Simultaneous seismic interpolation and statics estimation of land data

Alfaraj, Ali M.; Verschuur, D. J.Eric; Herrmann, Felix J.

DOI
10.1190/geo2023-0668.1

Publication date
2025

Document Version
Final published version

Published in
Geophysics

GEOPHYSICS, VOL. 90, NO. 1 (JANUARY-FEBRUARY 2025); P. V43–V52, 11 FIGS., 2 TABLES.
10.1190/GEO2023-0668.1

Simultaneous seismic interpolation and statics estimation of land data

Ali M. Alfaraj[1], D. J. (Eric) Verschuur[2], and Felix J. Herrmann[3]

ABSTRACT

Imaging and inversion of land seismic data affected by complex weathering layers near the surface are challenging. When the data are additionally subsampled for economical reasons such as monitoring of sequestrated carbon dioxide and hydrogen, the problem is further exacerbated due to the combined influence of subsampling and weathering layers. First, interpolation performs poorly because the weathering layers reduce the data's coherency. Second, near-surface corrections require knowledge of the subsurface model, separation between primaries and multiples, as well as subsurface velocity estimation, which are difficult to perform from subsampled data. To overcome these hurdles, we combine seismic interpolation and statics estimation into a joint single rank-reduction-based algorithm. To our knowledge, this is the first time that this has been done. Our method simultaneously accounts for the weathering and subsampling effects, which both contribute to the low-rank (LR) structure destruction typically associated with statics-free densely sampled data, to provide accurate reconstruction. Because an LR approximation is used for statics estimation, we also use it in rank-minimization interpolation as a cost-free initial solution to the optimization problem. As statics estimation and interpolation operate in the midpoint-offset domain, we avoid the cost of transformations back and forth from the source-receiver to the midpoint-offset transform domain. Consequently, our reconstruction, which indicates its potential on synthetic and field data, also is computationally efficient.

INTRODUCTION

Due to their low velocity and rapidly varying nature, the weathering layers near the surface dictate how land seismic data are acquired (Keho and Kelamis, 2012). An additional layer of complexity arises when collecting data in the presence of acquisition gaps, e.g., due to acquisition limitations such as obstacles. Another scenario that can result in acquisition gaps is the subsampling of sources and receivers to reduce acquisition costs, which is in high demand for exploration, engineering, and CO_2 monitoring purposes. One approach to do that is via compressive sensing (Donoho, 2006; Candès and Wakin, 2008), which avoids the strict Nyquist sampling criterion with randomized subsampling. In both cases, the data contain gaps that require interpolation to fill in the missing traces (Herrmann et al., 2012).

For interpolation, transform-domain-based methods are popular as they are capable of distinguishing and using properties of dense and subsampled data. Such domains include the Fourier (Zwartjes and Sacchi, 2007), curvelet (Hennenfent and Herrmann, 2008; Alfaraj et al., 2017), and Radon (Kabir and Verschuur, 1995; Trad et al., 2002) basis functions. An alternative approach is to use rank-based methods, which can process multidimensional large-scale seismic data in a computationally efficient manner (Trickett et al., 2010; Oropeza and Sacchi, 2011; Kreimer and Sacchi, 2012; Aravkin et al., 2014). However, rank-based methods that require repetitive computations of the singular values of large matrices can be computationally demanding. This issue can be mitigated with truncated singular value decomposition (SVD) or matrix factorization techniques (Recht et al., 2010; Recht and Ré, 2013; Cheng and Sacchi, 2016). Another challenging issue is rank-based interpolation of high frequencies due to their highly varying nature, which Zhang et al. (2019) improve by using common information among neighboring frequencies. Although the aforementioned reconstruction methods demonstrate their success with statics-free and therefore low-rank (LR) seismic data,

Manuscript received by the Editor 22 November 2023; revised manuscript received 22 August 2024; published ahead of production 7 October 2024; published online 7 January 2025.
[1]Saudi Aramco, EXPEC Advanced Research Center, Dhahran, Saudi Arabia; formerly Delft University of Technology, Department of Imaging Physics, Delft, The Netherlands. E-mail: ali.alfaraj.2@aramco.com (corresponding author).
[2]Delft University of Technology, Department of Imaging Physics, Delft, The Netherlands. E-mail: d.j.verschuur@tudelft.nl.
[3]Georgia Institute of Technology, School of Earth and Atmospheric Sciences, School of Computational Science and Engineering, Atlanta, Georgia, USA. E-mail: felix.herrmann@gatech.edu.

reconstruction of data affected by weathering layers, which lead to low data coherency, remains challenging.

