OXFORD

Journal of Geophysics and Engineering

Journal of Geophysics and Engineering (2021) **18**, 890–907 https://doi.org/10.1093/jge/gxab060

Evaluating strategies for estimation of local kinematic parameters in noisy land data: quality versus performance trade-offs

Andrey Bakulin[1,*], **Ilya Silvestrov**[1] **and Maxim Protasov**[2]

[1] EXPEC Advanced Research Center, Geophysics Technology, Saudi Aramco, 31311, Dhahran, Saudi Arabia
[2] Laboratory of Multiwave Seismic, Institute of Petroleum Geology and Geophysics SB RAS, 630090, Novosibirsk, Russia

*Corresponding author: Andrey Bakulin. E-mail: Andrey.Bakulin@aramco.com

Received 1 April 2021, revised 2 October 2021
Accepted for publication 20 October 2021

Abstract

Modern land seismic data are typically acquired using high spatial trace density with small source and receiver arrays or point sources and sensors. These datasets are challenging to process due to their massive size and relatively low signal-to-noise ratio caused by scattered near-surface noise. Therefore, prestack data enhancement becomes a critical step in the processing flow. Nonlinear beamforming had proved very powerful for 3D land data. However, it requires computationally intensive estimations of local coherency on dense spatial/temporal grids in 3D prestack data cubes. We present an analysis of various estimation methods focusing on a trade-off between computational efficiency and enhanced data quality. We demonstrate that the popular sequential «2 + 2 + 1» scheme is highly efficient but may lead to unreliable estimation and poor enhancement for data with a low signal-to-noise ratio. We propose an alternative algorithm called «dip + curvatures» that remains stable for such challenging data. We supplement the new strategy with an additional interpolation procedure in spatial and time dimensions to reduce the computational cost. We demonstrate that the «dip + curvatures» strategy coupled with an interpolation scheme approaches the «2 + 2 + 1» method's efficiency while it significantly outperforms it in enhanced data quality. We conclude that the new algorithm strikes a practical trade-off between the performance of the algorithm and the quality of the enhanced data. These conclusions are supported by synthetic and real 3D land seismic data from challenging desert environments with complex near surface.

Keywords: data enhancement, noisy seismic data, kinematic parameters estimation, beamforming

1. Introduction

Improving land seismic data quality has been an active field of research over the last several decades. A conventional and well-established way to obtain records with a good signal-to-noise ratio is to use source and receiver arrays directly during field acquisition. Arrays enhance weak reflection signals with high apparent velocity while suppressing strong near-surface arrivals with low speeds. With an increasing demand for more accurate seismic imaging for subtler low-relief structures and stratigraphic traps, the seismic industry moves toward higher density surveys expected to improve the resolution and information content of the data (Bagaini *et al.* 2010; Bakulin *et al.* 2018; Cordery 2020). Finer inline and cross-line samplings both for shots and receivers are achieved at the expense of decreasing or eliminating the field arrays. The ultimate goal

is to use geometries with very fine spacing intervals in all directions together with point source and point-receiver systems. It is acknowledged that the quality of a point-receiver data is degraded in such a case. However, the tremendously increased fold with advanced processing techniques is expected to compensate for this reduced quality and provide better images due to the preservation of tiny signals during recording and their summation in final migration. Since acquisitions with fine sampling in all directions remain out of reach for economic reasons, field arrays' simple mimicking during the data-processing stage is impossible. This leads to a requirement for new denoising and data-enhancement techniques that can deal with extremely noisy and relatively coarsely sampled data in a more advanced manner than the simple grouping of traces used before.

Different methods were developed in the past to attenuate coherent and incoherent noise in seismic records. Among them are classical prediction-filtering methods (Canales 1984; Gulunay 1986; Abma & Claerbout 1995; Wang 1999), various optimum-weighted stacking approaches (Robinson 1970; Rashed 2014), several filtering techniques in *f-x* and *f-k* domains (Stewart & Schieck 1989; Duncan & Beresford 1994; Naghizadeh 2012; Naghizadeh & Sacchi 2012), plane-wave destruction filtering (Fomel 2002), methods based on slant stacking and Radon transform (Kelamis & Mitchell 1989; Trad *et al.* 2003; Ibrahim & Sachi 2014), techniques using other sparse representations of seismic data including wavelets, seislets and curvelets (Herrmann & Hennenfent 2008; Fomel & Liu 2010; Mousavi *et al.* 2016), rank-reduction-based methods (Ulrych *et al.* 1988; Oropeza & Sacchi 2011; Chen & Sacchi 2015) and recent machine-learning-based algorithms (Gadylshin *et al.* 2020; Saad & Chen 2020; Gao *et al.* 2021).

Coherency-based data-enhancement techniques are widely used in the poststack domain (Marfut *et al.* 1998). They can be applied to improve the signal level in prestack data in pre- and postmigrated domains. The main techniques from this category are based on a local slant stacking of the data and selection of the most coherent components. More advanced approaches, coming mainly from the common-reflection surface (CRS) (Zhang *et al.* 2001; Baykulov & Gajewski 2009) and multifocusing theory (Berkovitch *et al.* 2011), use second-order approximations of the wavefront. This allows one to describe the events' kinematical trajectories better and stack locally along them to increase SNR. The moveout can be characterised by a global operator or a local one. The local operators' methods appear most flexible for enhancing challenging 3D land data where heterogeneity and static issues often invalidate any global moveout behavior. At the same time, these methods rely on intensive numerical search for optimal coherency over the entire 5D data domain, making them computationally expensive. By representing a traveltime surface locally as a general second-order curve (Hoecht *et al.* 2009; Buzlukov & Landa 2013), a nonlinear beamforming algorithm was introduced for enhancing massive challenging land datasets (Bakulin *et al.* 2020).

Different strategies for evaluating event coherency have been proposed in the past. One of the early optimisation techniques for CRS parameter search was based on a sequential refinement proposed by Jager *et al.* (2001) and Mann (2002). The following methods evolved to enhance and optimise the parameter search starting from simple two-step parameter refinement for 2D data and finishing with a global optimisation for 3D data by Garabito & Cruz (2019). Likewise, since early work on the multifocusing method (Berkovitch *et al.* 1994; Gelchinsky *et al.* 1999), there was steady progress in estimating kinematic parameters. The latest publications are devoted to global optimisation (Fam & Naghizadeh 2019).

This study discusses different estimation strategies with the application to the nonlinear beamforming (NLBF) algorithm. We provide insights into these strategies' quality and performance trade-offs in the context of 3D land seismic processing of massive datasets with petabytes of data. We present an optimised version of the algorithm to enhance 3D single-sensor data that uses a novel scheme based on the sequential estimation of dips and curvatures of kinematic surfaces on sparse grids with the subsequent interpolation to a dense grid. We provide a comparison based on synthetic and real-data examples showing the advantages of the proposed strategy compared to previously known schemes. A synthetic data example provides a direct assessment of the effect of variable signal-to-noise ratio on the accuracy of estimated kinematic parameters. We benchmark various approaches against the most comprehensive algorithm of global 5D optimisation on a real-data example.

2. Method

The main idea of the nonlinear beamforming method is to describe traveltime moveout locally as a second-order surface, estimate its parameters and perform local summation along this local moveout to improve a signal-to-noise ratio. Considering a data space with a coordinate vector $\vec{x} = (x_s, y_s, x_r, y_r)$ defined by source and receiver x and y coordinates, the traveltime t, can be locally represented using a Taylor series expansion as:

$$t(\vec{x}_0 + \Delta\vec{x}) = t(\vec{x}_0) + \vec{L}^T\Delta\vec{x} + \Delta\vec{x}^T N \Delta\vec{x}, \qquad (1)$$

where $\vec{L}$ is a first-derivative gradient vector, and N is a matrix of second derivatives. In total, 14 unknown coefficients of the vector $\vec{L}$ and matrix N define the local traveltime surface at a current position in 5D seismic volume. Estimating all these kinematic parameters is too costly from the

computational point of view. Simplification of equation (1) is usually required. In the current work, we fix two arbitrary directions in the data space and consider sections of the traveltime surface along the two remaining directions only. This reduces the number of unknown parameters to five. We estimate these local kinematic parameters by scanning different trajectories and finding one with the best coherency defined by a semblance function's maximum value.

2.1. Estimation of kinematic parameters

Considering only two variable spatial coordinates and simplifying equation (1), we obtain the following second-order traveltime approximation in a 2D plane with respect to some reference point (x_0, y_0):

$$\begin{aligned} t(x,y) = t(x_0,y_0) + A\cdot(x-x_0) + B\cdot(y-y_0) \\ + C\cdot(x-x_0)(y-y_0) + D\cdot(x-x_0)^2 \\ + E\cdot(y-y_0)^2, \end{aligned} \tag{2}$$

where x and y represent two selected coordinates in the data domain. A, B, C, D and E are five unknown kinematic parameters to be estimated. The estimation can be done by maximising semblance or other coherency measures in multi-dimensional parameter space. In this framework, different estimation strategies are possible. In the following, we will consider some of them in more detail.

2.1.1. 5D brute force strategy. If we have a wavefield $U(x, y; t)$ then using the notation from equation (2), a standard definition of semblance function can be written as follows:

$$S = \frac{\sum_{j=1}^{N}\left\{\sum_{i=1}^{M} U(x_i,y_i;t_j(x_i,y_i))\right\}^2}{M\sum_j^N\sum_i^M U(x_i,y_i;t_j(x_i,y_i))^2}, \tag{3}$$

where M is a number of traces within a spatial estimation window, N is a number of the time samples in a time window and moveout $t(x_i, y_i)$ is calculated using equation (2). A straightforward brute force strategy consists of probing all possible values of kinematic parameters A, B, C, D, E within predefined limits and choosing those values that provide the maximum of the semblance function. More sophisticated approaches based on global search methods (Garabito & Cruz 2019) can also be adapted into this scheme to improve the algorithm's performance. In this work, we use the full 5D strategy as a reference for the results' quality, i.e. we consider the 5D brute force search the best possible solution.

2.1.2. «2 + 2 + 1» strategy. The so-called «2 + 2 + 1» strategy consists of a sequential search first for two coefficients A and D, then for B and E and then for a remaining coefficient C, with the first four being fixed (Bakulin *et al.* 2020). In more detail, the algorithm can be written as follows:

(i) First, we fixed the value y_0 on a so-called estimation grid, where the parameters are to be estimated. Taylor series, in this case, is defined by two parameters only:

$$t(x,y_0) = t(x_0,y_0) + A\cdot(x-x_0) + D\cdot(x-x_0)^2. \tag{4}$$

So, in the direction of x-coordinate, we have to solve the optimisation problem of defining only two parameters A and D. Only a subset of data ensemble is used in this step, which may influence estimation robustness, as we show later. Since the number of parameters is decreased by three, and the data dimension is reduced by one, the total computation cost for this new problem is roughly four orders of magnitude less than the full 5D problem.

(ii) Second, we fix another direction, i.e. value x_0 and use a similarly sparse estimation grid. Taylor series, in this case, is defined by two other parameters:

$$t(x_0,y) = t(x_0,y_0) + B\cdot(y-y_0) + E\cdot(y-y_0)^2. \tag{5}$$

So in the direction of y-coordinate, we have to solve the optimisation problem of searching for E and D parameters. The computational cost of this step is similar to the first one.

(iii) Finally, after fixing parameters A, B, D, E estimated in point (x_0, y_0), we search for the remaining coefficient C using equation (2). A full 2D patch of data extracted from the vicinity of (x_0, y_0) is used. In this case, the data dimension is more extensive since we use an entire data ensemble but only estimate a single parameter C. As a result, the computational cost is similar to the first and second steps. Overall, the «2 + 2 + 1» scheme gives us a more efficient algorithm compared to the 5D brute force solution, with a computational cost being less by roughly four orders of magnitude because we reduced search space from 5D to 2D (three orders), and we reduced the number of traces used for estimation in the vicinity of each point (x_0, y_0) by approximately one order (from 2D ensemble to 1D line).

