acquisition remains sparser than sampling used by individual receivers in conventional large field arrays. Therefore, we cannot merely replicate the same level of noise removal and signal enhancement during processing as was achieved with large field arrays used in the past. New enhancement methods are required, with iterative processing approaches such as cycles of enhance-estimate-image (Bakulin and Erickson, 2017) becoming more popular. Supergrouping of adjacent traces in different domains after applying normal-moveout (NMO) correction (Bakulin et al., 2018b) has proven to be very efficient and useful for enhancement of reflections and has been incorporated successfully into novel iterative processing approaches of single-sensor data (Cordery, 2018). In this study, we strive to find a more data-driven enhancement algorithm that is free from the simplified assumptions about the global hyperbolicity of events used in supergrouping. This is particularly important for cases involving complex models and moveouts, long offsets, and anisotropy.

Data-driven multidimensional local stacking of neighboring traces may be very effective for enhancing signals of different complexity. Existing methods based on this technique include the common-reflection-surface (CRS) stack (Baykulov and Gajewski, 2009; Li et al., 2011; Spinner et al., 2012; Buzlukov and Landa, 2013; Müller et al., 2013; Xie, 2017) and multifocusing (Berkovitch et al., 2011; Rauch-Davies et al., 2013). The main idea of these approaches is to collect and stack signals from neighboring traces along traveltime surfaces that describe the moveout of the target waves. The moveout is described in terms of local wavefronts, and its parameters are estimated directly from the data using a coherency search. A recent successful application of the multifocusing method to a relatively small (270 km^2) land 3D seismic data set, acquired in a fold and thrust belt setting, has demonstrated the strong potential of multidimensional stacking techniques for prestack data enhancement (Curia et al., 2017). The challenge is to design algorithms that can incorporate 3D geometry and the vast data volumes generated by modern high-density land seismic acquisition. The typical high-channel-count surveys with small geophone arrays may have a trace density of approximately 15 million traces/km^2 (Figure 2). Denser point-receiver surveys acquired with 50,000–100,000 active channels can reach more than 100 million traces/km^2. For typical high-channel-count surveys acquired over an area of 1000 km^2, the size of the data can be approximately 70 terabytes, whereas for a single-sensor survey, it can reach 450 terabytes. In other words, for realistic large surveys of such types, the full data size can reach into the petabytes, meaning that efficiency and optimization are a must for any viable data-enhancement technique.

Buzlukov et al. (2010) and Buzlukov and Landa (2013), using ideas from the CRS method, proposed an original and efficient approach for enhancing prestack data in the common-offset domain. They used a general local representation of a traveltime surface as a second-order curve and estimated local kinematic attributes. Unfortunately, data enhancement in the common-offset domain has some disadvantages when applied to orthogonal land data acquisition due to irregular azimuthal and common-midpoint (CMP) sampling. Therefore, other domains for enhancement are of interest. Building on previous experience in enhancing 2D data (Bakulin et al., 2018a), here we introduce nonlinear beamforming (NLBF) in the cross-spread domain (CSD) — a new approach for enhancing challenging 3D prestack data acquired with modern orthogonal land seismic surveys. In NLBF, time delays are approximated locally by nonlinear (quadratic in our case) traveltime curves in contrast to the conventional plane-wave approximation. We start by studying and explaining the advantages of CSD for data enhancement. Then, we provide the details of the implemented algorithm and demonstrate synthetic and real data examples validating its signal preservation and signal enhancement properties.

CHOICE OF DOMAIN FOR DATA ENHANCEMENT

An important practical aspect for the effective implementation of data enhancement procedures based on local coherent stacking of traces is the selection of the working domain. A seismic data volume can be represented as a 5D cube with two source coordinates (x_s, y_s), two receiver coordinates (x_s, y_r), and time (t). For example, in conventional CMP processing, we usually look at a slice of this cube along the offset direction for a given CMP location. In this domain, using coherency panels for reflected events, we can approximately estimate NMO velocity, which is a global kinematic parameter used throughout the data processing workflow. The same reflection events can be observed and identified in other domains of the data volume. This is the basis for multidimensional stacking techniques that attempt to identify and use event coherency in different domains. Similarly to NMO velocity, which is a second-order approximation of reflection traveltimes using a Taylor-series expansion in the vicinity of zero offset, we can estimate kinematic parameters at nonzero offsets as well and, generally speaking, at each point of the 5D data cube. The ultimate solution should allow us to represent the kinematic properties of the recorded wavefield in detail and use them for many data processing purposes, with data enhancement being only one of them. In reality, the full representation in the 5D volume remains computationally prohibitive, so some simplifications usually are required to restrict the working domain and the number of estimated kinematic parameters. Typical domains for estimation of the kinematic parameters and data enhancement are common-shot, common-receiver, common-offset, CMP, or cross-spread. We examine these domains in more details and identify the most suitable one for enhancing seismic data using local stacking procedures when applied to modern 3D land surveys.

Modern trends in 3D land seismic data acquisition

Typical land surveys have an orthogonal layout with lines of sources lying perpendicular to lines of receivers (Vermeer, 2012), as shown in Figure 1. The shot- and receiver-station intervals are usually several times smaller than the shot- and receiver-line intervals. As a result, the trace distribution is dense in two directions (the inline direction for receivers and the crossline direction for shots), whereas the perpendicular directions (the crossline direction for receivers and the inline direction for shots) are sampled much more coarsely. As outlined in the "Introduction" section, the main trends in modern land 3D seismic acquisition can be summarized as follows:

- The size of field shot/receiver arrays used in the field is rapidly decreasing, and point-sensor surveys are becoming more popular.
- There is a general trend toward using denser source and receiver spatial distributions in the inline and crossline directions, mainly constrained by economic considerations.

Figure 2 shows a simple schematic classification of 3D land seismic surveys based on trace density defined by Cooper (2004). It illustrates typical acquisition parameters for legacy low-density data and the modern dense, high-channel-count, and single-sensor seismic surveys currently in use.

Land 3D seismic gathers for data enhancement

In the framework of data enhancement based on local coherency estimation and stacking, we formulate the desired criteria to identify the optimal working domain for such kind of algorithms when applied to typical orthogonal land 3D data sets:

1) Azimuth preservation, i.e., minimizing mixing of traces from different azimuths so that only traces in a small neighborhood around the target trace are used for estimation of local kinematic parameters and beamforming. This is an essential feature of any data-enhancement algorithm required for anisotropic processing and inversion.
2) A sufficient number of traces in each gather, meaning that there should be enough data to perform a robust estimation of kinematic parameters and summation.
3) Good density and regularity of traces, implying a uniform distribution of data within a multidimensional gather and also between gathers. This guarantees a similar quality of estimation and summation for all traces without gaps and deficient zones.
4) The simplicity of typical traveltime curves of target events that affect how accurately and efficiently estimation of local kinematic parameters can be performed.
5) Availability of gathers within a typical processing flow and low computational intensity to extract them from the entire data volume.

Let us now consider different typical domains that can be used for prestack data enhancement and analyze them based on these criteria. As an example, we will use data acquired using the acquisition parameters shown in Figure 1a.

Common-offset gathers with azimuthal binning

Common-offset gathers (COGs) are defined as a subspace of the whole data cube in which the midpoint coordinates x_m and y_m in the x- and y-directions vary, whereas the offset and azimuth are fixed. The data-enhancement algorithms have been developed and applied in this domain either without azimuth preservation (Buzlukov and Landa, 2013) or with partial preservation based on azimuthal binning (Xie and Gajewski, 2017). This is a popular domain for applying local stacking algorithms because criteria 4 in the previous list is satisfied easily in many geologic settings. In models of mild complexity, local moveout of reflected waves in the x_m and y_m directions is similar within the whole section and may be reliably approximated as planar surfaces with dip and curvature varying in limited intervals. The disadvantage of this domain is the nonuniform distribution of traces with respect to offset, as observed in Figure 3. For example, the number of traces in the short-offset (<2000 m) and far-offset (>6000 m) ranges is much smaller than in the intermediate range. As a result, the corresponding COGs will vary dramatically in trace density.

Figure 4 compares the trace distribution inside a fixed 1 km^2 area for 500 and 5000 m offsets. In both cases, the trace distribution is

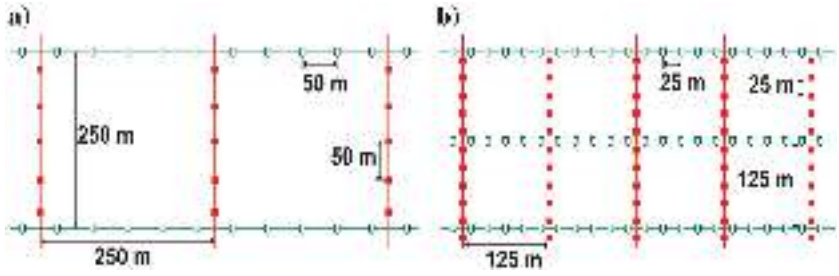

Figure 1. Examples of typical 3D land seismic acquisition geometries including (a) a legacy low-density survey and (b) a modern high-channel-count medium-density survey. Receiver stations and receiver lines are shown in green, whereas shot stations and shot lines are shown in red.

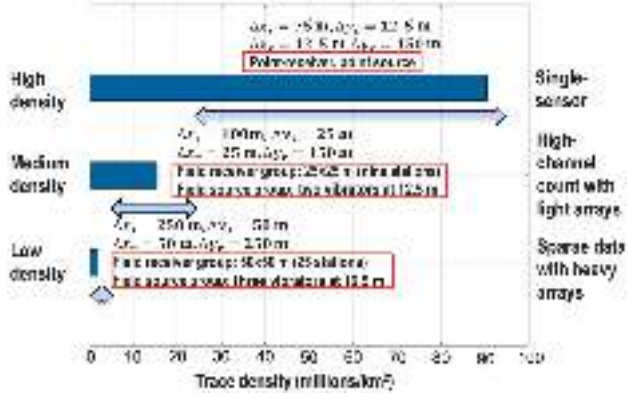

Figure 2. A simplified classification of 3D land seismic surveys based on trace density showing typical acquisition parameters along with field array configurations for sources and receivers. The terms Δx_r and Δy_r denote receiver spacings in the inline and crossline directions, respectively. Likewise, Δx_s and Δy_s denote shot spacings.