Trad (2009) studies 5D sparse Fourier interpolation and finds that interpolation becomes difficult when there are imperfections in the applied statics before interpolation, which we also illustrate. Stanton et al. (2013) use sparsity promotion in the Fourier domain with projection onto convex sets to reconstruct subsampled data contaminated with residual statics. Even though the method results in improved reconstruction of data synthetically shifted by ± 10 ms of residual statics, it requires data windowing to ensure sparsity, which is not a trivial task and may influence the results. When the windowed data are merged back together, there is a potential of introducing errors along the spatial and temporal dimensions as each window may obtain assigned different statics. Gholami (2014) applies phase retrieval using only the amplitude spectrum with sparsity-promoting regularization to interpolate data affected by residual statics. Sparsity promotion in the Fourier and curvelet domains also has been used for statics correction (Gholami, 2013), which also require data windowing. Mosher et al. (2017) acquire and reconstruct compressively sampled land data, where statics correction due to ice lakes is one of the challenging issues that need to be resolved. In complex near-surface (NS) regimes, where statics correction on its own is a major challenge, data reconstruction also is expected to suffer.

The current approach to tackle the combined effects of weathering and acquisition gaps is to first perform statics correction to make the data more coherent, followed by data interpolation of normal-moveout (NMO)-corrected data (Trad, 2009). However, interpolation and statics estimation depend on the knowledge of the subsurface model such as NMO velocity and separation between primaries and multiples (Taner et al., 1974; Wiggins et al., 1976; Cox, 1999). Because the velocity model is usually influenced by the statics and vice versa (Yilmaz, 2001), velocity estimation becomes more challenging when the data are additionally subsampled thus decreasing the number of traces in each common midpoint (CMP). Despite the extensive experience with NMO velocity estimation, the process remains labor intensive and time consuming, and it is prone to errors in the preceding situations or in the presence of multiples that can be confused with primaries. To avoid dependency on the subsurface model for statics estimation of periodically and densely sampled data, Alfaraj et al. (2023) use a model-independent LR-based approach. However, acquisition of such data is prohibitively expensive. In contrast, reduction of the acquisition costs with randomized subsampling influences the rank structure, which consequently affects the LR-based statics estimation.

The paper is organized as follows. We first study the singular values decay behavior in the combined presence of weathering and subsampling effects. Next, we propose a joint rank-reduction-based interpolation and statics correction method. We provide the details of the proposed algorithm and its components, which we subsequently apply to synthetic and field data. We then discuss data windowing, parameters selection, sensitivity to the rank selection, computational efficiency, limitations, and extensions of the method.

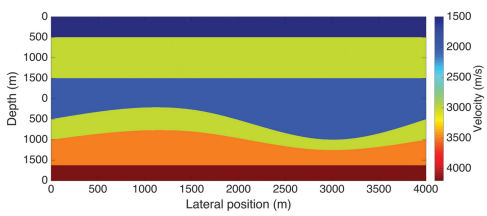

Figure 1. Velocity model used to simulate the synthetic data.

THE SINGULAR VALUES DECAY

Similar to other rank-reduction-based methods, our proposed reconstruction exploits the redundant nature of seismic data, which leads to an LR structure where the singular values decay rapidly. When the data are acquired with randomized subsampling (Donoho, 2006; Candès and Wakin, 2008; Herrmann et al., 2012) or when influenced by the weathering layers, where both processes render coherent energy incoherent, the LR structure is destroyed, leading to slowly decaying singular values (Kumar et al., 2015; Alfaraj et al., 2023). To observe their combined influence, we analyze the singular values decay behavior.