2.1.3. Objective functions for «2 + 2 + 1» strategy vs. 5D brute force strategy. Apart from the optimisation method itself, there is also an essential difference in objective function computation between the «2 + 2 + 1» and 5D approaches. Figure 1 shows a typical ensemble structure selected from a field dataset. The 5D strategy uses an entire ensemble of traces inside the red rectangle for estimation. In contrast, the sequential «2 + 2 + 1» strategy uses only an oriented strip of several lines along each coordinate direction denoted

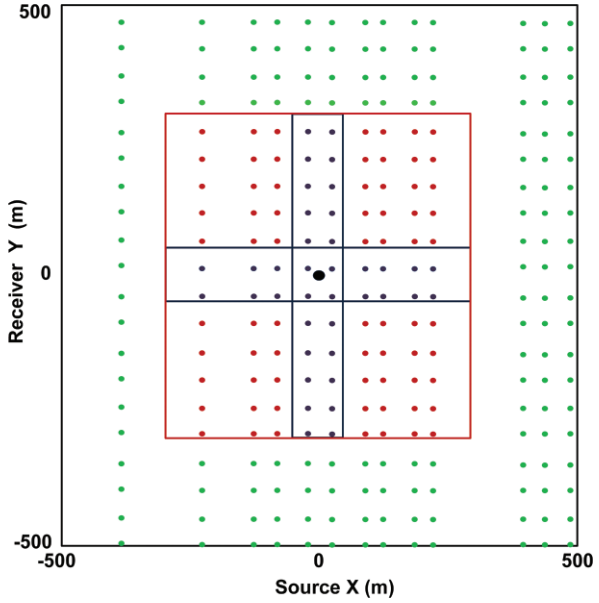

Figure 1. Distribution of the traces for a single ensemble in a cross-spread gather, shown for a field dataset used in this study. Each dot denotes a location of seismic traces with relative source X and receiver Y coordinates (origin (0,0) corresponds to the geometric center of the cross-spread). The red square encloses all traces of the single ensemble used for nonlinear beamforming. All these traces contribute during summation to produce enhanced output trace in the middle (black dot). In the 5D brute force strategy, all ensemble traces are also used for parameters estimation. In contrast, the «2 + 2 + 1» strategy only uses traces bounded by blue rectangles oriented along the x or y axes during the first and second estimation steps, respectively. In the presence of noise, the data subset's size may influence the estimated parameters.

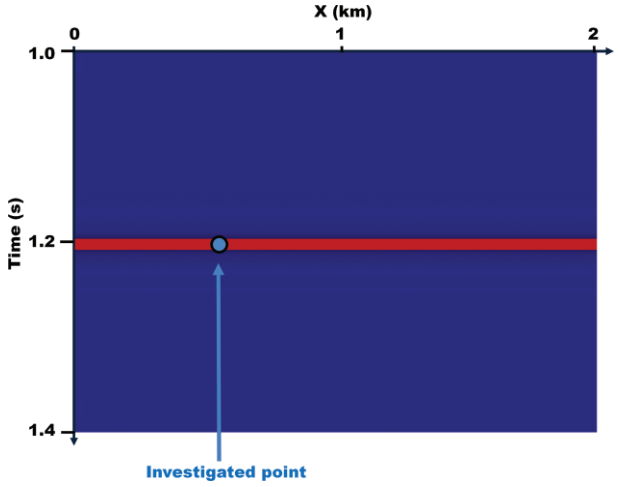

Figure 2. Synthetic data with the planar event and the investigated point on the event.

by blue rectangles. In an ideal case, a single well-sampled line would be sufficient to estimate dips/curvatures in each vertical plane x-t or y-t. However, each strip is designed to contain 2–3 central lines for field data, essentially forming a 'fat' line to deal with geometry imperfections, missing traces and noise. These choices, along with the data quality, may affect the estimation process. Let us demonstrate this point by analysing the semblance-based objective function's cross-sections using synthetic and real field data.

First, let us consider a simple plane event from synthetic data (figure 2) with zero dips and curvature values, i.e. $A = B = C = D = E = 0$. Figure 3 shows slices of the semblance from equation (3) depending on the chosen optimisation strategy. The first column presents 2D cross-sections of the objective function for «2 + 2 + 1» strategy along planes (A, B) and (D, E) and 1D cross-section along the direction of parameter C. Cross-sections are produced while fixing remaining parameters to their optimal values. For the noise-free synthetic event, picking the maximum semblance value along each transection provides a good approximation of the actual kinematic trajectory, i.e. estimated $A = B = C = D = E = 0$. Recall that the semblance objective function in equation (3) uses summation over a limited subset or fat lines (figure 1). The middle column (figure 3d–f) shows the same cross-sections but with the full objective function (1) used for the 5D strategy, using all ensemble traces instead of narrow strips in figure 1. We stress again that the middle column uses a different objective function. However, cross-sections are done at the same parameter planes as estimated by the original «2 + 2 + 1» strategy from figure 3a–c. Finally, the right column (figure 3g–i) shows cross-sections of the same full 5D objective function but dissected through the values of optimal parameters estimated from the 5D brute force strategy. For synthetic data with an ideal signal and absence of noise, both strategies estimate correct parameters. Besides, cross-sections of objective function exhibit very similar behavior (figure 3). This synthetic example validates a simplified sequential «2 + 2 + 1» strategy widely used in various approaches (Mann 2002; Buzlukov & Landa 2013) as a computational timesaver.

Now let us perform a similar analysis of objective function using real data with significant noise. First, we select a point of interest located along with a relatively coherent and robust event at the near offset (blue dot in figure 4). We calculate similar cross-sections as was done in the synthetic example (figure 5). Cross-sections for the left and center columns are comparable, suggesting that objective functions are similarly shaped for «2 + 2 + 1» (figure 5a–c) and 5D brute force strategy (figure 5d–f) when dissections are taken along the optimal parameters estimated by «2 + 2 + 1» approach. Although the 5D brute force approach (figure 5g–i) provides more focused and less oscillating results for the (A, D) pair and C value, the actual estimated parameters between 5D and «2 + 2 + 1» approaches are similar.

We chose a point in a noisier area at a far offset in a final test, where the signal-to-noise ratio is low (red dot in figure 4). In contrast to previous examples, we see very different behavior of the objective functions from the «2 + 2 + 1» and the 5D approaches (figure 6). Differences between the left (figure 2a–c) and

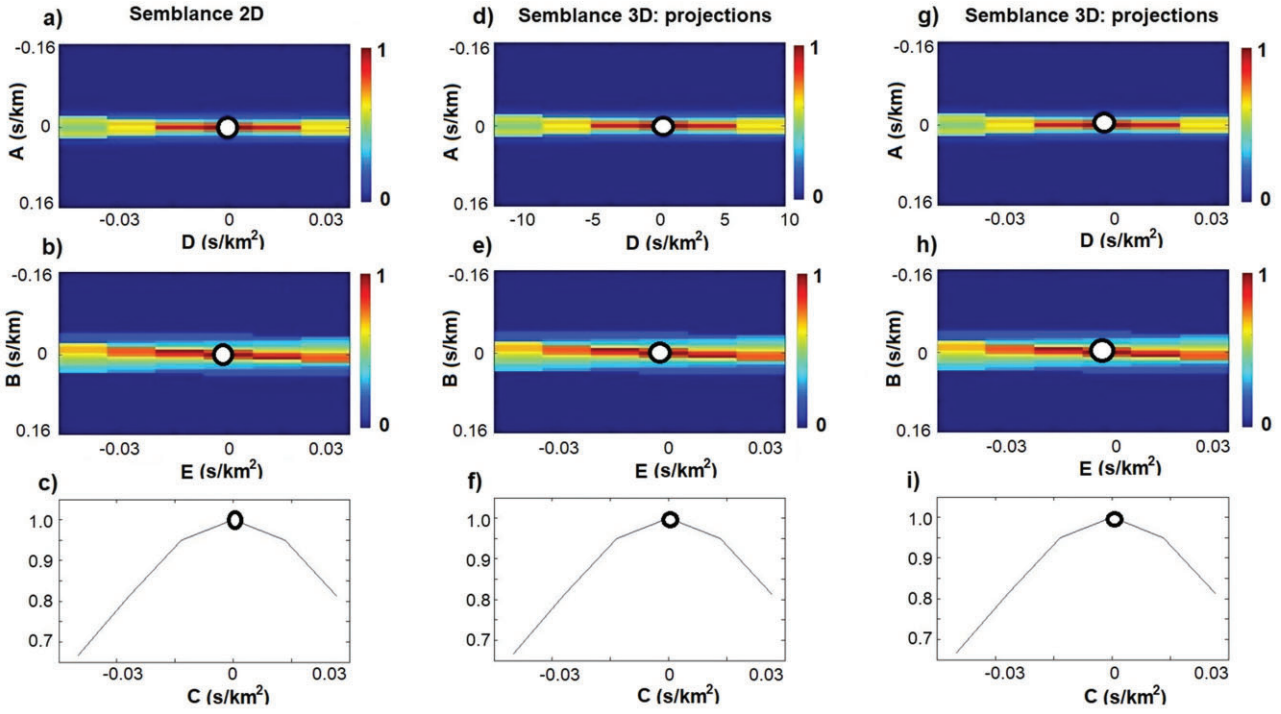

Figure 3. Cross-sections of the objective function for a synthetic dataset consisting of a single planar event. The left (a, b, c) column of images shows cross-sections of the «2 + 2 + 1» objective function (data from 'fat' lines parallel to each coordinate axis as shown in figure 1). Cross-sections are taken through the optimal values of the parameters as obtained with the «2 + 2 + 1» strategy. From top to bottom, each display (a, b, c) shows the first, second and third steps of the conventional «2 + 2 + 1» approach correspondingly. The middle column (d, e, f) shows the 5D objective function's slices but is still drawn through the same estimated values from the «2 + 2 + 1» approach as in the left column. Finally, the right column (g, h, i) shows slices of 5D objective function but drawn through the optimal parameters estimated with the 5D brute force approach. White dots denote the optimal values of the parameters for each case.

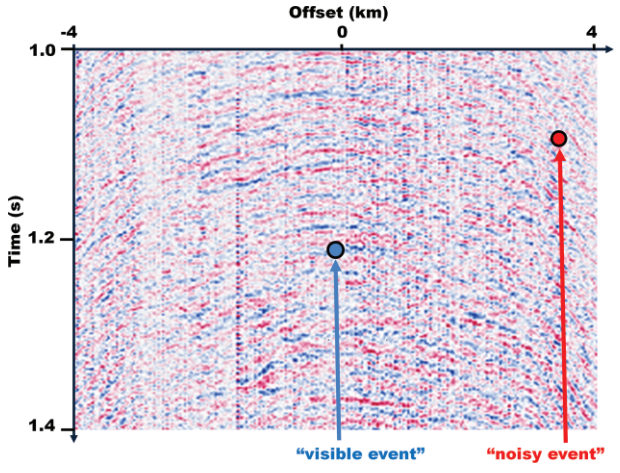

Figure 4. Real data and two investigated points located on the 'visible event' and the 'noisy event'.

right (figure 6g–i) columns are not unexpected in case of complex objective function. The sequential «2 + 2 + 1» approximation may fail to capture the objective function complexity in the presence of noise. However, differences between the left (figure 6a–c) and the middle (figure 6d–f) columns confirm that the data ensemble's size is a significant culprit. The bigger the ensemble, the more resilient the estimation approach is to noise. The mere fact that cross-sections change shape for «2 + 2 + 1» and 5D objective functions (compare figure 6a–c and figure 6d–f) suggest that despite forming 'fat' lines in «2 + 2 + 1» approach (figure 1), the limited size of the data subsets still leads to less reliable estimation. Therefore, we might expect differences in the estimation results between the two strategies, especially on the noisy data (figure 7). Similar to the coordinate descent method, additional iteration strategies may improve «2 + 2 + 1» schemes. Nevertheless, usage of an undecimated data ensemble (e.g. full 2D data patch) appears the most reliable approach to overcome the effects of intense noise on the estimation. This study shall propose a new alternative strategy that uses an entire undecimated ensemble to estimate all five parameters but sequentially in contrast to the 5D brute force solution. It will maintain the computational efficiency of sequential optimisation while mitigating noise effects similar to the 5D approach.