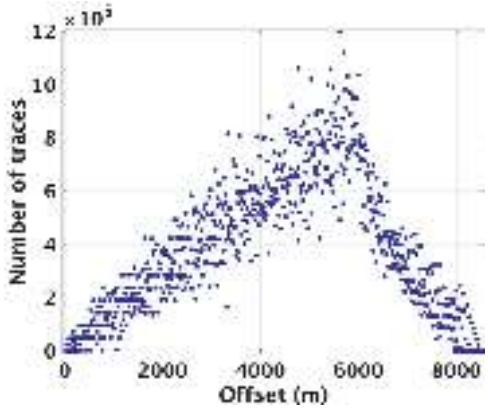

Figure 3. Offset distribution of traces in the typical low-density 3D land seismic data set with geometry from Figure 1a displays significant variability of trace density with offset for COGs. The mid-offset bins are well-populated, whereas the short- and long-offset bins are more sparse. The traces from all azimuths are shown.

quite irregular. As expected from Figure 3, COG for short offsets (Figure 4a) consists of a small number of sparsely sampled traces. Therefore, it is difficult to expect reliable results from data enhancement in this case. For the long-offset COG (Figure 4b), the number of points is sufficiently large; however, its irregular distribution in the plane would result in the variable quality of local kinematic parameters estimation and beamforming. The distributions presented in Figure 4 have been taken from the full-fold areas with the most regular organization of the recorded traces. If we approach the edges of the acquisition area, the number of traces in the COGs will decrease considerably. In this example, azimuthal binning in the 0°–45° range is used. The COGs constructed without taking into account azimuth would have much denser trace distribution, but the general trends mentioned above would remain the same. In addition, without azimuth binning, such gathers will contain traces located physically far from each other, and their mixing will smear the anisotropic effects as well as the surface-consistent factors we are attempting to resolve in the data.

In summary, traces in COGs have very predictable and favorable local moveout behavior, allowing effective constraints for the search intervals of local kinematic parameters. However, COGs have some significant disadvantages including the highly irregular trace distribution within a single gather. In addition, the trace density varies considerably between the gathers, and it might become insufficient for robust data enhancement of near- and far-offset traces. Finally, azimuth preservation would require additional sectoring, sorting, and storing.

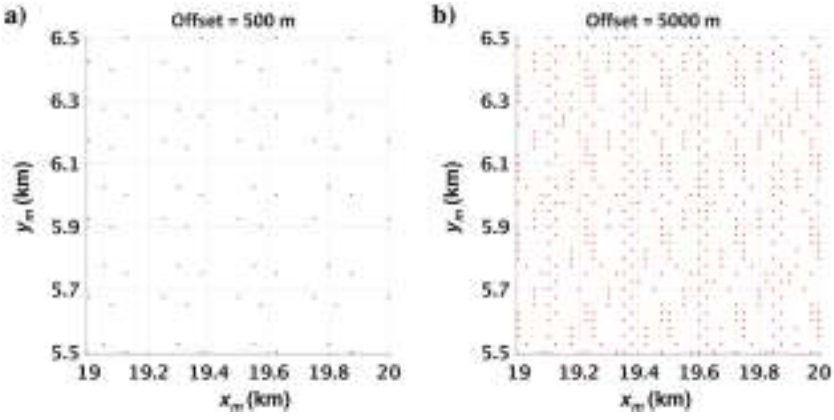

Figure 4. Distribution of traces within COGs with azimuthal binning shown for (a) small- (500 m) and (b) medium-offset ranges (5000 m).

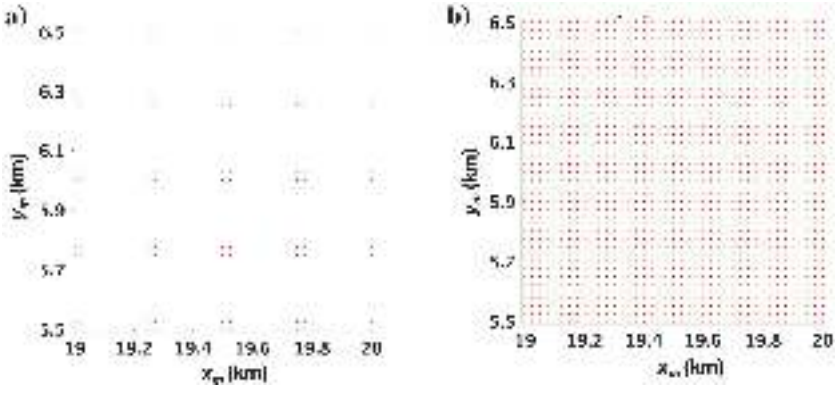

Figure 5. Distribution of traces within OVT gathers for $h_x = 2000$ m and $h_y = 2500$ m. When the dimension of the tile (here 100 m) is smaller than twice the distance between the inlines (250 m in this case) as shown in (a) then the trace distribution is relatively coarse and irregular. Increasing the dimension of the tile to 500 m (b) allows us to obtain a fine and regular distribution, but it has the ramification of mixing quite different offsets into a single ensemble for enhancement.

Offset-vector tiles

Offset-vector tiles (OVTs) are another choice of domain for orthogonal surveys that is similar in some aspects to the common-offset domain with azimuthal binning. The easiest way to think about common-offset, common-azimuth data from the survey is to consider all of the traces with a constant offset vector $\mathbf{h} = (h_x, h_y)$. The OVT gather is constructed by additional binning of the offset space and by taking all of the traces with offset vectors lying inside the tile $(h_x \pm \Delta x/2, h_y \pm \Delta y/2)$ with dimensions Δx and Δy in the inline and crossline directions, respectively. When the size of the tile is small, few traces are located inside each OVT gather, as shown in Figure 5a. In this case, the considerable distance between traces and the small overall number make estimation of the kinematic parameters and signal improvement through summation challenging. To obtain fine and regular sampling in this case, one needs to use OVTs with sizes equal to the distances between the source and receiver lines, as shown in Figure 5b. This provides excellent coverage but leads to the mixing of traces with different offsets into one gather, which results in smearing of signals within the OVT. In addition, the moveout of traces from these different offsets can vary significantly, especially for short offsets. Effectively, this breaks the main advantage of OVT gathers consisting of simple and relatively flat moveouts. It is worth mentioning that during the processing workflow, OVT gathers typically are constructed at the final stage before migration and are not readily accessible at earlier steps when data enhancement usually is applied. This also limits the use of the OVT domain for prestack data enhancement.

Common-depth-point gathers with azimuthal binning

The next domain we will discuss is characterized by two coordinates (x_m, offset), whereas the y_m and azimuth ranges are fixed. Here, a set of common-depth-point (CDP) gathers is taken along one inline direction. The local moveout behavior in this domain is not as favorable as in the case of COGs. Indeed, in CDP gathers, hyperbolic moveout in the offset direction exists together with planar moveout in the x_m direction. The dip and curvature of hyperbolas vary significantly along with the offset. Therefore, intervals of parameter estimation should be wide enough or offset dependent. This fact considerably reduces the effectiveness of any moveout estimation scheme. To narrow search intervals in the offset direction, preliminary NMO corrections

may be applied before data enhancement, mitigating the issue. Figure 6 shows trace distributions for inline CDP gathers. For areas with a full fold and an offset range between 2 and 3 km, the trace distribution is quite irregular (Figure 6a). Despite the overall good trace density, there are still some local gaps remaining. Distribution becomes even sparser in areas of the reduced fold (Figure 6b), usually associated with the edges of the survey. Such deterioration would substantially inhibit reliable and uniform data enhancement.

Common-source and common-receiver domains

The most straightforward domains are the common-source and common-receiver domains, in which the shot (or receiver) position is fixed, and the trace coordinate is represented by the inline and crossline coordinates of receivers (or shots) at the surface. The distribution of traces in these domains is regular and can be readily observed in Figure 1. The disadvantage of these two domains is the much coarser sampling of traces in one direction compared to another, leading to a lack of accuracy in the estimation of kinematic parameters in the crossline directions and difficulties in signal preservation during summation due to the nonuniform grouping of traces. The moveouts of the events are also appreciable and offset dependent, so global NMO correction is required to reduce search intervals of kinematic parameters.

Cross-spread gathers

Finally, we consider the case of CSD. A cross-spread gather is formed by a single receiver line and a single orthogonal source line (Figure 7a). For 3D orthogonal shooting geometries shown in Figure 1, CSD has the densest and most regular sampling in both directions. As a result, they are widely used in 3D processing, mainly as a domain for effective noise removal, and usually are available at the early stages of the land processing flow. Figure 7 shows the corresponding trace distribution away from the edges of the survey. The physical proximity of the traces within the ensemble guarantees similar azimuths without any additional binning. Even at the edges of a survey, trace distribution remains regular (Figure 8). The presence of existing acquisition gaps does not affect the uniformity of enhancement in neighboring areas. The only disadvantage of CSD is the considerable variability of local moveout within a gather. By definition, a single cross-spread gather consists of traces taken from many neighboring areal common-shot/common-receiver gathers. It is naturally comprised of "near-offset" and "far-offset" segments and anything in between. The moveout of reflected arrivals in these cases is not as predictable as in the COG or CDP domains. This disadvantage can be mitigated by applying preliminary NMO corrections before enhancement. The main idea is to make reflection events approximately flat and estimate residual moveout (RMO) rather than absolute moveout.

After analysis of the outlined domains, we conclude that CSD after NMO correction is the most appropriate domain for effective data enhancement based on local multidimensional stacking. The ability of CSD to naturally preserve azimuths in combination with dense and regular trace distribution outweighs the difficulty caused by more complex moveout behavior (which can be alleviated by performing preliminary NMO correction). The required high efficiency of a data enhancement algorithm for massive data volumes can be achieved by straightforward parallelization and independent enhancement of different cross-spread gathers on separate cluster nodes.