For this analysis, we simulate densely sampled data with 10 m source and receiver intervals using the velocity model shown in Figure 1 with finite-difference modeling (Thorbecke and Draganov, 2011). To exploit the redundant nature of the seismic data, we use the midpoint-offset domain, where the midpoint m is defined as

$$m = \frac{s + r}{2} \tag{1a}$$

and the offset h as

$$h = s - r, \tag{1b}$$

where s and r are the source and receiver coordinates, respectively. Because field data are not always on a desired grid, the computational grid will be different from the acquisition grid. In this case, binning becomes necessary. For field data example, we place the traces at their nearest midpoint-offset grid. After processing the data, the traces can be sorted back to the original domain.

Because the midpoint-offset transform rotates a matrix (complex-valued frequency slice) in the source-receiver domain by 45°, the matrix columns become more linearly dependent in the midpoint-offset transform domain. This can be visually seen when comparing frequency slices in the source-receiver and midpoint-offset domains (Alfaraj et al., 2023), where the columns in the latter exhibit lower variability compared with the former. Consequently, the singular values in the midpoint-offset domain decay rapidly compared with those in the source-receiver domain. Alternatively, when using windowed data, which we avoid in this work, as elaborated further in the "Discussion" section, Trickett (2003) shows that a grid of traces made up of the sum of plane waves having at most k dips results in matrices with the rank of at most k.

When the data become incoherent, the LR structure is destroyed because linear dependency of the matrix columns becomes lower. To replicate complex NS conditions, we shift the simulated data with up to ± 52 ms of statics composed of surface- and nonsurface-consistent elements. We then subsample the data by randomly removing 75% of the shots to mimic a compressive sensing acquisition design (Figure 2a and 2b). Removing 75% of the shots may sound harsh, but that is analogous to an average of 40 m periodic shot spacing, which is common for data acquisition in large areas. Although it is optimal to collect densely sampled data, it also is important to reduce acquisition costs. For example, affordable data

DOI:10.1190/geo2023-0668.1

are necessary for engineering purposes or subsurface monitoring of CO_2 and reservoirs. However, in complex NS regimes, the singular values (Figure 2e and 2f) become slowly decaying due to the effects of statics and subsampling compared with the statics-free densely sampled data. Therefore, we can use rank-reduction methods to estimate the latter from subsampled data affected by the weathering layers.

To overcome the challenging effects of subsampling and NS weathering layers, we combine seismic data reconstruction and statics estimation into a single joint algorithm. To the best of our knowledge, this is the first time that this has been done. With this approach, NS correction improves interpolation and vice versa.

The proposed algorithm uses rank-reduction-based interpolation and statics correction taken from Kumar et al. (2015) and Alfaraj et al. (2023), respectively. The combined approach not only results in improved reconstruction but also increases the computational efficiency compared with carrying out separate stepwise interpolation and statics correction. Because we need to compute an LR approximation for NS correction, we use it as an accurate *cost-free* initial solution for the interpolation rank-minimization optimization problem. We also lower the number of required transformations between the acquisition and transform domain as interpolation and NS correction operate in the midpoint-offset domain. Moreover, the proposed method requires less preprocessing as it is applied to data without NMO correction. Avoiding NMO correction also preserves the shallow reflectors, which are otherwise destroyed to avoid NMO stretch.

To demonstrate its potential, we apply our method to synthetic and field data affected by complex weathering layers and noise. We show on both data sets that the proposed method can better handle the reconstruction of randomly subsampled data affected by statics resulting in densely sampled data with high resolution.

METHODOLOGY

Promotion of the LR structure requires solving for the NS and subsampling effects, which we do simultaneously without the need of knowledge of the subsurface model resulting in improved reconstruction and computational efficiency.

LR-based NS correction

The LR-based statics estimation and correction iteratively estimates multiscale short-wavelength statics in the midpoint-offset domain (Alfaraj et al., 2023). Using a multiscale approach enables estimation of the statics from multiple frequency bands and ranks without knowledge of the subsurface model (Alfaraj et al., 2021). The method promotes an LR structure, which corresponds to matrices that can be approximated by an LR matrix in the appropriate domain. It is based on the SVD (Golub and Reinsch, 1971):