2.1.4. Sequential «dips + curvatures» strategy. We propose an alternative estimation strategy for kinematic parameters comprising of two sequential steps. During the first step,

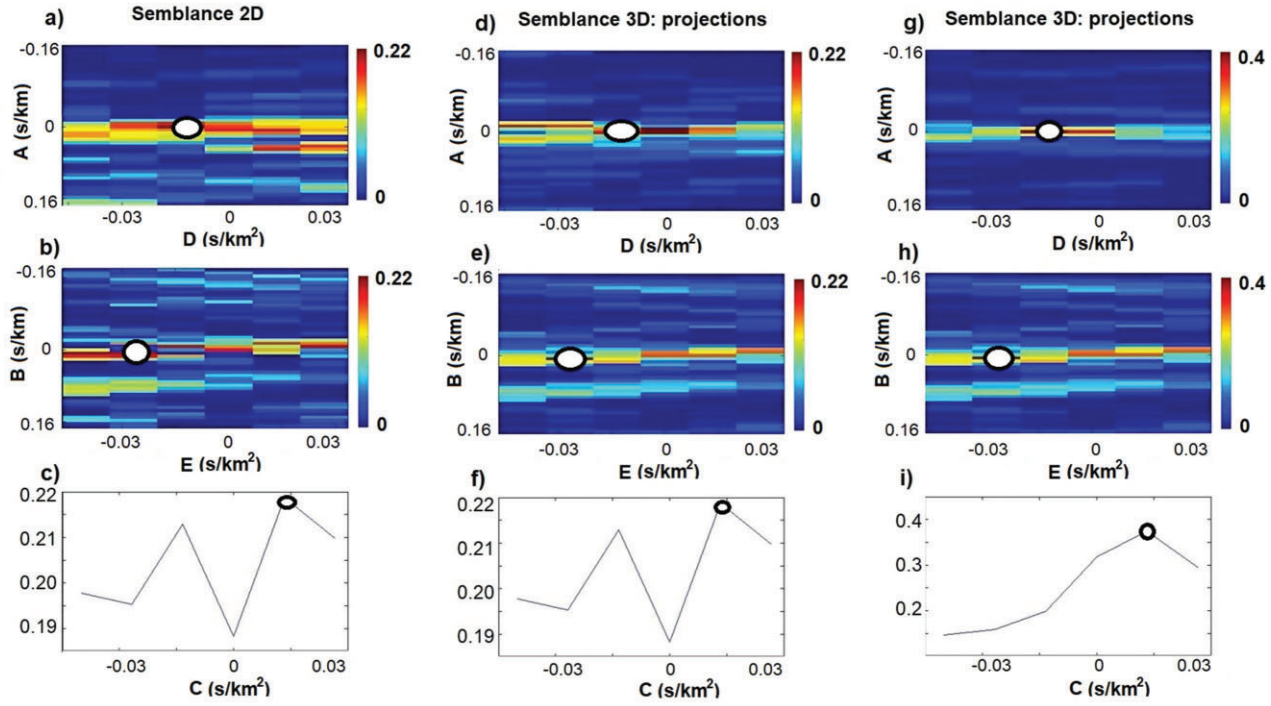

Figure 5. Same as figure 3, but for the 'visible event' from real data (figure 4a): (a, b, c) show, correspondingly, the first, second and the last steps of the «2 + 2 + 1» approach (left); (d, e, f) show cross-sections of 5D objective function at the estimated values from the «2 + 2 + 1» approach (middle) and (g, h, i) show cross-sections of 5D objective functions at the estimated values from the 5D brute force approach (right).

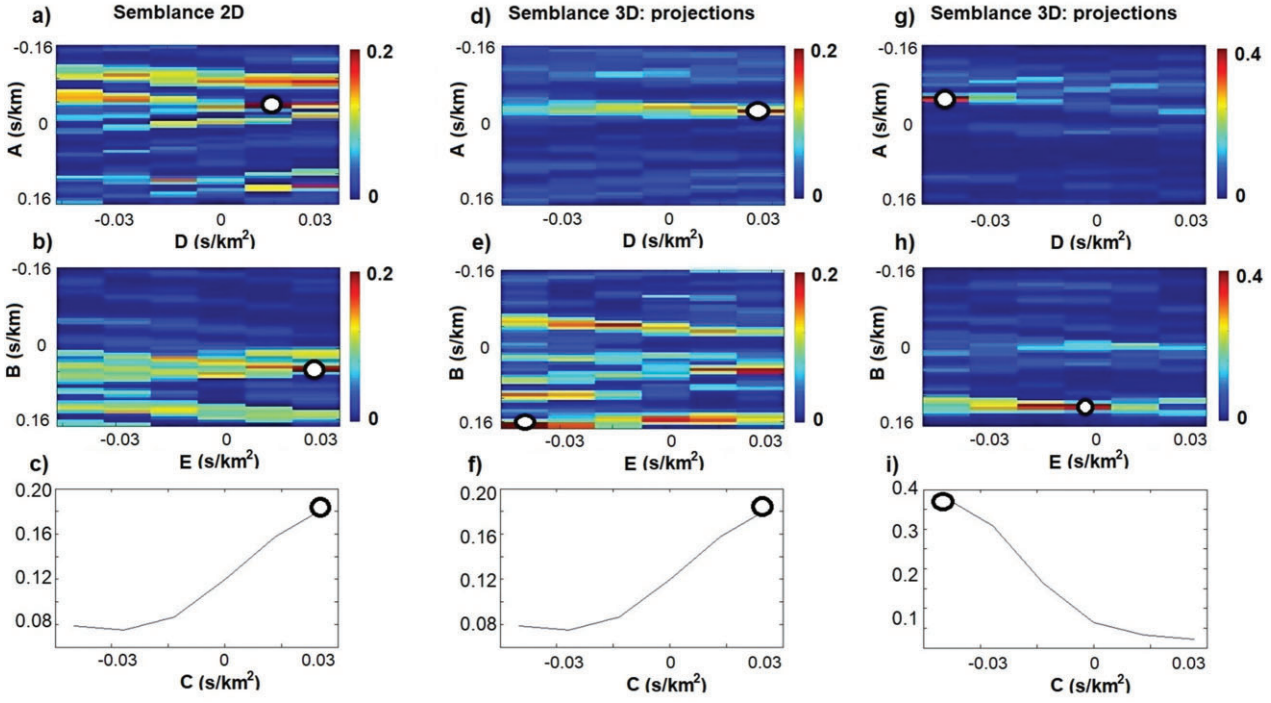

Figure 6. Same as figure 5, but for the 'noisy event' from real data (figure 4b): (a, b, c) show, correspondingly, the first, second and the last steps of the «2 + 2 + 1» approach (left); (d, e, f) show cross-sections of 5D objective function at the estimated values from the «2 + 2 + 1» approach and (g, h, i) show cross-sections of 5D objective functions at the estimated values from the 5D brute force approach (right).

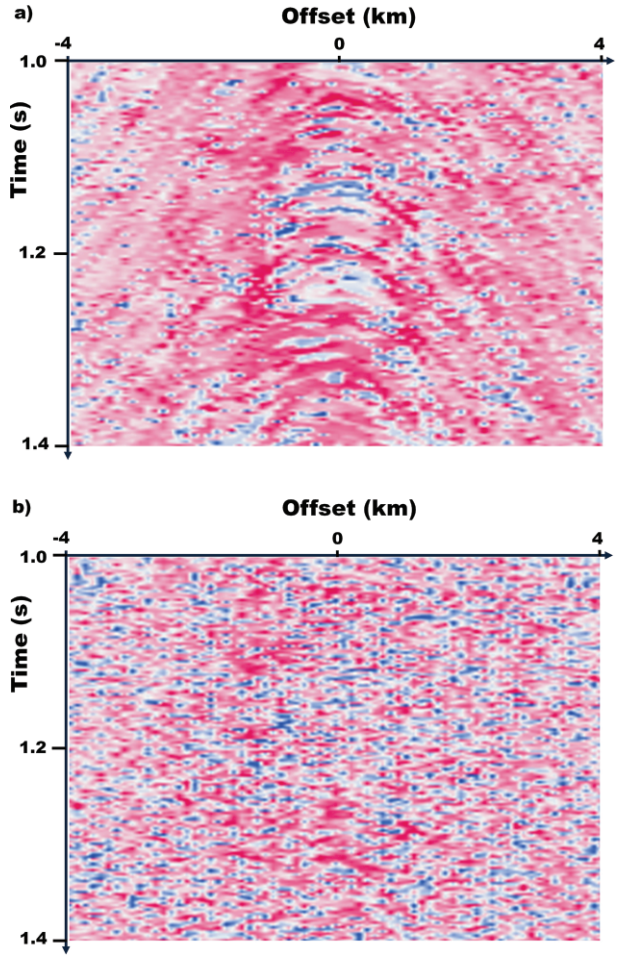

Figure 7. Parameter *B* obtained by different estimation strategies: (a) the 5D brute force approach and (b) the «2 + 2 + 1» strategy.

we find only the dips but in both coordinate directions. Likewise, during the second step, we fix estimated dips and estimate all three curvatures. Compared with the «2 + 2 + 1» strategy, both steps of this approach use the entire undecimated data ensemble (full 2D patch) instead of limited subsets used in the «2 + 2 + 1» method. During the first step, curvatures *C*, *D* and *E* are set to zero. The brute force strategy searches for optimal dips *A* and *B* that maximise semblance-based objective function in equation (3). Likewise, we fix dips A and B during the second step and use brute force strategy to find curvatures *C*, *D* and *E*. This alternative sequential strategy finds all five parameters with much less computational effort than the 5D brute force approach.

2.2. *Interpolation of kinematic parameters*

After choosing an appropriate strategy, one needs to select a grid in the data domain to estimate these parameters. Irrespective of the strategy, using every trace and time sample is not computationally feasible for massive datasets. A faster method is a so-called operator-oriented approach (Hoecht *et al.* 2009). The 'operator' defines a traveltime surface used for the local moveout correction. Kinematic parameters are estimated and stored as samples of so-called parameter traces at the decimated uniform grid in data space. These parameters define traveltime trajectories for summation. They are assumed valid for the neighborhoods between the estimated grid points. The coarser the spatial grid of parameter traces, the faster algorithm performs. However, a too-coarse grid may not accurately capture complex moveouts, leading to incoherent summation and loss of data quality for enhanced traces located far from the operator traces. To mitigate these effects in this study, we also invoke parameter interpolation: estimating parameters on coarser regular or random grids, then interpolating them to a dense original grid. While advanced interpolation and inpainting techniques can be considered (Gadylshin *et al.* 2020), we find that the simplest linear interpolation method is suitable when coarse and dense grids are regular. The ratios of the coarsened and initial grid's spatial steps are denoted by *kx* and *ky*. Likewise, instead of estimating parameters at every time sample, we coarsen the time axis with a ratio of $kt > 1$ to further reduce the number of estimations.

2.3. *Proposed optimal solution*

An optimal strategy should provide a reasonable trade-off between enhanced data quality and computational efficiency for practical applications. For real data, a reasonable trade-off is achieved with the sequential strategy «dips + curvatures» done on a sparser grid followed by the parameter interpolation as described previously (figure 8). This scheme is far more computationally affordable compared to the global 5D search algorithms. At the same time, it provides more robust kinematic parameters than a sequential «2 + 2 + 1» strategy (Hoecht *et al.* 2009; Bakulin *et al.* 2020), which may get trapped in a local maximum and limit the enhancement power for data with a low signal-to-noise ratio. We support these conclusions with numerical tests using real land data from the complex desert environment.