NONLINEAR BEAMFORMING IN THE CROSS-SPREAD DOMAIN

This section describes the algorithm for 3D data enhancement based on nonlinear time-delay beamforming in CSD. In its general form, the method can be written as a local summation:

$$u(\mathbf{x}^0, t_0) = \sum_{\mathbf{x}\in B_0} w(\mathbf{x}, \mathbf{x}^0) u(\mathbf{x}, t_0 + \Delta t(\mathbf{x}, \mathbf{x}^0)). \quad (1)$$

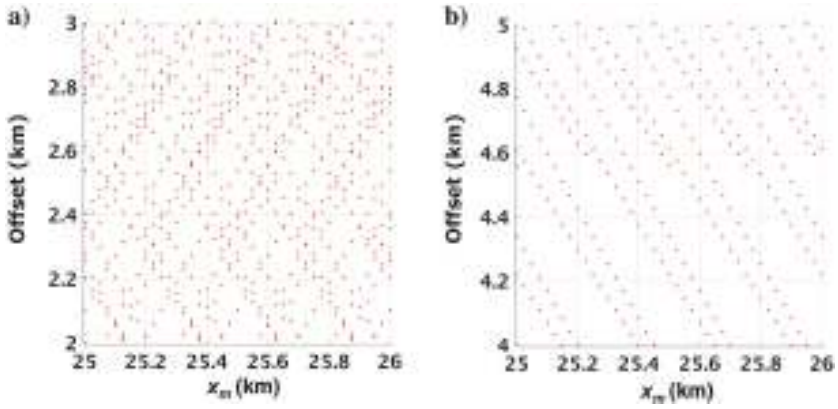

Figure 6. Trace distribution in inline CDP gathers (x_m, offset) with azimuthal binning of 0°–45°: (a) surface region from the full-fold area covering offset range of 2–3 km and (b) region from an area with limited fold (the edge of the survey).

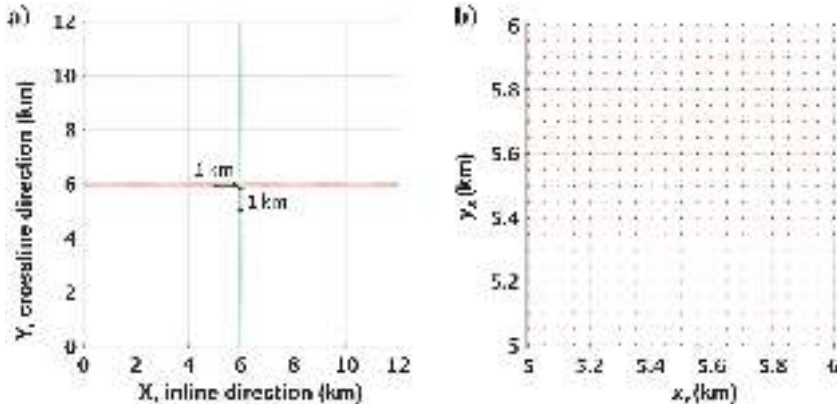

Figure 7. (a) A cross-spread is formed by a single receiver line (red) and an orthogonal source line (green). By design, the coordinates x_r and y_s vary within the cross-spread gather, whereas y_r and x_s are fixed. (b) The corresponding trace distribution for a 1 km^2 region shown with arrows in (a). The red dots show positions of the traces.

Here, $u(\mathbf{x}, t)$ represents a trace with coordinates given by vector $\mathbf{x} = (x_r, y_s)$ with varying crossline shot coordinate y_s and inline receiver coordinate x_r, keeping the inline coordinate of the shots and the crossline coordinates of the receivers fixed (Figure 9a). The summation is done over a local vicinity B_0 of the enhanced trace with coordinates $\mathbf{x}^0 = (x_r^0, y_s^0)$. The beamforming weights $w(\mathbf{x}, \mathbf{x}^0)$ are used to preserve signal energy and additionally to suppress noise. In the following examples, we use a simple fold normalization, although more sophisticated approaches also can be adopted. The traveltime moveout $\Delta t(\mathbf{x}, \mathbf{x}^0)$ with respect to the current time sample t_0 of the enhanced trace is approximated using a second-order surface:

$$\begin{aligned}\Delta t &= t(x_r^0 + \Delta x_r, y_s^0 + \Delta y_s) - t(x_r^0, y_s^0) \\ &= A\Delta x_r + B\Delta y_s + C\Delta x_r \Delta y_s + D\Delta x_r^2 + E\Delta y_s^2, \end{aligned} \quad (2)$$

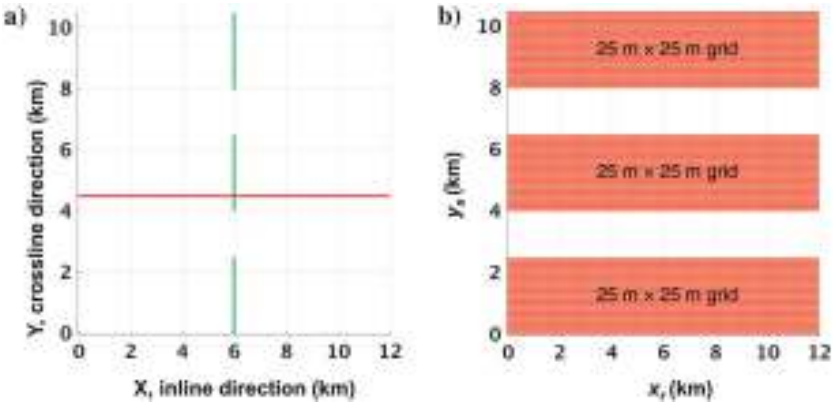

Figure 8. Acquisition geometry showing (a) a cross-spread layout taken from the edge of a typical 3D low-density data set and (b) the full trace distribution in the (x_r, y_s) plane. The red rectangles show the areas where we have uniform trace distribution with a grid of 25 × 25 m.

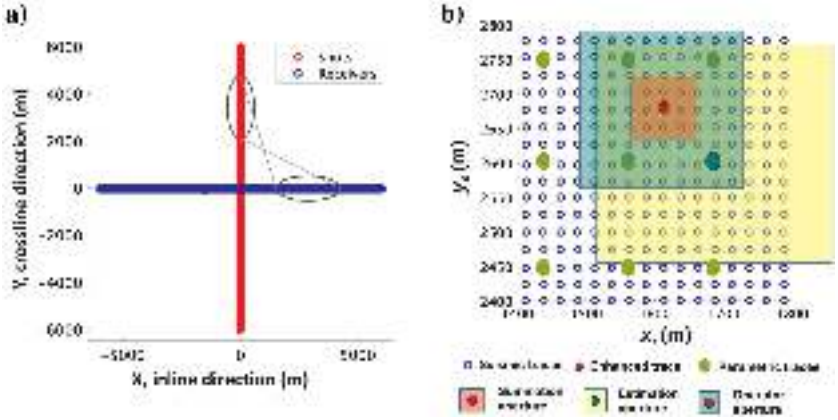

Figure 9. Implementation of the data enhancement algorithm in CSD: (a) a diagram of a single cross-spread gather from a typical 3D high-channel-count land survey. An ensemble for enhancement is formed by selecting traces with specific small ranges of source Y and receiver X coordinates shown by two ellipses. In (b), the ensemble selected from (a) is redrawn on a 2D plane used for data enhancement. Traces within the ensemble are from the same physical proximity at the surface; as a result, they possess similar offsets and azimuths. The green dots schematically show the locations of the parameter traces, where local kinematic parameters are estimated. The location of the target enhanced trace is shown in red. Different colors denote the estimation, operator, and summation apertures, respectively.

where Δx_r and Δy_s are the distances from the position of the current trace in the corresponding directions, and A, B, C, D, and E are the local kinematic parameters representing first- and second-order derivatives of the local wavefront.

Operator-oriented approach

The first step in the beamforming procedure is the estimation of local kinematic parameters that describe coherent local events. The most straightforward approach is to estimate them at each trace of the data volume, but considering the vast amount of data for enhancement, this is usually not feasible. In this work, we adopt the so-called operator-oriented approach (Hoecht et al., 2009). The term "operator" is used here to define a traveltime surface along which the local moveout correction is done. The operator-oriented and the conventional trace-oriented schemes use different strategies for where to estimate the local kinematic parameters, how to introduce moveout corrections, and how to perform summation. For the conventional trace-oriented scheme (Figure 10), the kinematic parameters are estimated directly at the position of the target trace and are used to construct the operator at the target time sample itself. The local moveout correction is done with respect to this sample, which results in one unique operator per output sample. As a result, the estimation of parameters should be done at every location containing the input trace. By contrast, in the operator-oriented scheme (Figures 9b and 11), special auxiliary parameter traces are introduced for the estimation and construction of the traveltime operators. These auxiliary parameter traces do not necessarily coincide with the actual traces. The estimation of the kinematic parameters in these auxiliary parameter traces with coordinates $\mathbf{x}^p = (x_r^p, y_s^p)$ is achieved by performing local summation similar to equation 1:

$$u(\mathbf{x}^p, t_p) = \sum_{\mathbf{x} \in \Pi_{\text{est}}} w(\mathbf{x}, \mathbf{x}^p) u(\mathbf{x}, t_p + \Delta t(\mathbf{x}, \mathbf{x}^p)) \quad (3)$$

and searching for the values of A, B, C, D, and E that provide maximum semblance along the corresponding local traveltime curve in equation 2. The set of traces used in the calculation of semblance is defined as $\Pi_{\text{est}} = \{\mathbf{x} = (x_r, y_s): |x_r - x_r^p| < A_{\text{est}}, |y_s - y_s^p| < A_{\text{est}}\}$, where estimation aperture A_{est} is the maximum distance in the inline and crossline directions from the parameter trace coordinates. The estimation of parameters is done in selected time samples t_p of parameter traces and by using all actual data traces within an area constrained by the estimation aperture (Figure 9b). The data enhancement process for the trace with coordinates $\mathbf{x} = (x_r, y_s)$ can be written as a double summation of the following form:

$$u(\mathbf{x},t) = \sum_{\mathbf{x}^p \in \Pi_{\text{op}}} w(\mathbf{x},\mathbf{x}^p) \sum_{\hat{\mathbf{x}} \in \Pi_{\text{sum}}} w(\mathbf{x},\hat{\mathbf{x}}) u(\hat{\mathbf{x}}, t - \Delta t(\mathbf{x},\mathbf{x}^p) + \Delta t(\hat{\mathbf{x}},\mathbf{x}^p)). \quad (4)$$