$$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^H, \tag{2}$$

where $\mathbf{X} \in \mathbb{C}^{n_m \times n_h}$ is a midpoint-offset frequency slice of dimensions n_m (number of midpoints) by n_h (number of offsets). In equation 2, $\mathbf{S} \in \mathbb{R}^{k \times k}$ is the block diagonal matrix containing the nonnegative real-valued singular values $\mathbf{S} = \mathrm{diag}(\sigma_1, \sigma_2, \sigma_3, \ldots, \sigma_k)$, where $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \cdots \geq \sigma_k \geq 0$ and $k = \min\{n_m, n_h\}$, and $\mathbf{U} \in \mathbb{C}^{n_m \times k}$ and $\mathbf{V} \in \mathbb{C}^{n_h \times k}$ are the orthogonal matrices that hold the left $\mathbf{u}$ and right $\mathbf{v}$ singular vectors, respectively, whereas H denotes the Hermitian (conjugate transpose). An LR approximation of $\mathbf{X}$ can be obtained by selecting a few singular vectors that correspond to the few largest singular values. The statics can be estimated by calculating the lag of the maximum crosscorrelation between subsampled data with statics and LR approximated data. The process is performed on band-pass-filtered data in the time domain after multiscale LR approximation of frequency slices. This enables exploiting the relationship among multiple frequency bands and avoiding spurious statics from corrupted frequencies. The method starts with low frequencies for statics estimation as they are less influenced by short-wavelength statics. Then, it applies these statics to the full-band data. As a result, the coherency of the high frequencies is improved, which allows for updating the statics when including them in the next iteration.

The method shows its potential for statics estimation and correction of periodically and densely sampled data (Alfaraj et al., 2023). Figure 3 shows the results after three multiscale iterations when also including randomized subsampling. The data in the time domain become more coherent (Figure 3b and 3e) compared with the incoherent data when we began (Figure 3a and 3d). However, there are imperfections in the estimated statics, which are noticeable when comparing the statics-corrected data (Figure 3b and 3e) with the statics-free data (Figure 3c and 3f). These imperfections are more obvious from the statics error, as shown in Figure 3g and 3h. Next, we investigate whether this error influences data interpolation or not.

Rank-minimization-based interpolation

To interpolate the data, we use rank-minimization-based interpolation with matrix factorization due its computational efficiency (Recht et al., 2010; Kumar et al., 2015). For each midpoint-offset frequency slice $\mathbf{X} \in \mathbb{C}^{n_m \times n_h}$, we solve the following basis pursuit denoising problem:

$$\underset{\mathbf{X}}{\text{minimize}} \; \|\mathbf{X}\|_* \quad \text{subject to} \; \|\mathcal{A}(\mathbf{X}) - \mathbf{b}\|_2 \leq \epsilon, \tag{3}$$

where $\mathbf{b} \in \mathbb{C}^{n_p}$ is the observed frequency slice on the subsampled source-receiver grid $n_{\bar{s}} \times n_{\bar{r}}$ organized as a vector with dimension

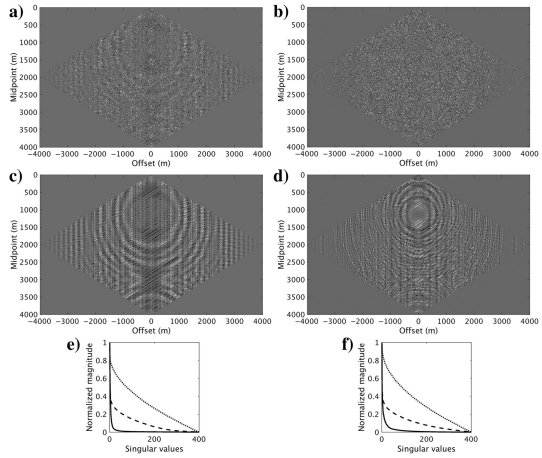

Figure 2. Frequency slices (real part) in the midpoint-offset domain after 75% randomized subsampling of shots (a and b) affected by ±52 ms of nonsurface-consistent statics and (c and d) statics-free data along with (e and f) their singular values decay at (a, c, and e) 18 and (b, d, and f) 49 Hz. Dotted curves correspond to (a and b), dashed curves correspond to (c and d), and solid curves correspond to statics-free frequency slices, as shown in Alfaraj et al. (2023).