3. Synthetic data example

Let us compare the robustness of different estimation strategies as a function of signal-to-noise ratio using a controlled synthetic event superimposed with variable noise taken from real data. We take the geometry of real cross-spread gather shown in figure 4. The signal is represented by a single hyperbolic reflection event with a Ricker's wavelet of dominant frequency 30 Hz. The synthetic gather is generated by combining a signal with noise taken from real land data. To verify robustness to noise, estimation is repeated for different levels of SNR from −60 to 40 dB. We specifically evaluate NLBF estimation using «2 + 2 + 1» and «dips + curvatures»

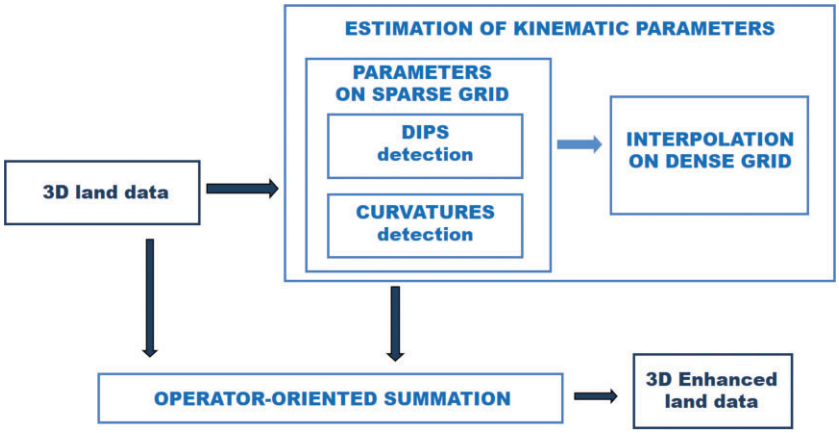

Figure 8. Flow chart of the new proposed optimal strategy for 3D seismic data enhancement.

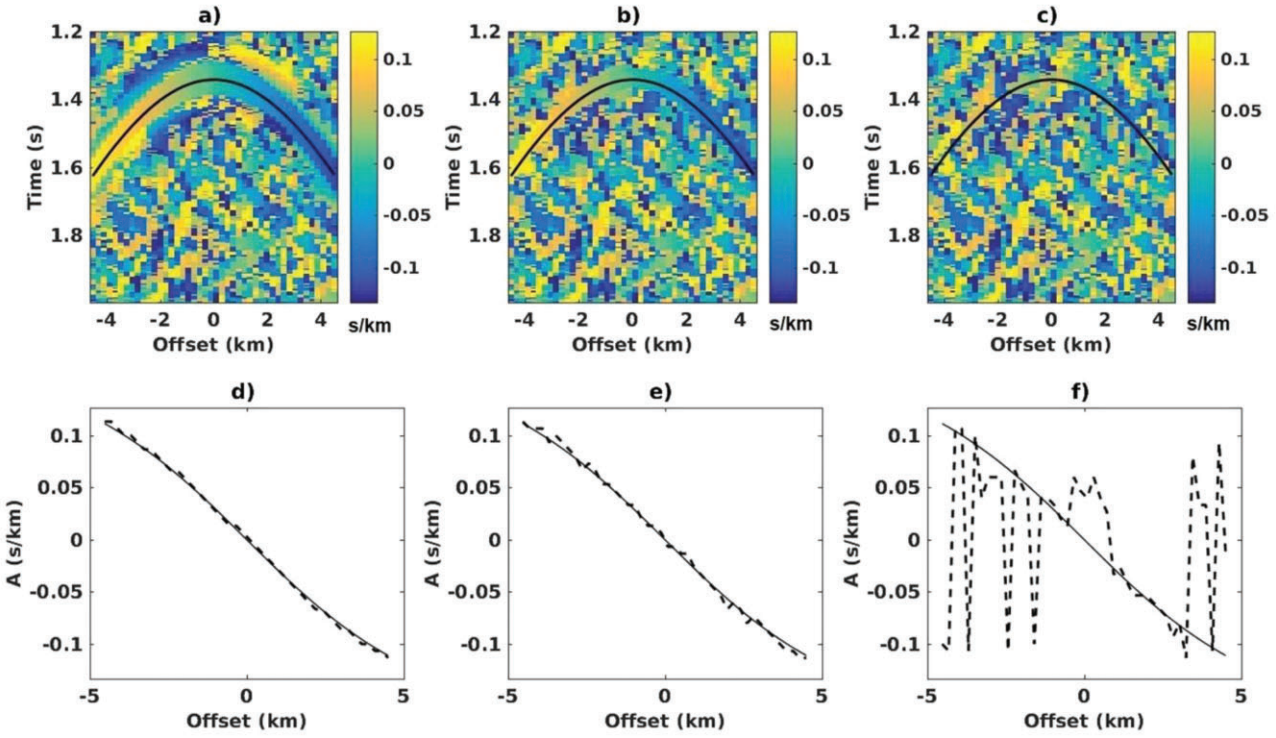

Figure 9. Parameter *A* (dip) obtained with the «2 + 2 + 1» estimation strategy without interpolation ($kt = 1$, $kx = ky = 1$) for synthetic gather with single hyperbolic reflection and variable levels of noise taken from real data: (a) SNR = 10 dB, (b) SNR = −15 dB and (c) SNR = −25 dB. Profile of the estimated parameter *A* along the hyperbolic event (dashed line) is compared with the ground truth (solid line) for different levels of noise: (d) SNR = 10 dB, (e) SNR = −15 dB and (f) SNR = −25 dB. These observe good estimation above −15 dB and unreliable estimation at −25 dB.

strategies with and without interpolation. Comparing estimated parameter *A* with the ground truth, one can see that both strategies have similar robustness for a higher SNR (higher than −15 dB). Accurate estimation is achieved along with the entire event with acceptable errors (figures 9–11). However, at lower SNRs, the «2 + 2 + 1» strategy loses its robustness earlier than the «dips + curvatures» strategy as seen by comparing figures 9f and 10f for −25dB. Large and frequent deviations become abundantly present along the event in figure 9f. This happens without the interpolation of the estimation grid. In contrast, the «dips + curvatures» strategy maintains similar estimation quality even at −25 dB without interpolation (figure 10f). Estimation errors increase with interpolation. However, they remain acceptable for practical applications (figure 11f). Similar behavior is observed for other kinematic parameters (not shown).

A similar trend is observed in figure 12, analysing mean absolute percentage error (MAPE) in individual parameters along the entire event as a function of SNR. One can

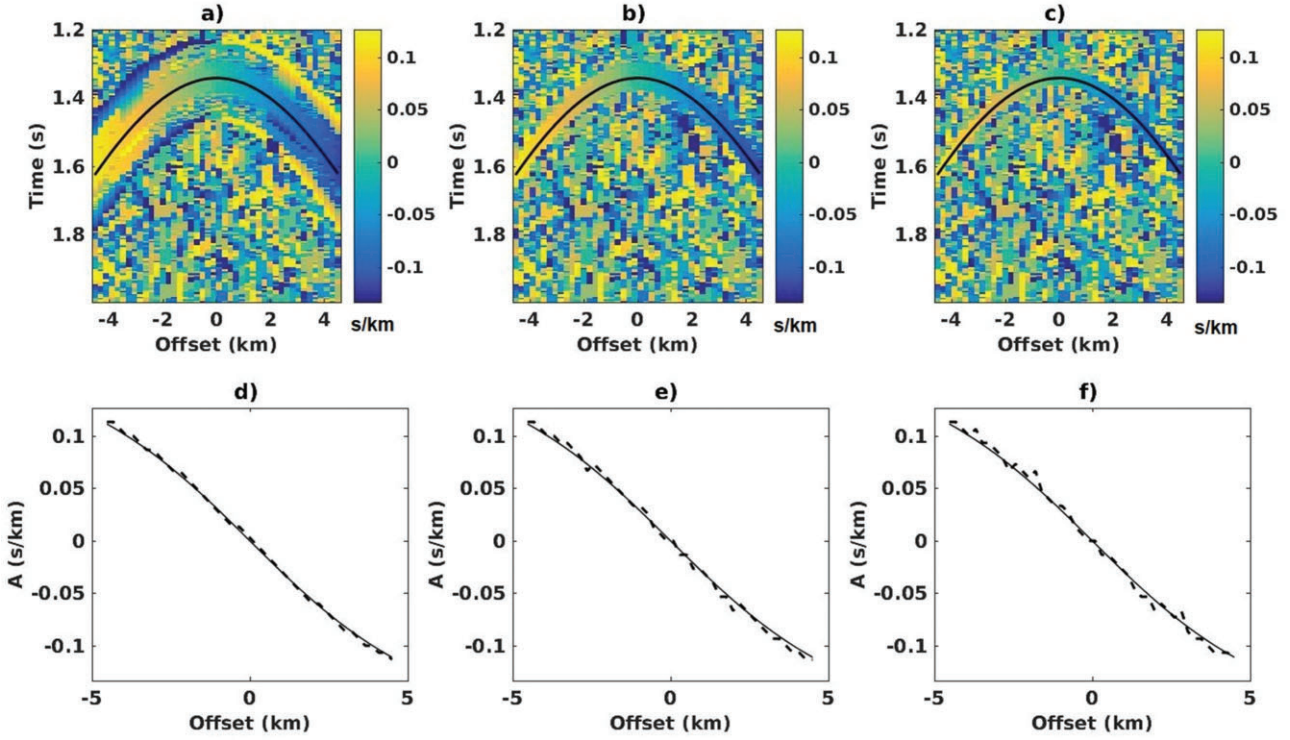

Figure 10. Same as figure 9, but obtained with the «dips + curvatures» estimation strategy without interpolation ($kt = 1, kx = ky = 1$). These observe reliable estimation all the way to −25 dB.

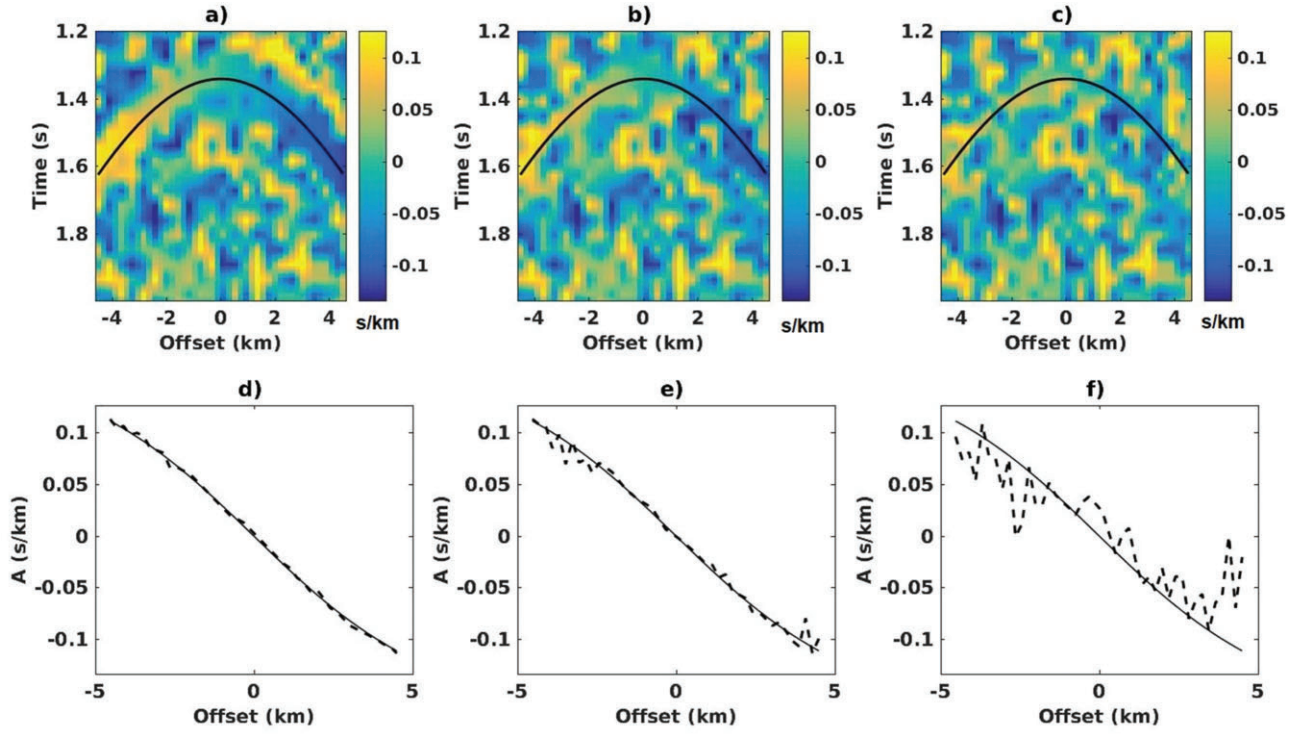

Figure 11. Same as figure 10 but obtained with the «dips + curvatures» estimation strategy with interpolation ($kt = 11, kx = ky = 2$). These observe an increased error at −25 dB (f) but are still of acceptable estimation quality for data enhancement.