Here, $\Pi_{\text{sum}} = \{\hat{\mathbf{x}} = (\hat{x}_r, \hat{y}_s) : |\hat{x}_r - x_r| < A_{\text{sum}}, |\hat{y}_s - y_s| < A_{\text{sum}}\}$ is the set of actual traces in the vicinity of the output trace defined by summation aperture A_{sum} and $\Pi_{\text{op}} = \{\mathbf{x}^p = (x_r^p, y_s^p) : |x_r^p - x_r| < A_{\text{op}}, |y_s^p - y_s| < A_{\text{op}}\}$ is the set of parameter traces within operator aperture A_{op} from the output trace that defines traveltime curves for this summation. In this double-summation scheme, several different operators can contribute to one target sample (Figure 11). The number of operators depends on the operator aperture A_{op}, which defines the area of influence such that all parameter traces falling into this area are chosen to contribute during beamforming according to their estimated operators. One of the main advantages of the operator-oriented scheme is the possibility of using a sparser grid of parameter traces for estimation to improve the computational performance of the algorithm. Another potential advantage is the ability to partially resolve the conflicting dips problem. Because several different operators may contribute to each output time sample of the target trace, there is a chance for them to belong to different moveouts and contribute energy from distinct local events with conflicting dips (Hoecht et al., 2009). However, because we do not select among the contributing operators, their influence may be unbalanced. More "proper" handling of the conflicting dips situation requires explicit identification of different events (Mann, 2001; Müller, 2009) and is more computationally demanding.

Choice of apertures

The proper choice of apertures from equations 3 and 4 is paramount to achieving computational efficiency and obtaining a good quality of the enhanced output data set. There are three different apertures involved: the summation aperture (A_{sum}), estimation aperture (A_{est}), and operator aperture (A_{op}). The summation aperture defines the area used for beamforming (stacking). The estimation aperture specifies the neighborhood of traces used for estimating the local wavefront parameters. Finally, the operator aperture outlines the area of influence from which operators are drawn for beamforming. In addition to apertures, another critical parameter of the algorithm is the spacing between adjacent parameter traces that defines the density of the estimation grid.

In the current implementation, the summation aperture is selected as the most intuitive primary parameter specified by the user. Theoretical considerations for choosing the size of the summation aperture based on the size of the Fresnel zones can be found in Mann (2002) and Faccipieri et al. (2016). In Appendix A, we present a simplified example showing the accuracy of the traveltime approximation in equation 2 as related to the size of the aperture. For real data, the selection of summation aperture is a practical trade-off between achieving the desired signal-to-noise ratio improvement and limiting oversmoothing of the output signals. By definition, the summation aperture controls the number of neighboring traces involved in local summation to produce the enhanced output trace. Too small of a summation aperture may fail to deliver the desired signal-to-noise ratio. However, if the summation aperture is too large, signals could be become overly smoothed and suffer a loss of higher frequencies. These consequences can escalate rather quickly in land seismic data when stacking traces recorded in different near-surface conditions having substantial variations in waveforms and are affected by variable static delays. The choice of summation aperture for optimal beamforming relies on the assumption that signal waveforms vary in a limited manner within the aperture. However, this cannot be guaranteed in practice with real data. Therefore, each particular case requires individual consideration. In general, our experience shows that good practical beamforming results can be obtained using summation apertures of approximately 100–500 m. To take into account the divergence of the wavefront propagating through the subsurface, the summation aperture is

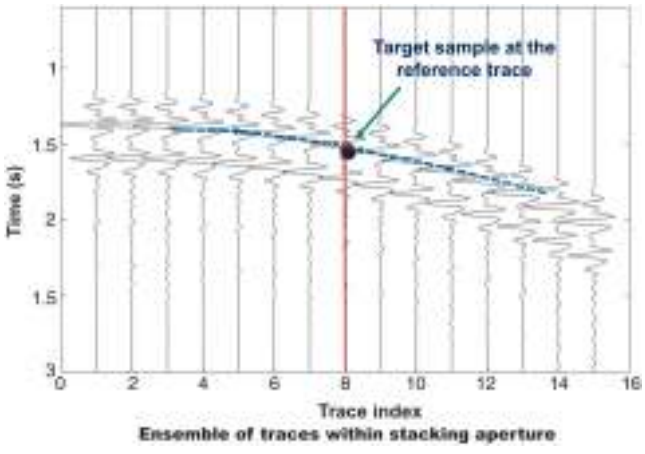

Figure 10. In the trace-oriented estimation approach, a single traveltime operator (blue line) is centered directly at the target sample (black dot coinciding with red dot). The summation of the neighboring traces is done along this single operator. In essence, each parameter trace coincides with each reference trace.

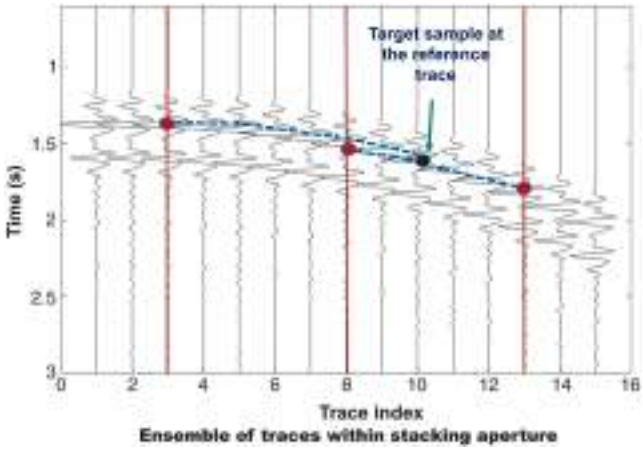

Figure 11. In the operator-oriented approach, local kinematic attributes are estimated on the sparser grid of locations of the so-called parameter traces. Each time sample of the single parameter trace (red lines) contains five kinematic parameters defining the local operator at this fixed traveltime. Local moveout operators are launched from the parameter traces (red dots) and continued until they intersect the target trace near a time sample of interest (black dot). The summation is done along the local traveltime surfaces (blue lines) using contributions from multiple nearby operators (within the area of influence defined by the operator aperture). The beamformed signal value is placed into the target sample of the reference trace (black dot).

linearly increasing with depth. The actual aperture number defines the beamforming area at some target two-way time provided by the user and decreases in a linear fashion reaching 10%–15% of this value at the surface.

The density of the parameter estimation grid is decided by a trade-off between the computational cost and the quality of data enhancement. If computational resources are unlimited, then parameter estimation can be performed in every trace. With limited computer resources, it is desirable to reduce the computational effort required, especially for the massive high-channel-count and single-sensor data sets. Using numerical experiments on synthetic and real data, we conclude that a reasonable trade-off for practical applications is reached when we have only three estimation points within the summation aperture in each direction on a 2D plane. In this case, enhanced data quality experiences very minor degradation compared to the ideal case, whereas a coarser grid significantly improves the computational performance of the algorithm. However, using fewer than two operator traces within two summation apertures usually leads to noticeable deterioration that should be avoided. As a result, we arrive at the following practical recipe between summation aperture A_{sum} and spacing of the estimation grid Δh:

$$\frac{A_{\text{sum}}}{3} \le \Delta h \le A_{\text{sum}}. \tag{5}$$

To relate estimation and operator apertures, we set a reasonable requirement that all traces in an ensemble to be summed must have been used for estimation of parameters used for stacking. This can be summarized using the following conditions further demonstrated in Figure 12:

$$A_{\text{est}} \ge 2A_{\text{sum}}, \quad A_{\text{op}} \le A_{\text{est}} - A_{\text{sum}}. \tag{6}$$

In practice, we usually take A_{est} from the interval $[2A_{\text{sum}}, 2.5A_{\text{sum}}]$, which places A_{op} in the interval $[A_{\text{sum}}, 1.5A_{\text{sum}}]$. Even though the

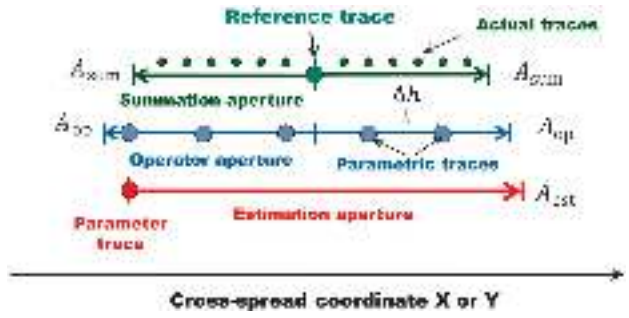

Figure 12. Relationships between three types of apertures demonstrated along one direction of the cross-spread coordinate plane. A reference trace to be enhanced is shown as a large green dot. The smaller green dots denote neighboring actual traces to be used in local summation. The blue dots show parameter traces used for operator construction (for a given reference trace). The red dot depicts the leftmost parameter trace within the operator aperture. As explained in the text, the estimation aperture shown for the leftmost parameter trace (red dot) encloses all actual traces that would be used for the final summation.

above-mentioned relations are based on common-sense considerations rather than on rigorous theoretical examination, they prove to be quite useful in application to a wide range of synthetic and real data sets.