DOI:10.1190/geo2023-0668.1

$n_p = n_{\bar{s}} \times n_{\bar{r}} < n_s \times n_r$, ϵ is the noise level, and $\|\mathbf{X}\|_* = \sum_{i=1}^{k} \sigma_i$ is the nuclear norm of $\mathbf{X}$, with σ_i containing the singular values defined in equation 2. The $\mathcal{A}$ is the sampling-transform operator composed of a midpoint-offset transform-domain operator $\mathcal{S}$ and a measurement operator $\mathbf{M}$:

$$\mathcal{A} = \mathbf{M}\mathcal{S}^{\mathsf{T}}, \tag{4}$$

where $\mathcal{S}^{\mathsf{T}}$ (equation 1) transforms the data from the midpoint-offset domain to the source-receiver domain. Therefore, $\mathcal{A}$ maps the data from $\mathbb{C}^{n_m \times n_h} \rightarrow \mathbb{C}^{n_p}$, where $n_p \ll n_m \times n_h$. The transform-domain operator rotates the matrix in the source-receiver domain by 45°, which makes its columns close to linearly dependent and therefore of LR nature. The measurement operator is composed of ones and zeros, which correspond to the measured and unmeasured samples, respectively. For a compressive sensing scenario, the sensing matrix $\mathbf{M}$ aims to destroy the coherency of the data, e.g., with uniform randomized sampling, which results in slowly decaying singular values, as shown in Figure 2e and 2f. To solve equation 3, we use an extension of the spectral-projected gradient ℓ_1 solver (Berg and Friedlander, 2008) that solves a sequence of least absolute shrinkage and selection operator subproblems (Aravkin et al., 2013):

$$\underset{\mathbf{X}}{\text{minimize}} \, \|\mathcal{A}(\mathbf{X}) - \mathbf{b}\|_2 \quad \text{subject to} \quad \|\mathbf{X}\|_* \leq \delta, \tag{5}$$

where $\|\mathbf{X}\|_* \leq \delta$ is the projection on the nuclear norm ball. It requires computation of the singular values with SVD (equation 2) followed by thresholding. Calculation of the singular values at each iteration per frequency slice is computationally demanding for large-scale matrices, which can be avoided by using a more efficient matrix factorization technique (Kumar et al., 2015).

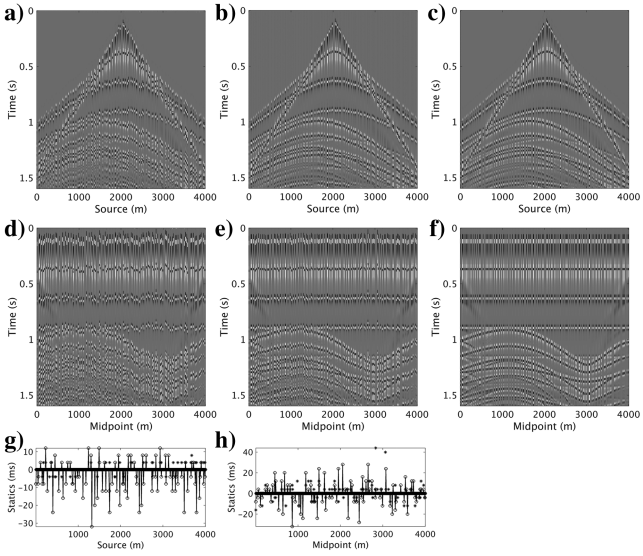

Figure 3. (a–c) Common-receiver gather (CRG) and (d–f) common-offset gather (COG) of (a and d) 75% randomly subsampled data with statics after (b and e) NS correction at the third iteration and (c and f) statics-free data. (g and h) The estimated total statics of the displayed (g) CRG and (h) COG plotted in solid lines with circle markers and their error plotted with asterisks.