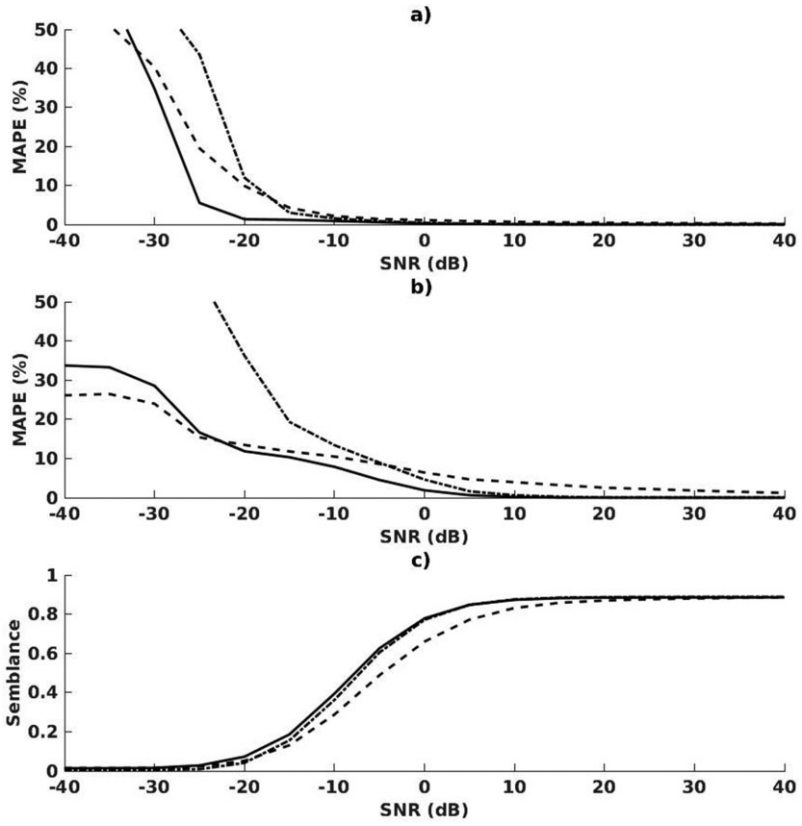

Figure 12. Mean absolute percentage error (MAPE) for estimated parameters as a function of SNR: (a) MAPE for parameter *A* (dip) and (b) MAPE for parameter *D* (curvature). (c) Absolute values of the semblance function versus SNR. Solid line—«dips + curvatures» without interpolation ($kt = 1$, $kx = ky = 1$), dashed line—«dips + curvatures» with interpolation ($kt = 11$, $kx = ky = 2$), dash-dotted line—«2 + 2 + 1» estimation strategy without interpolation ($kt = 1$, $kx = ky = 1$). Noise borrowed from real data is used in all examples. Observe smaller errors for «dips + curvatures» at the lower end of the SNR range.

Table 1. Comparison of computational effort and errors in estimated parameters for different estimation strategies measured as MAPE of all five parameters averaged over SNR interval of [−20 0] dB.

Strategy	*kx*,*ky*	*kt*	Number of traces used for estimation	Estimation time	MAPE
«2 + 2 + 1»	2	11	80 ('fat' line)	0 h 1 min	13.3%
«2 + 2 + 1»	1	11	80 ('fat' line)	0 h 3 min	12.5%
«2 + 2 + 1»	1	1	80 ('fat' line)	0 h 23 min	10.7%
«Dips + curvatures»	2	11	840 (2D patch)	0 h 33 min	6.7%
«Dips + curvatures»	1	11	840 (2D patch)	2 h 13 min	6.0%
«Dips + curvatures»	1	1	840 (2D patch)	23 h 56 min	4.3%

observe smaller MAPE values for the «dips + curvatures» as compared to the «2 + 2 + 1» strategy without interpolation (figure 12). Specifically, in the practically interesting SNR range of less than 0 dB (where NLBF should be applied), we observe that the «dips + curvatures» strategy with interpolation has a lower level of error than the «2 + 2 + 1» strategy without interpolation.

To include other parameters into account, we now compute MAPE for all five parameters further averaged over the SNR interval [−20 0] dB. The upper bound (0 dB) represents SNR when enhancement is rarely needed. The lower bound of −20 dB is chosen for this specific example as SNR corresponding to the smallest trustable semblance values (figure 12c). Table 1 shows MAPE versus the computational times for different estimation strategies. It may help analyse the trade-off between quality and computational speed. We conclude that for desired fast computational times, the «dips + curvatures» strategy with interpolation ($kt = 11$, $kx = ky = 2$) provides a smaller error in estimated parameters than the «2 + 2 + 1» strategy without interpolation. Since computational effort is comparable for both strategies, this suggests that «dips + curvatures» as the best candidate in terms of quality/speed ratio for data with low SNR.

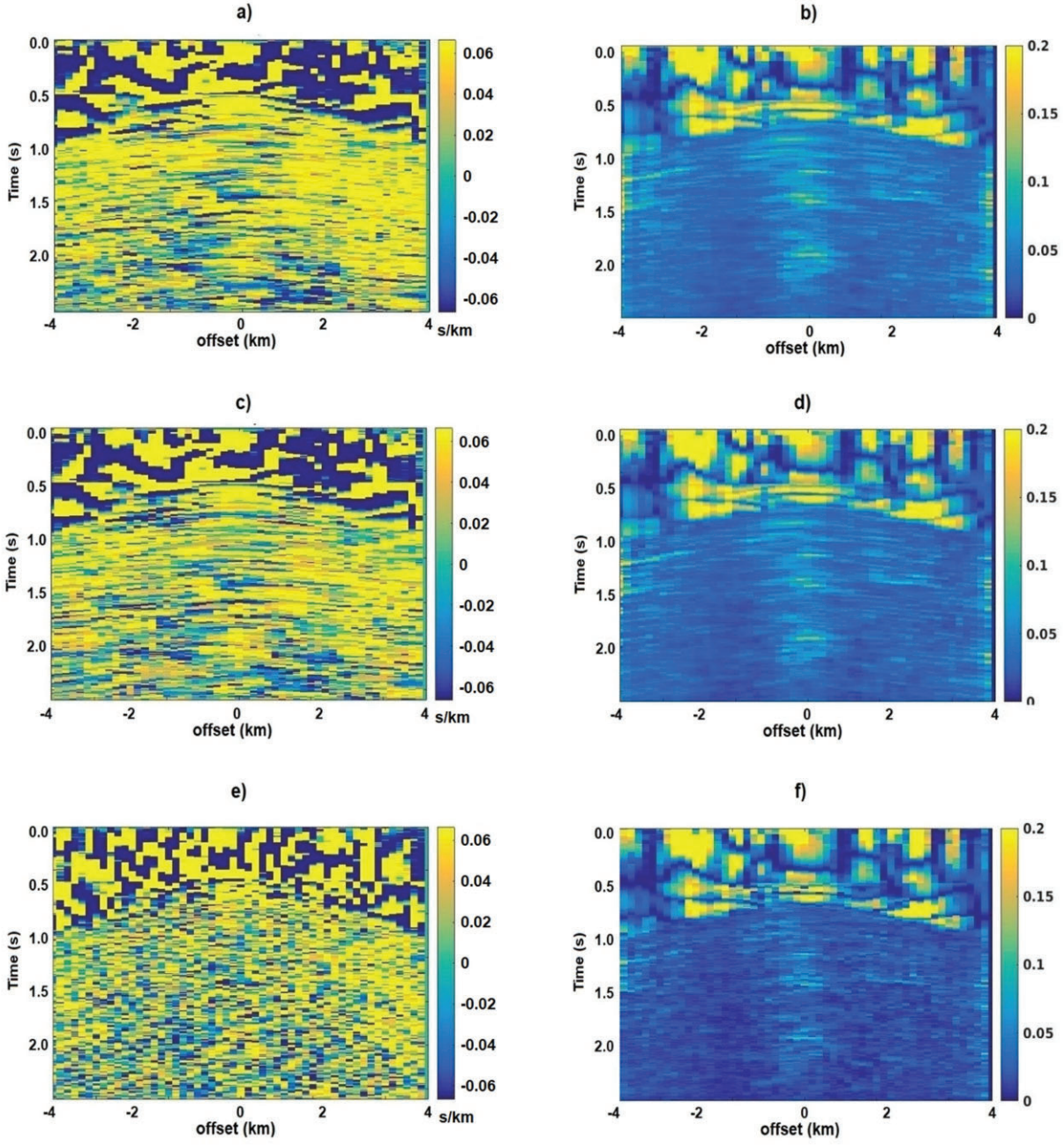

Figure 13. Parameter *B* (left) and maximum semblance (right) obtained with different estimation strategies: (a, b) 5D brute force estimation, (c, d) sequential estimation of dips and curvatures and (e, f) the «2 + 2 + 1» approach. Brute force and sequential approaches provide a more stable estimation of parameter *B* and higher semblance values than the «2 + 2 + 1» procedure.

4. Real-data examples

4.1. Comparison of different estimation strategies

Let us now analyse and understand the advantages and disadvantages of the described estimation strategies based on a series of real-data numerical experiments using nonlinear beamforming. The goal is to identify the optimal strategy in terms of quality and computational cost.