Estimation approach based on "2 + 2 + 1" scheme

We estimate the local kinematic parameters A, B, C, D, and E by scanning many different beamforming surfaces and finding the best coherency defined by the maximum value of a semblance function. Optimization can be implemented either as a simultaneous 5D search for all five parameters or sequential estimation of one parameter after another as in the coordinate-descent method (Mann, 2002). The advantage of the simultaneous search is a more stable and accurate evaluation of the parameters, especially in the case of noisy data. The sequential strategy is more computationally attractive; however, it may become trapped in a local maximum and produce erroneous results for low signal-to-noise-ratio data. Here we implement the sequential hybrid 2 + 2 + 1 strategy, searching first for a pair of parameters A and D in the inline directions for receivers, then for pair of parameters B and E in the crossline direction for sources, and finally for the remaining parameter C in all directions with the first four parameters being fixed (Buzlukov and Landa, 2013). If optimization of the semblance function is done by a brute-force 5D search, probing each parameter bounded by a predefined search interval, the computational cost of the simultaneous method can be roughly estimated as

$$N_A \cdot N_B \cdot N_C \cdot N_D \cdot N_E \cdot N_R \cdot N_S, \tag{7}$$

where N_A, N_B, N_C, N_D, and N_E denote the number of probing values and N_R and N_S designate the number of receivers and sources in the estimation aperture. For the hybrid sequential $2+2+1$ method, we recall that source and receiver directions are scanned independently. As a result, the number of calculations can be approximated as

$$N_A \cdot N_D \cdot N_R + N_B \cdot N_E \cdot N_S + N_C \cdot N_R \cdot N_S. \tag{8}$$

Assuming that all multipliers are of the same order, we can characterize the performance of 5D and $2+2+1$ method as $O(N^7)$ and $O(N^3)$, respectively, meaning the latter is four orders of magnitude more efficient. As a result, the 5D simultaneous search does not appear computationally feasible for large data sets, so we opt for the more efficient hybrid $2+2+1$ approach for real data applications. In practice, the search for parameters is not entirely independent for sources and receivers, and several additional adjacent lines of receivers and sources usually are taken to stabilize the search and avoid the local maximum problem. A comparison of beamformed data with 5D and $2+2+1$ approaches usually does not reveal significant differences for a wide range of real data examples; therefore, we conclude that the $2+2+1$ method strikes a sensible balance between computational cost and data quality for large data sets.

Local summation procedure

The final step in the NLBF data enhancement procedure is multidimensional local summation or beamforming of traces along the estimated surfaces to produce an output data set with an increased

signal-to-noise ratio. The simplest method in which the loop in equation 4 is done over the target traces with accompanying signal accumulation is inefficient and slow due to repeated uses of the same operators over and over again. Instead, we have developed a fast summation approach that can be characterized as a parameter-trace-oriented algorithm. First, the signal is accumulated inside samples of the traces where the parameters were estimated and only then is moved to the samples of the target output traces. Compared to the straightforward approach, this fast method significantly increases the speed of the summation phase. A detailed description of this fast summation algorithm is presented in Appendix B.

SYNTHETIC DATA EXAMPLE

The main goal of the synthetic experiment is to investigate signal preservation properties of the developed NLBF algorithm. We use the SEAM Arid Model (Oristaglio, 2015) and regenerate data without the free surface using 2D finite-difference acoustic algorithm for one cross section of the model shown in Figure 13. This allows us to avoid major challenges of the model caused by complex near-surface conditions and to focus only on the preservation of reflection events representing signal to be beamformed. Source and receiver spacing intervals are chosen to be 25 m. An example of a synthetic common-shot gather is presented in Figure 14a. Because we use 2D data in this example and do not have full orthogonal cross-spreads, we apply NLBF independently to each common-shot gather only along the direction of the receiver coordinate. This implies that only parameters A and D are used and optimized in equation 2, whereas other coefficients are not required for summation. The beamforming is applied after preliminary NMO correction with a summation aperture of 400 m. The spacing between parameter traces is chosen to be 140 m, according to equation 5. The operator and estimation apertures are selected based on equation 6 and are equal to 400 and 800 m, respectively. Because parameter estimation has been done after NMO correction, the search intervals are chosen to be quite narrow: $\pm 10 \times 10^{-5}$ s/m for parameter A (equal to 40 ms maximum moveout across the summation aperture of 400 m) and $\pm 5 \times 10^{-7}$ s/m^2 for parameter D (20 ms moveout). The common-shot gather after NLBF is shown in Figure 14b. The difference between gathers before and after NLBF (Figure 15c) exhibits low residual energy along the trajectories of the primary reflection events, whereas diffraction events with steeper dips are attenuated by the algorithm as expected. The stacked sections obtained using the original and enhanced data are also very similar (Figure 15), validating that prestack reflections are well-preserved and remain unharmed by NLBF, whereas the diffractions are partially suppressed by the algorithm.

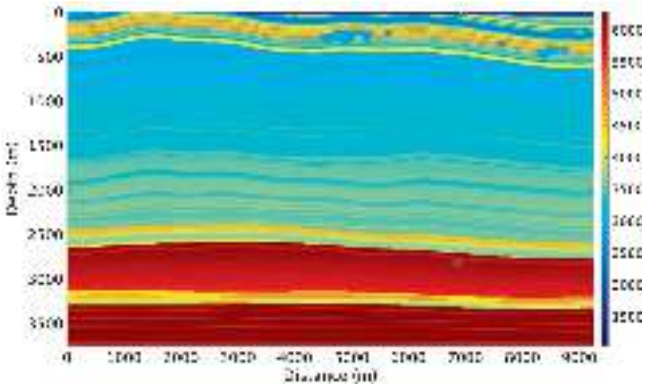

Figure 13. Cross section from the SEAM Arid Model used for generating synthetic example. Observe karst bodies in the near surface.

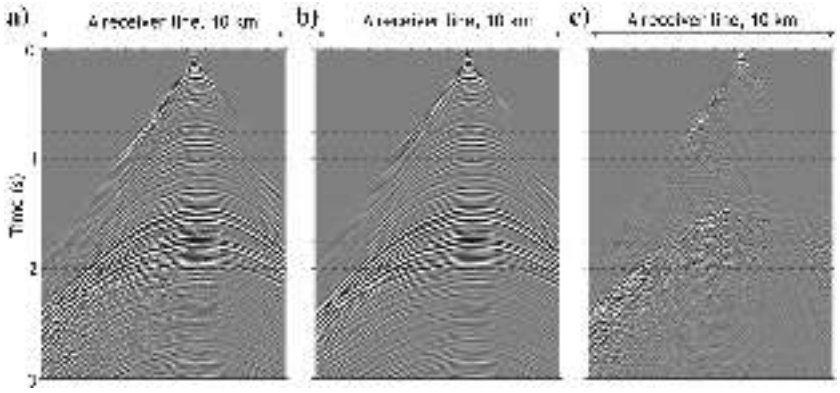

Figure 14. Synthetic common-shot gather for the SEAM Arid Model: (a) the original data, (b) the enhanced data after nonlinear beamforming with a summation aperture of 400 m, and (c) the difference between the two shown with the same scaling. Observe the good preservation of reflections and suppression of some diffracted energy from shallow karsts.

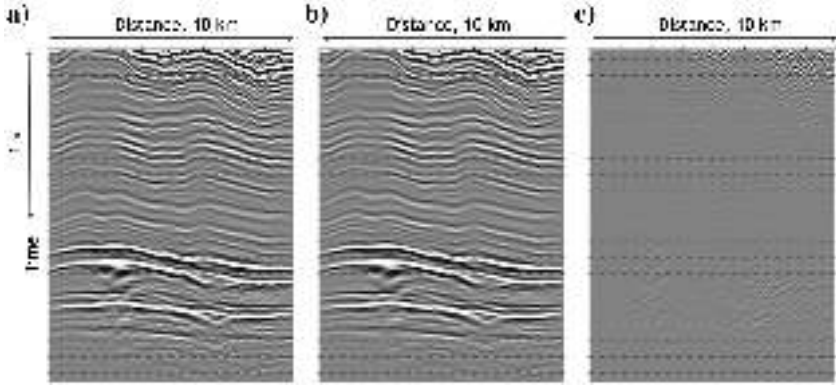

Figure 15. (a) Stack sections for the SEAM Arid Model: (a) original data, (b) enhanced data after nonlinear beamforming with summation aperture of 400 m, and (c) the difference between the two shown with the same scaling. Observe the good subtraction of reflected energy implying signal preservation. The difference image is dominated by diffracted energy that was suppressed by beamforming focused on enhancing reflections only in this example.

REAL DATA EXAMPLES

Single-sensor data: Enhancement of reflections

In this real data example, we use single-sensor data acquired onshore on the Arabian Peninsula with an orthogonal 3D source/receiver geometry. The receivers are separated by 20 m along the receiver lines, with the lines separated by 120 m. Likewise, sources are sampled at 20 m along source lines (orthogonal to the receiver lines) with 80 m between shot lines. The nominal fold of this data set is 2128. In the first example related to this data, we focus on enhancing reflection events. The data after heavy conventional preprocessing and noise attenuation exhibit weak reflections, as can be observed in Figures 16a and 17a. Here, NLBF is done in CSD with a summation aperture of 250 m. The beamforming is performed after NMO correction with intervals of $\pm 8 \times 10^{-5}$ s/m for parameters A and B (or 20 ms maximum moveout across the summation aperture) and $\pm 1.6 \times 10^{-7}$ s/m^2 for parameters C, D, and E (or 10 ms across the summation aperture). As shown in Figure 16b, NLBF enhances reflections in CSD. The level of enhancement can be further evaluated after resorting the data into the CMP gathers (Figure 17). We stress that sorting the data into different domains after noise attenuation or signal enhancement often results in increased "apparent" noise that has not been attacked in the CMP domain.

In contrast, data after NLBF exhibit robust improvement of the signals observed in the CMP domain as well. The reflections have become much stronger and more coherent, making the data more manageable for velocity analysis and estimation of surface-consistent scalars and deconvolution operators. A significant amount of noise has been removed from prestack gathers, whereas no evidence of signal leakage is observed in the difference between the original and the enhanced data (Figure 17c).

The resulting stacked sections obtained using data before and after the enhancement (and using the same velocity model) allow us to perform further cross-check of the NLBF results (Figure 18). Sections after NLBF show a significant reduction of noise and greatly improved continuity of reflections in shallow and deep sections. At the same time, amplitude variation across the section appears more geologically plausible for the area under examination. The difference between the two sections shows the noise that was robustly removed, whereas residual signal leakage is very weak and considered to be acceptable for such challenging data.

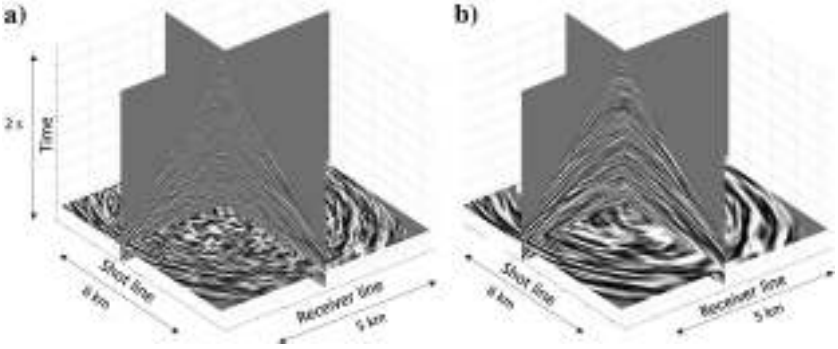

Figure 16. (a) Cross-spread gather from land single-sensor acquisition after heavy preprocessing exhibits weak reflection signals. (b) Nonlinear beamforming in CSD dramatically improves the signal-to-noise ratio and reveals visible reflections.

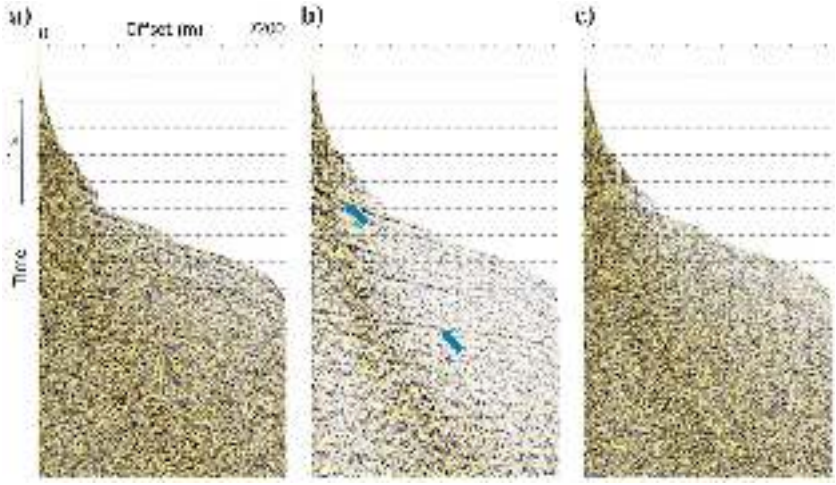

Figure 17. (a) The CMP gather from land single-sensor acquisition after heavy preprocessing and noise attenuation does not reveal strong coherent events. (b) After applying nonlinear beamforming in CSD and resorting to CDPs, the reflection events become stronger and more coherent. (c) The difference between the two gathers showing noise that has been removed from the data during the beamforming process.

Single-sensor data: Enhancement of first arrivals

Another application of the NLBF algorithm is the improvement of early arrivals for first-break picking and/or full-waveform inversion. In this experiment, we use the same single-sensor data set as in the previous case. Examples of common-shot gathers used for first-break picking after conventional preprocessing and muting are shown in Figure 19 (left column). The automatic picking results before enhancement are shown in blue. First breaks are barely visible and are challenging to pick, especially for far offsets. At offsets larger than 2.2 km, first arrivals are missed completely, and the picking algorithm is merely tracking the mute function. As before, NLBF is applied to these data in CSD with a summation aperture of 150 m. In contrast to the previous experiments, no moveout correction has been applied before data enhancement. Therefore, larger search intervals for the parameters are used: $\pm 1 \times 10^{-3}$ s/m for parameters A and B (or 150 ms maximum moveout across the summation aperture) and $\pm 8.9 \times 10^{-7}$ s/m^2 for parameters C, D, and E (or 20 ms moveout across the summation aperture). The right column in Figure 19 shows that the first breaks after enhancement become much more coherent and can be reliably obtained using the same automatic picking tool. Specifically, the enhancement enables robust picking of traveltimes at long offsets beyond 2200 m. The picks before the enhancement are entirely incorrect in this range and merely follow the mute function, as can be observed

in Figure 20. After beamforming, the picks start to track the first arrivals that now are revealed, and they become useful for near-surface characterization with the help of refraction tomography or other methods.

DISCUSSION

From the numerical examples, we observe that the proposed algorithm provides strong data enhancement capabilities in the case of very noisy land seismic data. We successfully apply it to the enhancement of reflection signals and early arrivals in single-sensor data. The presented synthetic and real data examples validate the good amplitude and signal preservation character of the algorithm even when using relatively large summation apertures. The quality of enhanced data is strongly controlled by the input parameters, in particular, the size of the summation apertures, as well as minimum and maximum dips and curvatures of the targeted events. Smaller apertures provide better signal preservation but may lead to insufficient data enhancement. In the case of very noisy data, large apertures usually are required for better signal accumulation, thus demanding usage of second-order terms for a more accurate description of the traveltimes. In practice, there is always a trade-off between the desired level of enhancement and the signal preservation properties of the algorithm. The maximum and minimum input dips and curvatures of the events strongly affect the computational cost of the algorithm and also can act as a filter enhancing or de-emphasizing specific events. In its simplest form, intervals for these parameters can be chosen with a wide safety margin encompassing the range of local dips and curvatures expected in real data. However, large search intervals may lead to the undesired consequence of enhancing other unwanted arrivals. Indeed, the NLBF algorithm has no notion of "signal" and "noise" and would enhance any locally coherent energy located inside the selected cones of parameters. As a result, careful choice of the search intervals should be done based on all available a priori information about the subsurface and using simplified formulas similar to equation A-5 or more advanced CRS or multifocusing approximations. The presented synthetic example suggests that relatively small search intervals targeting reflections after NMO correction would lead to the attenuation of steeper diffractions in the data. If diffractions are required for further processing, then larger dips and curvatures should be considered to preserve the diffraction energy. Diffraction preservation also is related to the conflicting dips problem because the diffracted energy usually is overlain by stronger reflected events. Although the presented

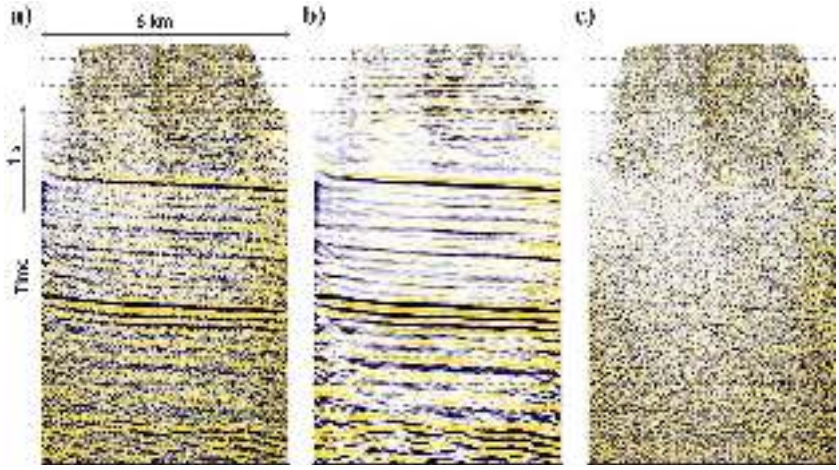

Figure 18. Stack sections from a land single-sensor survey (a) before and (b) after nonlinear beamforming in CSD and (c) the difference between the two.

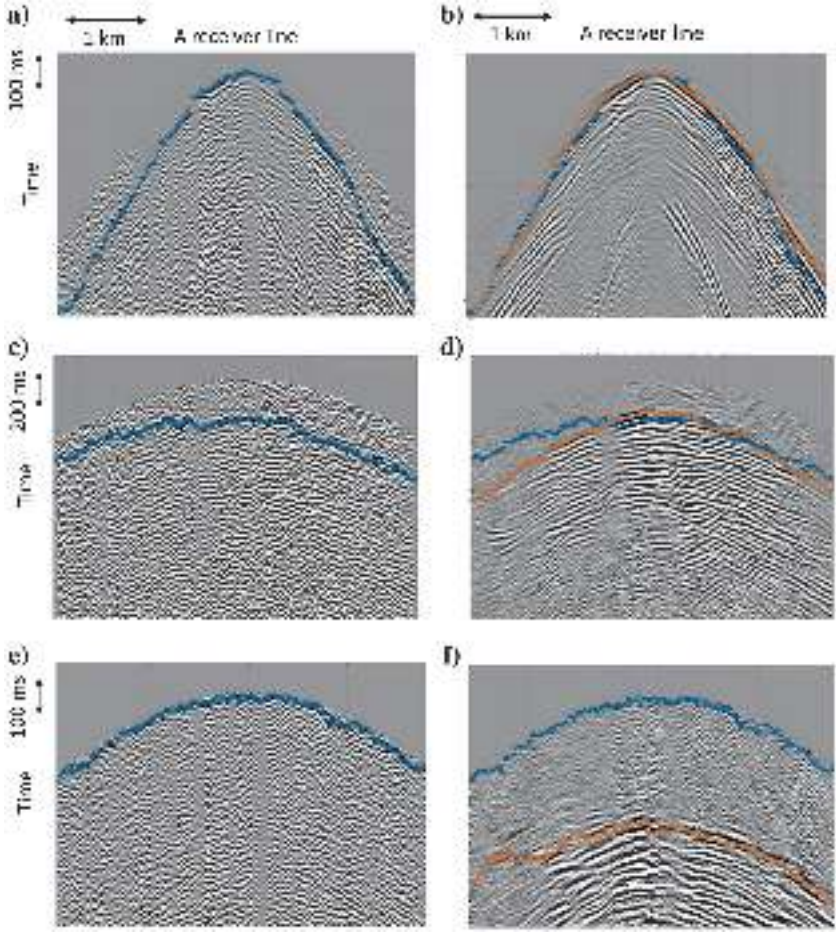

Figure 19. Common-shot gathers used for first-break picking on a land single-sensor survey (a, c, and e) before and (b, d, and f) after nonlinear beamforming conducted in CSD. The offset ranges are (a and b) 625–2395 m, (c and d) 2207–3197 m, and (e and f) 2747–3591 m. The automatic picking results before the enhancement are shown in blue and after the enhancement in red.