As a first example, we use legacy land seismic data acquired with 72-geophones arrays and five vibrators per sweep in a challenging desert environment with poor data quality. The data have passed through major denoising steps, including high-amplitude and linear-noise attenuation (figure 14a). As can be seen, the data's noise level is relatively high, and the reflection events are hardly visible even with large field arrays. We applied nonlinear beamforming in the cross-spread domain using different estimation strategies. The estimated parameter *B* (dip) and corresponding maximum semblance values are shown in figure 13. The remaining parameters exhibit similar behavior (not shown). The brute force 5D

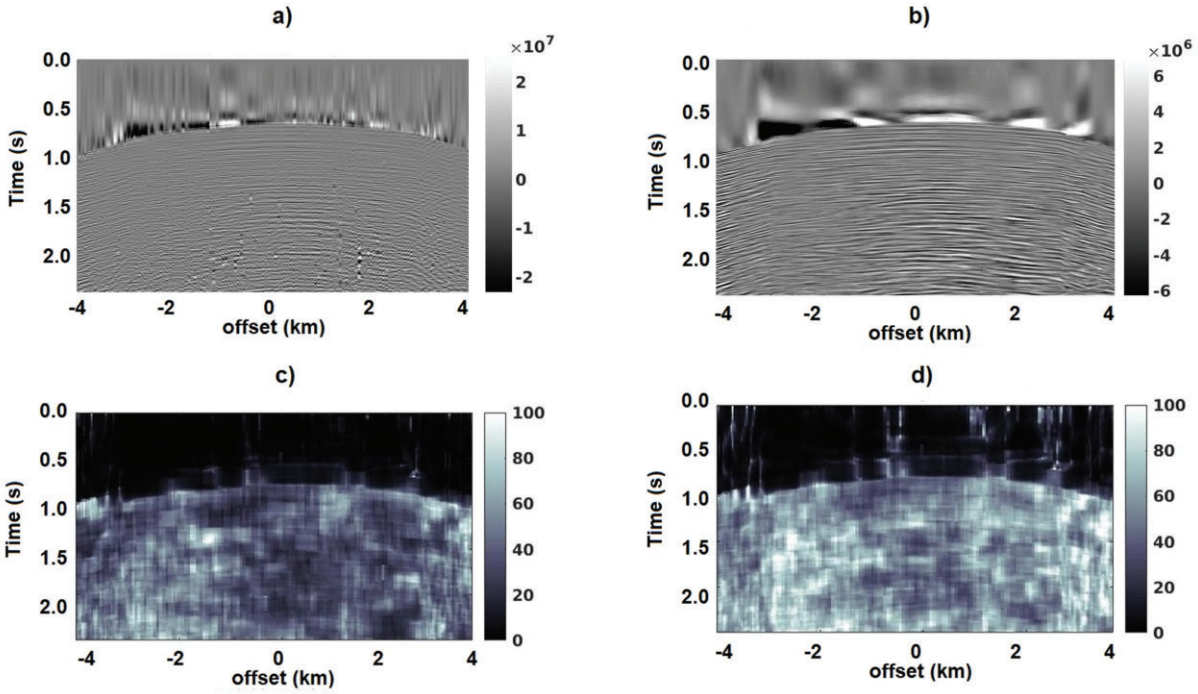

Figure 14. Typical common-shot gathers before (a) and after enhancement with NLBF (b) using 5D brute force solution (used as reference data). NRMS metric computed for reference data and data obtained by various estimation strategies: the (c) 'dips + curvatures' and (d) «2 + 2 + 1» approach. Sequential estimation (c) provides the enhancement result closer to the reference as manifested by the lower values of the NRMS compared to the «2 + 2 + 1» approach.

Table 2. Comparison of computational effort and enhanced data quality for three estimation strategies used for NLBF. Data from the 5D brute force approach is considered the best achievable reference. The closeness of the enhanced data to the reference is evaluated using the NRMS metric, with smaller values indicating better similarity.

Strategy / Metric	«2 + 2 + 1»	«Dips + curvatures»	5D brute force
Estimation time	11 min	8 h 40 min	1666 h 40 min
Number of traces used for estimation	80 ('fat' line)	840 (2D patch)	840 (2D patch)
Average semblance	0.011	0.020	0.025
NRMS	40%	29%	0% (reference)

estimation strategy shows the semblance function's stable behavior. The estimated parameter B corresponds with the weak reflection events visible in figure 14. A similar, although the slightly noisier result, is obtained using the strategy with the sequential estimation of dips and curvatures. The «2 + 2 + 1» strategy provides the most unstable result with the lowest optimal semblance value and large spatial variations in dip parameter *B* that are not consistent with the structure of reflected events. To make a quantitative comparison, we compute average semblance values for the whole gather (Table 2). We also use the normalised mean-squared (NRMS) metric (Kragh & Christie 2002) to evaluate the data's closeness to those enhanced by the 5D brute force estimation strategy, which is considered the best achievable result in this case. This strategy provides the biggest average semblance comparing to other solutions, which justifies its use as a reference. Differences between the data obtained using the 5D brute force approach and two other strategies were calculated using the NRMS metric computed in a sliding window. As shown in figure 14, strategy with the sequential estimation of dips and curvatures provides data closer to the reference solution than the «2 + 2 + 1» strategy. The average NRMS metric value calculated over the entire gather is around 29% for the «dips + curvatures» strategy. In contrast, it is higher at 40% for the «2 + 2 + 1» approach. Lower NRMS means closer similarity to a reference solution

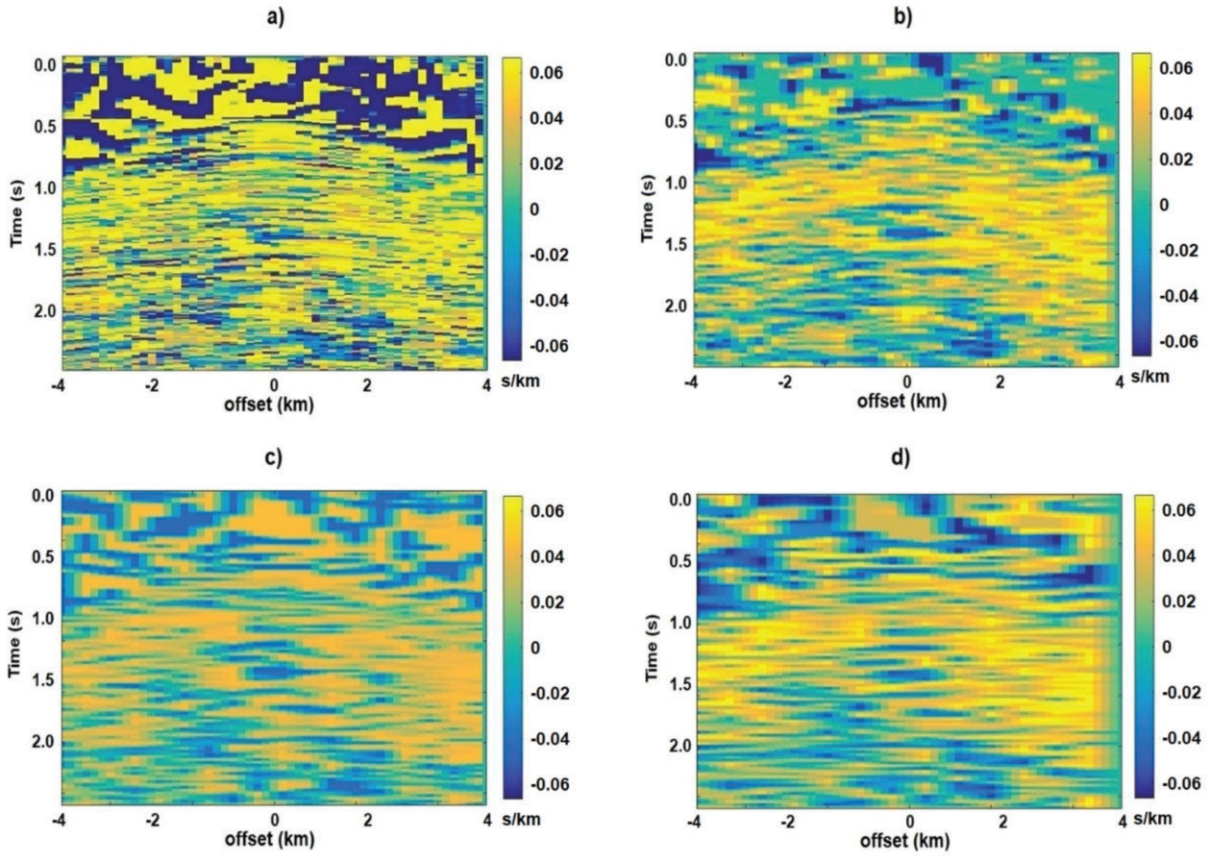

Figure 15. Parameter *B* estimated by the «dips + curvatures» approach using various decimation/interpolation parameters in space and time: (a) no interpolation ($kt = 1, kx = ky = 1$); (b) $kt = 11, kx = ky = 2$; (c) $kt = 11, kx = ky = 3$ and (d) $kt = 11, kx = ky = 4$. For each case with decimated estimation grids (b–d), we linearly interpolate computed parameters to the same dense grid as for the case without interpolation (a).

confirming that the latter approach provides better quality enhancement.

As a next step, we examine a computation cost for all the strategies. Table 2 shows a comparison of running times required to estimate kinematic parameters for one cross-spread gather. Although the full 5D brute force approach provides the best results in terms of quality, it is unacceptably slow for practical applications at this time. The «2 + 2 + 1» strategy provides the best result in performance and shows a 40× faster computational time compared to the «dips + curvatures» strategy.

4.2. Comparison of attributes interpolation strategies

Let us now study the influence of attributes interpolation on the NLBF results. We performed a series of experiments with different estimation grid sparsity by varying the coefficients *kx*, *ky* and *kt*, defined previously. Since the semblance value during a coherency search is calculated in a particular time window, the half-window size appears to be the right candidate for a decimated grid step in the time direction. We used the time window of 22 time samples in the numerical examples, resulting in a 44 ms semblance window at 2-ms sampling. This is comparable to one period at a dominant frequency. Therefore, $kt = 11$ or a decimated time step of 22 ms is selected for interpolation tests. We use equal estimation grid steps for spatial directions along the *x* and *y* axes and investigate coefficients $kx = ky = 2$, 3 and 4, resulting in 200, 300 and 400 m intervals in the estimation grid. The estimated values of parameter *B* for each of these grids are presented in figure 15. For $kx = 2$, estimated parameters closely resemble the values calculated on an original densest estimation grid before any interpolation in time or space. Further increasing the spatial grid interval with $kx = 3, 4$ leads to visually over-smoothed behavior suggesting loss of resolution.

To quantitatively understand the influence of such interpolation on the quality of the enhanced data and estimation cost, we compare computational times, average semblance and NRMS values (figure 16, Table 3). Like Table 2, we continue to use the same reference solution from the 5D strategy without interpolation. Analysing the results from Tables 2 and 3, we conclude that the strategy with a sequential estimation of dips and curvatures on a sparse grid ($kx = ky = 2$, $kt = 11$) delivers a better solution than the «2 + 2 + 1»

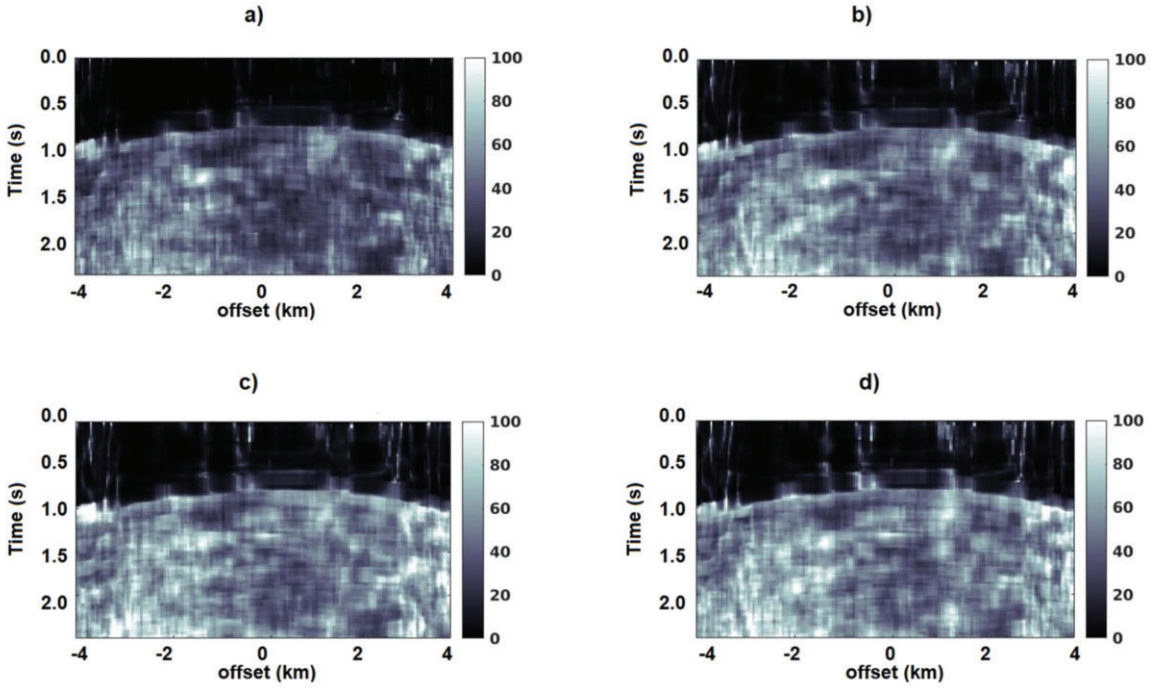

Figure 16. NRMS difference of enhanced data and reference data from 5D approach (no interpolation). Similar to figure 14, we evaluate the «dips + curvatures» approach using various decimation/interpolation parameters in space and time: (a) no interpolation ($kt = 1$, $kx = ky = 1$); (b) $kt = 11$, $kx = ky = 2$; (c) $kt = 11$, $kx = ky = 3$ and (d) $kt = 11$, $kx = ky = 4$. For each case with decimated estimation grids (b–d), we linearly interpolate computed parameters to the same dense grid as for the case without interpolation (a). Interpolation with $kt = 11$, $kx = ky = 2$ (b) provides an average NRMS of 35% (Table 3) that is used in this study as an acceptable threshold of similarity with the reference data. In contrast, cases (c, d) deliver average NRMS above 35%.