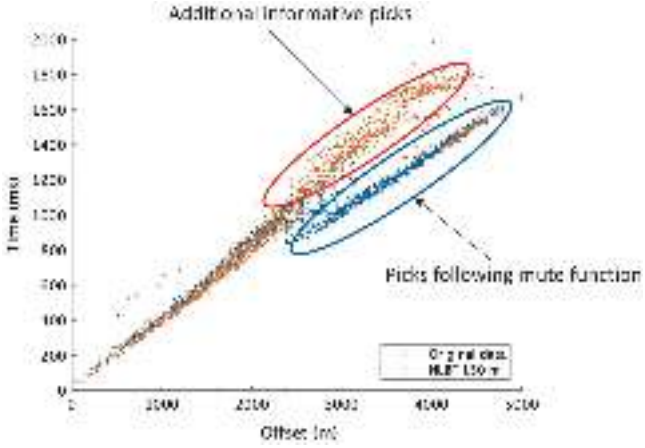

Figure 20. First-break picks obtained using an automatic picking algorithm on the original data (blue) merely follow the linear mute function at offsets larger than 2200 m. In contrast, picks after data enhancement using NLBF (red) appear more reliable and provide additional information on far offsets required to characterize the deeper subsurface.

version of the NLBF algorithm has some potential to resolve conflicting dips due to the multioperator stacking approach, strong reflections naturally dominate during coherency search, further de-emphasizing diffractions. This could be addressed by the more accurate design of input parameter intervals based on realistic traveltime approximations or by an iterative technique in which reflection enhancement is followed by an enhancement step focused on diffractions. For enhancement of first arrivals in the presence of a simple near-surface, application of linear moveout correction might be prudent to reduce search intervals significantly. However, with a complex near surface containing karsts, velocity inversions, or topography, large search intervals are required to capture highly variable moveouts accurately. The choice of input parameters ultimately relies on the end-user taking into account the framework in which the enhanced data will be used at subsequent processing steps. For example, in the context of full-waveform inversion, low-frequency data are enhanced with NLBF using larger summation apertures (Kim et al., 2019). Derivation of deconvolution operators or residual static solutions, however, might require significantly smaller apertures to avoid corrupting surface-consistent information in the data. These considerations lead to the development of adaptive multiscale processing (AMP) (Bakulin et al., 2019) in which, instead of prescribing single universal enhancement for all frequencies, we strive to achieve different levels of enhancement appropriate for each frequency band and as a result address the fundamental issue of frequency-dependent signal-to-noise typical for broadband acquisition. Here, AMP splits broadband data into multiple frequency bands and provides us with additional flexibility to apply larger summation apertures for enhancing the lower frequencies that demand them most and adjust to smaller apertures at higher frequencies. At low frequencies, this allows us to overcome the weakness of broadband sources (acquisition challenge), whereas, at medium and high frequencies, it purges near-surface diffractions (geologic challenge). Data in each band can be processed independently, applying scale-specific NLBF enhancement and recombining bands at the end for broadband imaging.

CONCLUSION

We present NLBF in CSD for enhancing challenging 3D prestack land data with orthogonal acquisition geometry. Important practical features of the algorithm are the ability to efficiently handle large volumes of high-density data and seamless incorporation into the conventional land processing workflow. Using a synthetic data example, we verify that NLBF possesses good signal preservation even for relatively large summation apertures required in practice. In the case of challenging single-sensor field data, the method allows us to successfully uncover reflected and refracted events on prestack gathers that previously exhibited only hints of coherent energy. By varying the local summation aperture, we could adjust the level of enhancement desired for a specific processing step. We anticipate that, with the current industry trend toward increasing trace density and decreasing the size of source and receiver arrays, NLBF will enable more efficient processing of challenging modern 3D high-channel-count and single-sensor data, ultimately leading to better land seismic imaging and more reliable reservoir characterization.

ACKNOWLEDGMENT

The authors would like to thank Evgeny Landa for the inspiring discussions and suggestions.

DATA AND MATERIALS AVAILABILITY

Data associated with this research are confidential and cannot be released.

APPENDIX A

DESCRIBING MOVEOUTS AFTER IMPERFECT NMO CORRECTION WITH SECOND-ORDER TRAVELTIME APPROXIMATION

In this appendix, we perform simplified analysis of the second-order traveltime approximation considered in the main text. Let us assume a case of a homogeneous medium with one horizontal reflector. Assuming that source x-coordinate x_s is equal to zero, receiver and source y-coordinates are identical, and considering traveltime moveout in X plane only, we can write familiar hyperbolic traveltime equation

$$t^2 = T_0^2 + \frac{x_r^2}{V_{\text{True}}^2}, \tag{A-1}$$

where T_0 is a two-way traveltime at zero offset, x_r is the x-coordinate of receiver, and V_{True} is a correct velocity above the reflector. The aim of an NMO correction is to move the observed traveltimes of the reflected waves to the corresponding zero-offset traveltime. Let us assume that we apply NMO correction with some velocity V_{NMO}; then, we obtain the following equation:

$$t_0^2 = t^2 - \frac{x_r^2}{V_{\text{NMO}}^2} = T_0^2 + x_r^2\left(\frac{1}{V_{\text{True}}^2} - \frac{1}{V_{\text{NMO}}^2}\right) = T_0^2 + x_r^2 K, \tag{A-2}$$

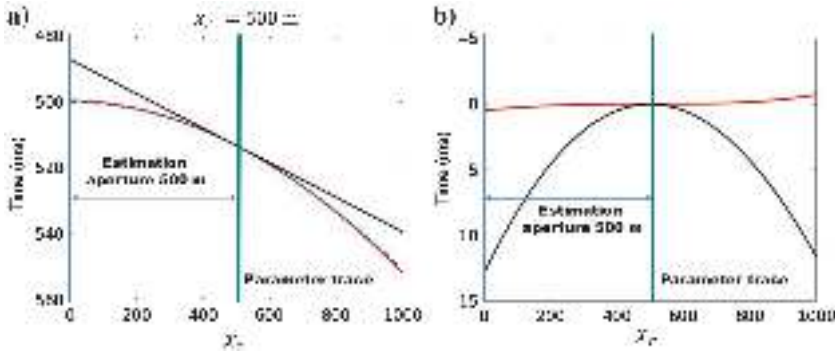

Figure A-1. (a) True residual moveout curve (dashed blue line) along with its first-order (black line) and second-order (red line) approximations at an offset $x_r = 500$ m and (b) the corresponding traveltime errors associated with these approximations. The residual moveout is computed in CSD after introducing NMO correction using incorrect velocity (10% error). Observe the much better fit provided by the second-order approximation compared to the linear one.

where t_0 is the actual traveltime after NMO correction and $K = (1/V_{\text{True}}^2 - 1/V_{\text{NMO}}^2)$ is the coefficient showing the accuracy of the NMO velocity. When $V_{\text{NMO}} = V_{\text{True}}$, we obtain a perfect flat event in the common-shot gather. When incorrect velocity is used, we shall observe some RMO:

$$t_0 = \sqrt{T_0^2 + x_r^2 K}. \tag{A-3}$$

Expanding this moveout in Taylor series in the vicinity of the receiver x_r^0, we obtain the following second-order approximation:

$$\begin{aligned} t_0(x_r^0 + \Delta x_r) &\approx t_0(x_r^0) + \frac{\partial t_0}{\partial x_r}\Delta x_r + \frac{1}{2}\frac{\partial^2 t_0}{\partial x_r^2}\Delta x_r^2 \\ &= t_0(x_r^0) + A\Delta x_r + D\Delta x_r^2, \end{aligned} \tag{A-4}$$

where coefficients A and D are precisely the same as in equation 2 and can be related as

$$A = \frac{K x_r^o}{t_0}, D = \frac{1}{2t_0}\left(K - \frac{K^2 x_r^o}{t_0^2}\right). \tag{A-5}$$

Let us consider a simple numerical example with a correct velocity $V_{\text{True}} = 1800$ m/s, zero-offset two-way traveltime $T_0 = 0.5$ s, and NMO velocity with a 10% error. Figure A-1a shows an exact RMO curve (equation A-3) and its approximation at the point of $x_r^0 = 500$ m using second-order expansion (equation A-4) and using first-order expansion when coefficient D is zeroed. The discrepancy between the true and the approximated traveltimes can be used to quantify errors in describing actual trajectories by used second-order approximation (Figure A-1b). As expected, the error is increasing with aperture and, at 500 m, it can reach 13 ms in case of the first-order approximation. The second-order approximation is much more accurate and provides an error of less than 1 ms. This straightforward illustrative example indicates that traveltime approximation could be acceptable for a chosen CSD even when NMO corrections are performed with imperfect initial velocities. Real moveouts can be distorted further by near-surface complexities or strong lateral variation of subsurface velocity. Therefore, the selection of summation apertures requires taking into account not only noise level in the data but also mentioned considerations along with the context of the specific workflow in which the enhanced data will be used.

APPENDIX B

FAST LOCAL SUMMATION ALGORITHM USING THE "PARAMETER-TRACE-ORIENTED" APPROACH

This appendix describes a computationally efficient algorithm we have developed to perform local summation of traces for the operator-oriented summation approach. First, let us estimate an approximate number of floating-point operations required to perform local summation for a whole data set by the straightforward implementation of equation 4. Inner summation includes calculation of terms $\Delta t(\mathbf{x}, \mathbf{x}^p)$ and $\Delta t(\hat{\mathbf{x}}, \mathbf{x}^p)$. Each of these terms demands $C_1 = 12$ operations (eight multiplications and four summations according to equation 2). Assume for simplicity that our stacking weights $w(\mathbf{x}, \mathbf{y})$ are equal to 1, an average number of seismic traces in the summation aperture is equal to $n$sum, and an average number of operators used in summation equals $n$op. Then for a fixed time sample $u(\mathbf{x}, t)$, one performs $n\text{op} \cdot n\text{sum} \cdot 2C_1$ operations during inner summation. Taking into account that we have nt time samples per trace and $n$traces seismic traces in the gather, the total number of floating-point operations is

$$n\text{traces} \cdot nt \cdot n\text{op} \cdot n\text{sum} \cdot 2C_1. \tag{B-1}$$

Fast summation achieves better computational performance by using extra memory for storing intermediate operations that are reused many times during beamforming for the efficiency of the numerical algorithm. We are continuing to perform the same summation as in equation 4, but we rearrange the terms to reduce the number of computations.