Table 3. Effect of interpolation on computational effort and enhanced data quality for the strategy with the sequential estimation of dips and curvatures. Using sparser grids for estimation with subsequent interpolation leads to a significant reduction in computing time/cost. NRMS values increase only by a small percentage during interpolation, indicating only slightly worsening similarity but still acceptable for practical purposes.

	5D brute force	«Dips + curvatures»	«Dips + curvatures» ($kx = ky = 2$, $kt = 11$)	«Dips + curvatures» ($kx = ky = 3$, $kt = 11$)	«Dips + curvatures» ($kx = ky = 4$, $kt = 11$)
Estimation time	1666 h 40 min	8 h 40 min	12 min	5 min	3 min
Average semblance	0.025	0.020	0.020	0.019	0.020
NRMS	0% (reference)	29%	35%	40%	40%

strategy in terms of semblance values (0.2 vs. 0.11) and quality while using similar computational effort (12 vs. 11 min). Therefore, the proposed sequential estimation of dips and curvatures followed by interpolation provided an optimal approach for real seismic data striking an excellent trade-off between computational time and achieved quality. We further stress that «dips + curvatures» increase the estimation process's robustness, especially for challenging real data from a desert environment with single sensors and small field arrays.

4.3. *Enhancement of first arrivals for land single-sensor data*

As a second real-data example, we consider enhancing early arrivals for single-sensor land data with the orthogonal acquisition geometry. Receiver spacing is 20 m along the lines with a cross-line interval of 120 m. Likewise, source intervals are 20 m (inline) and 80 m (cross-line). Only a minor high-amplitude noise attenuation was applied to the data in this example to preserve first-arrival energy. At far offsets, the signal level is very week, and first arrivals are merely invisible in raw gathers (figure 17a). Figure 17d shows first-arrival traveltimes picked using an automatic algorithm on a cross-spread gather. Figure 18 shows the mean of these first-break picks values as a function of offset. We note that starting from an offset of approximately 2250 m, the picking algorithm completely fails and follows a mute function applied to the data before picking. Hence, it does not provide reliable information at far offsets that can be useful, particularly for full-waveform inversion applications.

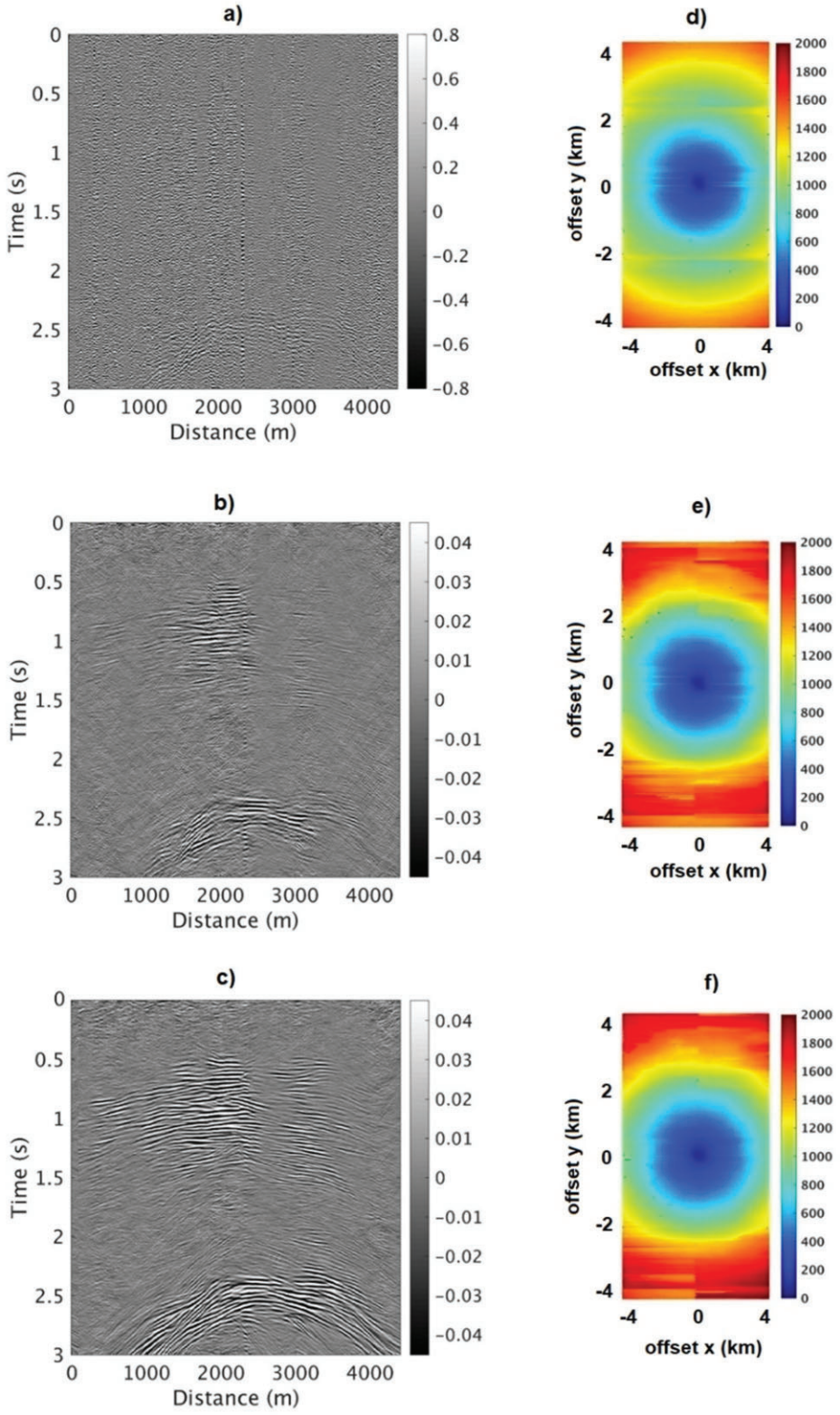

Figure 17. Far-offset receiver line from a 3D common-shot gather of land single-sensor data after minor preprocessing (a) and corresponding map of first-arrival traveltimes picked using an automatic algorithm (d). The same gather is shown after enhancement with nonlinear beamforming using the «2 + 2 + 1» approach without interpolation (b) and «dips + curvatures» with interpolation strategy (c). The corresponding traveltime maps obtained from the enhanced data are displayed in (e) and (f), respectively.

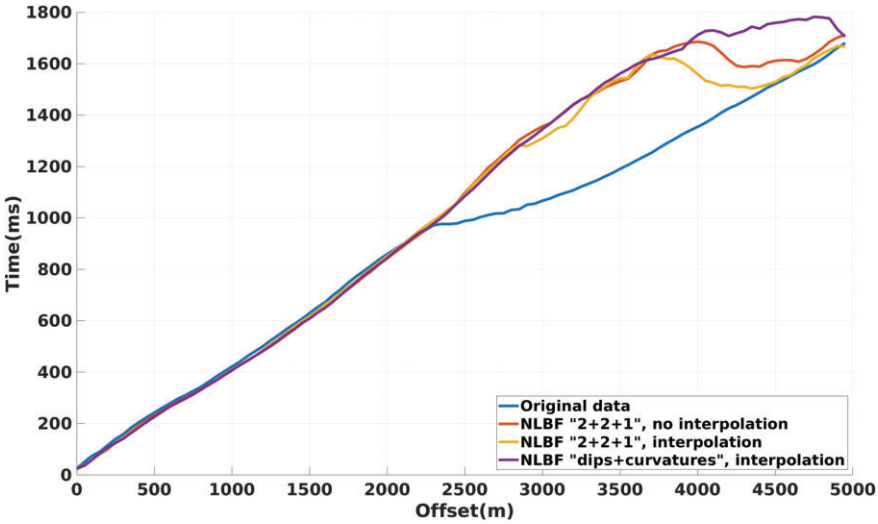

Figure 18. Mean values of first-break picks from the original data (blue) and the data after NLBF with different estimation strategies: the «2 + 2 + 1» approach without (brown) and with interpolation (yellow); the «dips + curvatures» approach with interpolation (purple). Interpolation parameters are $kt = 12$, $kx = ky = 2$.

We apply nonlinear beamforming to this data using the «2 + 2 + 1» strategy both with and without parameters interpolation and the «dips + curvatures» strategy with interpolation. The data quality improves dramatically after NLBF in all the cases compared to the original raw data (figure 17). The «dips + curvatures» strategy provides a higher quality enhancement that results in more coherent and clear events than the «2 + 2 + 1» strategy. The corresponding first arrivals maps obtained using the enhanced data show a more reliable shape on far offsets than the original raw data (figure 17). This time, the traveltimes follow actual enhanced first arrivals and not the mute function (figures 17 and 18). The more detailed comparison of mean values in figure 18 shows that the data quality after enhancement using the «2 + 2 + 1» approach with interpolation degrades faster than the same approach but with interpolation. The «dips + curvatures» strategy with interpolation helps obtain the most reliable picks in this case that tend to approach the longest offset available in the data. Moreover, the observed improvements are directly related to the estimated parameters' enhanced quality (figure 19). Table 4 illustrates that the required computational time is comparable between the «dips + curvatures» strategy with interpolation $kt = 12$, $kx = ky = 2$ (14 hours 30 minutes), and «2 + 2 + 1» without interpolation $kt = kx = ky = 1$ (16 hours 19 minutes). At the same time, the quality of the former one is better. We conclude that the «dips + curvatures» approach gives us the best trade-off in data quality and computation speed.