Let us assume that, for all traces within the estimation aperture of a fixed parameter trace with coordinate $\mathbf{x}^p$, we calculate all possible moveouts $\Delta t(\mathbf{x}, \mathbf{x}^p)$ (where $\mathbf{x} \in \Pi_{\text{est}}$) and store them in memory. The number of floating-point operations required for such computations is

$$nt \cdot n\text{est} \cdot C_1, \tag{B-2}$$

where nest is the number of traces in the estimation aperture. To efficiently use these beforehand calculated moveouts during the summation, we use a partial summation algorithm (see Table B-1). The first inner loop runs over all traces $\mathbf{x}_1$ within estimation aperture of a parameter trace $\mathbf{x}^p$. For each trace $\mathbf{x}_1$, we do partial summation: nested loop over trace $\mathbf{x}_2$ from the estimation aperture of a given parameter trace and within summation aperture of $\mathbf{x}_1$. The number of floating-point operations required for partial summation is

Table B-1. Partial summation loop of the proposed computationally efficient fast summation algorithm based on a parameter-trace-oriented approach.

```
loop over target trace x1 ∈ Πest
  if x1 ∈ Πop
    loop over input traces x2 ∈ Πest
      if x2 within summation aperture of x1
        do partial summation for x1 following formula 4
      endif
    end loop
  endif
end loop
```

$$nt \cdot n\text{sum} \cdot n\text{sum} \cdot C_1. \tag{B-3}$$

As one may recognize, we do absolutely the same arithmetic operations as in equation 4 but in a different order. Combining equations B-2 and B-3 for all parameter traces, we obtain the total number of operations in the fast summation approach as

$$n\text{op} \cdot (nt \cdot n\text{est} \cdot C_1 + nt \cdot n\text{sum} \cdot n\text{sum} \cdot C_1). \tag{B-4}$$

Dividing the number of operations in a straightforward implementation (equation B-1) by the number of operations in the proposed one (equation B-4) and making simplifications, we obtain the following estimate of the performance speedup coefficient:

$$K = \frac{n\text{traces} \cdot nt \cdot n\text{op} \cdot n\text{sum} \cdot 2C_1}{n\text{op} \cdot (nt \cdot n\text{est} \cdot C_1 + nt \cdot n\text{sum} \cdot n\text{sum} \cdot C_1)} \geq \frac{n\text{traces}}{n\text{sum}}. \tag{B-5}$$

The number of traces inside summation ensemble $n$sum is relatively small and depends on the acquisition density and the size of the summation aperture. In contrast, $n$traces is the total number of traces in the seismic gather and is usually several orders of magnitude larger than $n$sum. This fact results in $K \gg 1$ and confirms the superior performance of the proposed fast summation implementation over a straightforward summation algorithm.

REFERENCES

Bagaini, C., T. Bunting, A. El-Emam, A. Laake, and C. Strobbia, 2010, Land seismic techniques for high-quality data: Oilfield Review, **22**, 28–39.

Bakulin, A., and K. Erickson, 2017, Enhance-estimate-image: New processing approach for single-sensor and other seismic data with low prestack signal-to-noise ratio: 87th Annual International Meeting, SEG, Expanded Abstracts, 5001–5005, doi: 10.1190/segam2017-17495425.1.

Bakulin, A., P. Golikov, M. Dmitriev, D. Neklyudov, P. Leger, and V. Dolgov, 2018b, Application of supergrouping to enhance 3D prestack seismic data from a desert environment: The Leading Edge, **37**, 200–207, doi: 10.1190/tle37030200.1.

Bakulin, A., I. Silvestrov, and M. Dmitriev, 2019, Adaptive multiscale processing of challenging 3D seismic data for first-break picking, FWI and imaging: 89th Annual International Meeting, SEG, Expanded Abstracts, 3979–3983, doi: 10.1190/segam2019-3214616.1.

Bakulin, A., I. Silvestrov, M. Dmitriev, D. Neklyudov, M. Protasov, K. Gadylshin, V. Tcheverda, and V. Dolgov, 2018a, Nonlinear beamforming for enhancing pre-stack data with challenging near surface or overburden: First Break, **36**, 121–126.

Baykulov, M., and D. Gajewski, 2009, Prestack seismic data enhancement with partial common-reflection-surface (CRS) stack: Geophysics, **74**, no. 3, V49–V58, doi: 10.1190/1.3106182.

Berkovitch, A., K. Deev, and E. Landa, 2011, How non-hyperbolic multifocusing improves depth imaging: First Break, **27**, 95–103.

Buzlukov, V., R. Baina, and E. Landa, 2010, Prestack data enhancement using local traveltime approximation: 72nd Annual International Conference and Exhibition, EAGE, Extended Abstracts, C011.

Buzlukov, V., and E. Landa, 2013, Imaging improvement by prestack signal enhancement: Geophysical Prospecting, **61**, 1150–1158, doi: 10.1111/1365-2478.12047.

Cooper, N., 2004, A world of reality — Designing land 3D programs for signal, noise, and prestack migration: The Leading Edge, **23**, 1007–1014, doi: 10.1190/1.1813357.

Cordery, S., 2018, Seismic processing workflow for broadband single-sensor single-source land seismic data: U.S. Patent Application 20180284307.

Curia, D., P. Borghi, J. Noble, A. Berkovitch, D. Justo, and M. Alayón, 2017, The impact of multifocusing in the processing of land 3D seismic data in a fold and thrust belt setting: Ranquil Norte Block, Neuquén Basin, Argentina: The Leading Edge, **36**, 770–774, doi: 10.1190/tle36090770.1.

Faccipieri, J. H., T. A. Coimbra, L.-J. Gelius, and M. Tygel, 2016, Stacking apertures and estimation strategies for reflection and diffraction enhancement: Geophysics, **81**, no. 4, V271–V282, doi: 10.1190/geo2015-0525.1.

Gülünay, N., and N. Benjamin, 2008, Poststack driven prestack deconvolution (PPDEC) for noisy land data and radial trace mixing for signal enhancement: 78th Annual International Meeting, SEG, Expanded Abstracts, 2507–2511, doi: 10.1190/1.3063864.

Hoecht, G., P. Ricarte, S. Bergler, and E. Landa, 2009, Operator-oriented interpolation: Geophysical Prospecting, **57**, 957–979, doi: 10.1111/j.1365-2478.2009.00789.x.

Ji, Y., E. Kragh, and C. Bagaini, 2010, Noise attenuation methods for point-receiver land seismic data: 80th Annual International Meeting, SEG, Expanded Abstracts, 3545–3549, doi: 10.1190/1.3513586.

Kim, Y. S., T. Fei, M. Dmitriev, and Y. Luo, 2019, An offshore full waveform inversion with automatic salt flooding: An OBN case study: 81st Annual International Conference and Exhibition, EAGE, Extended Abstracts, WeR0812.

Li, D., Z. Li, and X. Sun, 2011, Improving the quality of prestack seismic data with the CO CRS stacking method: 81st Annual International Meeting, SEG, Expanded Abstracts, 3648–3652, doi: 10.1190/1.3627958.

Mann, J., 2001, Common-reflection-surface stack and conflicting dips: 63rd Annual International Conference and Exhibition, EAGE, Extended Abstracts, P077.

Mann, J., 2002, Extensions and applications of the common-reflection-surface stack method: Ph.D. thesis, University of Karlsruhe.

Müller, N.-A., 2009, Treatment of conflicting dips in the 3D common reflection-surface stack: Geophysical Prospecting, **57**, 981–995, doi: 10.1111/j.1365-2478.2009.00803.x.

Müller, N.-A., M. Brucher, and F. Clement, 2013, 3D common-offset CRS stack: Simplified formulation and application to Foothills data: 83rd Annual International Meeting, SEG, Expanded Abstracts, 4262–4266, doi: 10.1190/segam2013-0115.1.

Özbek, A., 2000, Multichannel adaptive interference cancelling: 70th Annual International Meeting, SEG, Expanded Abstracts, 2088–2091, doi: 10.1190/1.1815857.

Oristaglio, M., 2015, SEAM update: The Leading Edge, **34**, 466–468, doi: 10.1190/tle34040466.1.

Pecholcs, P. I., R. Al-Saad, M. Al-Sannaa, J. Quigley, C. Bagaini, A. Zarkhidze, R. May, M. Guellili, S. Sinanaj, and M. Membrouk, 2012, A broadband full azimuth land seismic case study from Saudi Arabia using a 100,000 channel recording system at 6 terabytes per day: Acquisition and processing lessons learned: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: 10.1190/segam2012-0438.1.

Rauch-Davies, M., A. Berkovitch, K. Deev, and E. Landa, 2013, Non-hyperbolic multifocusing imaging for pre-stack signal enhancement: 83rd Annual International Meeting, SEG, Expanded Abstracts, 4618–4622, doi: 10.1190/segam2013-0788.1.

Regone, C., M. Fry, and J. Etgen, 2015, Dense sources vs. dense receivers in the presence of coherent noise: A land modeling study: 85th Annual International Meeting, SEG, Expanded Abstracts, 12–16, doi: 10.1190/segam2015-5833924.1.

Spinner, M., C. Tomas, P. Marchetti, C. Gallo, and S. Arfeen, 2012, Common-offset CRS for advanced imaging in complex geological settings: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: http://dx.doi.org10.1190/segam2012-1099.1.

Vermeer, G. J. O., 2012, 3D seismic survey design: SEG.

Xie, Y., 2017, 3D prestack data enhancement with a simplified CO CRS: 79th Annual International Conference and Exhibition, EAGE, Extended Abstracts, WeP713.

Xie, Y., and D. Gajewski, 2017, 5-D interpolation with wave-front attributes: Geophysical Journal International, **211**, 897–919, doi: 10.1093/gji/ggx334.