5. Conclusions

Data-enhancement algorithms based on local stacking, such as nonlinear beamforming, are computationally expensive. They require local coherency search on massive spatial/temporal grids in 3D prestack data cubes. Efficient implementation of such algorithms poses an inevitable trade-off between the algorithm's computational performance and the achieved quality of the enhanced data. We compare several existing and new estimation strategies for nonlinear beamforming currently in wide use to process data from the desert environment with complex near surfaces. Specifically, we examine the conventional «2 + 2 + 1» estimation strategy versus the new sequential strategy evaluating dips and then curvatures. Synthetic and real-data examples show that sequential estimation of dips and curvatures with interpolation provides an optimal trade-off between computational speed and data quality. Table 4 summarises our semiquantitative assessment and ranking of various estimation strategies for practical applications. The new strategy, dubbed «dips + curvatures», has proven to be significantly more robust when applied to real data with a low signal-to-noise ratio. This robustness is achieved using the entire undecimated data ensemble for both estimation steps instead of decimated 'fat' lines along each coordinate as in the «2 + 2 + 1» approach. Using two challenging 3D land seismic datasets, we demonstrate that the new strategy outperforms the conventional «2 + 2 + 1» approach and provides results closer to the reference data obtained with the 5D brute force approach currently out of computational reach for massive seismic volumes. To compensate for larger ensemble sizes and a more expensive optimisation scheme in the «dips + curvatures» approach, we propose parameter computation on a coarser grid in space and in time, followed by interpolation back to the original grid. With a not-too-aggressive grid decimation followed by straightforward linear parameter interpolation, we achieve significant

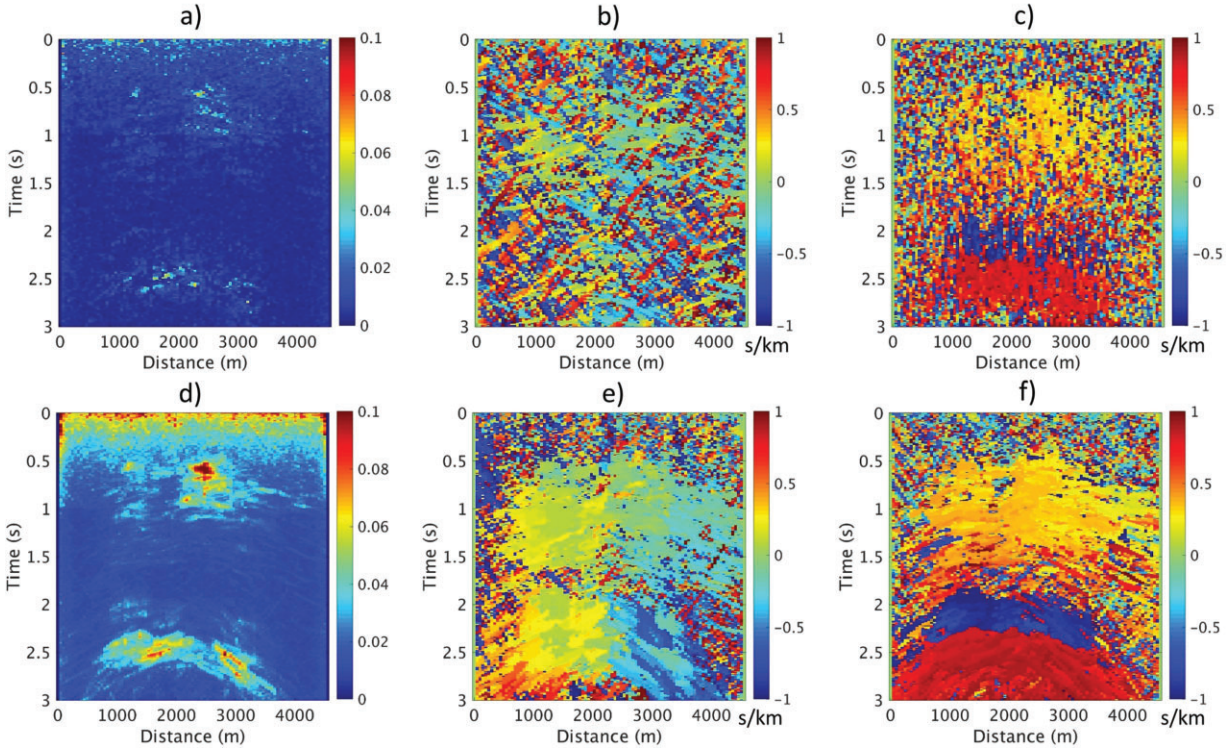

Figure 19. Maximum semblance (left) and dip parameters *A* (middle) and *B* (right) obtained using various estimation strategies: the «2 + 2 + 1» (a, b, c) and the sequential estimation of dips and curvatures (d, e, f). Observe the more geologically plausible behavior of parameters (d, e, f) with images resembling arrivals on the enhanced data (figure 17c). In contrast, (a, b, c) show more erratic and unstable behavior caused by limitations of the «2 + 2 + 1» approach.

Table 4. Semiquantitative assessment and ranking of various estimation strategies for practical applications in terms of quality vs. performance.

Option	Quality	Performance	Usage priority (qaulity vs. performance)
5D brute force	Best achievable	Not considered	Quality reference
«2 + 2 + 1»	Average	Good	2
«Dips + curvatures»	Good	Acceptable	4
«2 + 2 + 1» with interpolation	Acceptable	Good	3
«Dips + curvatures» with interpolation	Good	Good	1

additional speedup making a new strategy computing time/cost comparable to the conventional «2 + 2 + 1» approach. Simultaneously, comparing parameters and data quality obtained with «dips + curvatures» and a conventional strategy, we find them similar for synthetic data without noise.

In contrast, the new strategy delivers considerably better enhanced data quality for real 3D land seismic data with strong noise. This is also supported by synthetic data examples with different levels of real-data noise. Also, data quality remains similar when comparing the «dips + curvatures» approach with and without interpolation as measured by a sensitive NRMS repeatability metric experiencing only a slight reduction. Therefore, interpolating more accurately computed parameters with a new strategy is highly robust, unlike the conventional approach. Interpolation can still be usable for aggressive decimation, but more advanced parameter interpolation schemes may be required. In summary, we introduce an optimised version of the NLBF algorithm to enhance 3D seismic data based on the sequential estimation of dips and curvatures on the sparse grid followed by an interpolation to dense original grid. Applications to challenging 3D land single-sensor datasets demonstrate clear advantages in data quality and first-break picks derived from enhanced data.

Conflict of interest statement. The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

Abma, R. & Claerbout, J., 1995. Lateral prediction for noise attenuation by t-x and f-x techniques, *Geophysics*, **60**, 1887–1896.

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

Bakulin, A., Golikov, P., Dmitriev, M., Neklyudov, D., Leger, P & Dolgov, V., 2018. Application of supergrouping to enhance 3D prestack seismic data from a desert environment, *The Leading Edge*, **37**, 306–313.

Bakulin, A, Silvestrov, I., Dmitriev, M., Neklyudov, D., Protasov, M., Gadylshin, K. & Dolgov, V., 2020. Nonlinear beamforming for enhancement of 3D prestack land seismic data, *Geophysics*, **85**, V283–V296.

Baykulov, M. & Gajewski, D., 2009. Prestack seismic data enhancement with partial common-reflection-surface (CRS) stack, *Geophysics*, **74**, V49–V58.

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

Berkovitch, A., Gelchinsky, B. & Keydar, S., 1994. Basic formula for multifocusing stack, *56th EAGE Meeting, Expanded Abstracts*, 140–143.

Buzlukov, V. & Landa, E., 2013. Imaging improvement by prestack signal enhancement, *Geophysical Prospecting*, **61**, 1150–1158.

Canales, L., 1984. *Random noise reduction: SEG Technical Program Expanded Abstracts*, 525–527.

Chen, K. & Sacchi, M., 2015. Robust reduced-rank filtering for erratic seismic noise attenuation, *Geophysics*, **80**, V1–V11.

Cordery, S., 2020. An effective data processing workflow for broadband single-sensor single-source land seismic data, *The Leading Edge*, **39**, 401–410.

Duncan, G. & Beresford, G., 1994. Slowness adaptive *f-k* filtering of prestack seismic data, *Geophysics*, **59**, 140–147.

Fam., H. & Naghizadeh, M., 2019. Multi-Focusing stacking using the Very Fast Simulated Annealing global optimization algorithm, *GeoConvention 2019, Calgary, Extended Abstracts*, 1–6.

Fomel, S., 2002. Applications of plane-wave destruction filter, *Geophysics*, **67**, 1946–1960.

Fomel, S. & Liu, Y., 2010. Seislet transform and seislet frame, *Geophysics*, **75**, V25–V38.

Gadylshin, K., Silvestrov, I. & Bakulin, A., 2020. Inpainting of local wavefront attributes using artificial intelligence for enhancement of massive 3-D prestack seismic data, *Geophysical Journal International*, **223**, 1888–1898.

Garabito, G. & Cruz, C., 2019. application of very fast simulated annealing and differential evolution in the search for FO-CRS wavefield attributes, *Geophysics*, **84**, O81–O92.

Gao, Y., Zhao, P., Li, G. & Li, P, 2021. Seismic noise attenuation by signal reconstruction: an unsupervised machine learning approach, *Geophysical Prospecting*, **69**, 984–1002.

Gelchinsky, B., Berkovitch, A. & Keydar, S., 1999. Multifocusing homeomorphic imaging: Part 2. Multifold data set and multifocusing, *Journal of Applied Geophysics*, **42**, 243–260.

Gulunay, N., 1986. FXDECON and complex Wiener prediction filter, *SEG Technical Program Expanded Abstracts*, 279–281.

Herrmann, F. & Hennenfent, G., 2008. Non-parametric seismic data recovery with curvelet frames, *Geophysical Journal International*, **173**, 233–248.

Hoecht, G., Ricarte, P., Bergler, S. & Landa, E., 2009. Operator-oriented CRS interpolation, *Geophysical Prospecting*, **57**, 957–981.

Ibrahim, A., Sacchi & M. D.,2014, Simultaneous source separation using a robust radon transform, *Geophysics*, **79**, V1–V11.

Jäger, R., Mann, J., Höcht, G. & Hubral, P., 2001. Common-reflection surface stack: image and attributes, *Geophysics*, **66**, 97–109.

Kelamis, P. & Mitchell, A., 1989. Slant-stack processing, *First Break*, **7**, 43–54.

Kragh, E. & Christie, P., 2002. Seismic repeatability, normalized RMS, and predictability, *The Leading Edge*, **21**, 640–647.

Mann, J., 2002. *Extensions and applications of the common-reflection surface stack method, PhD thesis*, University of Karlsruhe.

Marfurt, K., Kirlin, R., Farmer, S. & Bahorich, M., 1998. 3-D seismic attributes using a semblance-based coherency algorithm, *Geophysics*, **63**, 1150–1165.

Mousavi, S., Langston, C. & Horton, S., 2016. Automatic microseismic denoising and onset detection using the synchrosqueezed continuous wavelet transform, *Geophysics*, **81**, V341–V355.

Naghizadeh, M., 2012. Seismic data interpolation and denoising in the frequency-wavenumber domain, *Geophysics*, **77**, V71–V80.

Naghizadeh, M. & Sacchi, M., 2012. Multicomponent f-x seismic random noise attenuation via vector autoregressive operators, *Geophysics*, **77**, V91–V99.

Oropeza, V. & Sacchi, M., 2011. Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis, *Geophysics*, **76**, V25–V32.

Rashed, M., 2014. Fifty years of stacking, *Acta Geophys*, **62**, 505–528

Robinson, J. C., 1970. Statistically optimal stacking of seismic data, *Geophysics*, **35**, 436–446.

Saad, O. & Chen, Y., 2020. Deep denoising autoencoder for seismic random noise attenuation, *Geophysics*, **85**, V367–V376.

Stewart, R.R. & Schieck, D.G., 1989. 3-D F-K filtering, *SEG Technical Program Expanded Abstracts*, 1123–1124.

Trad, D., Ulrych, T. & Sacchi, M., 2003. Latest views of the sparse radon transform, *Geophysics*, **68**, 386–399.

Ulrych, T., Freire, S. & Siston, P., 1988. Eigenimage processing of seismic sections, *58th SEG Annual International Meeting, Expanded Abstracts*, 1261–1265.

Wang, Y., 1999. Random noise attenuation using forward-backward linear prediction, *Journal of Seismic Exploration*, **8**, 133–142.

Zhang, Y., Bergler, S. & Hubral, P., 2001. Common-reflection-surface (CRS) stack for common offset, *Geophysical Prospecting*, **49**, 709–718.