Rank-minimization interpolation demonstrates its capability in the reconstruction of marine data, which is usually not influenced by weathering layers. Therefore, they exhibit higher coherency compared with land data. This is shown in Figure 2, in which the LR structure destruction is low for the statics-free subsampled data compared with that influenced by statics. Consequently, interpolation of the incoherent randomly subsampled data with statics leads to a noisy reconstruction (Figure 4a and 4b) because the destruction of the LR structure due to the weathering layers is not considered. In this case, statics correction after interpolation will still be noisy. Interpolation after improving the coherency with the LR-based statics correction provides a better reconstruction (Figure 4c and 4d). However, the reconstruction is erroneous and contains noise as can be seen from the frequency slices and the time-domain gathers (Figures 4c, 4d, 5a, and 5c). Consequently, the reconstruction error is large (Figure 5e and 5g). That is a result of the imperfections in the estimated statics (Figure 3), which are due to the mixed effects of weathering and subsampling on the rank structure. Therefore, without fully accounting for the weathering layers, interpolation becomes challenging, as also indicated by Trad (2009) and Mosher et al. (2017).

Joint interpolation and statics estimation

To reconstruct sparsely acquired data in complex NS regimes, we propose to simultaneously correct for the weathering layers and interpolate the data. For interpolation with equation 3, $\mathbf{b}$ refers to the observed subsampled frequency slice arranged as a vector. Because the LR-based statics estimation and correction operate on matrices in the midpoint-offset domain on the desired grid of densely sampled data, we modify equation 3 to

$$\underset{\mathbf{X}}{\text{minimize}} \, \|\mathbf{X}\|_* \quad \text{subject to} \quad \|\mathcal{M}_{\text{mh}}(\mathbf{X}) - \hat{\mathbf{b}}_{\text{mh}}\|_2 \leq \epsilon, \tag{6}$$

where

$$\mathbf{b}_{\text{mh}} = \mathcal{M}_{\text{mh}}\mathcal{S}(\mathbf{M}^{\mathsf{T}}\mathbf{b}) \tag{7}$$

is the subsampled frequency slice on the desired midpoint-offset grid, $\hat{\mathbf{b}}_{\text{mh}}$ is the same slice but after NS estimation and correction, and $\mathcal{M}_{\text{mh}} = \mathcal{S}^{\circ}\mathbf{M}$ is the measurement operator in the midpoint-offset domain. Compared with equation 3, equation 6 removes the need to apply the adjoint of the transform domain operator $\mathcal{S}^{\mathsf{T}}$ at each iteration. Therefore, because the subsampled measured data after statics correction $\hat{\mathbf{b}}_{\text{mh}}$ are already in the midpoint-offset domain and the optimization problem is solved over many iterations per frequency slice in the same domain, we reduce the computational costs of going back and forth to the transform domain while solving the optimization problem.

The initial solution $\mathbf{X}_0$ in equation 6 is crucial as it determines the number of required iterations. By using LR approximated data as the initial solution,

$$\mathbf{X}_0 = \sum_{j=1}^{k} s^{(j)}\mathbf{u}^{(j)}\mathbf{v}^{(j)H}, \tag{8}$$

where $k \ll \min\{n_m, n_h\}$ and s, $\mathbf{u}$, and $\mathbf{v}$ are the singular values and left, and right singular vectors, respectively, we become closer to the desired solution. This approach decreases the number of iterations compared with starting with an empty or random matrix, as the LR approximation is an interpolator (see Figure 6), but an inaccurate

one when the neglected singular vectors below the selected singular value are of importance. Because LR approximation is already computed for statics estimation, it provides a cost-free initial solution for interpolation. Therefore, the joint approach detailed in Algorithm 1 increases the computational efficiency.

Throughout the algorithm, $(\bar{-})$, $(\hat{\ })$, and $(\tilde{\ })$ indicate the subsampled data, estimated data, and data in the frequency domain, respectively. The subscripts mh and sr represent the data in the midpoint-offset and source-receiver domains, respectively, whereas mh,i and sr,i represent the interpolated data. The inputs to the algorithm are subsampled data in the source-receiver domain $\bar{\mathbf{D}}_{\text{sr}}$ and the edges of the frequency bands $\mathbf{f_b}$ at which we estimate the statics starting from low frequency and increasing the frequency bandwidth as iterations progress. Interpolation and statics estimation require the ranks of all the frequencies as inputs, which we indicate by $\mathbf{k}$ and $\mathbf{K}$, respectively. We will elaborate on their selection in the "Rank selection" subsection of the "Discussion" section. We define n_{f_b} and n_f as the number of frequency bands and number of frequencies, respectively. The outputs of the algorithm are the interpolated and statics-corrected data in the source-receiver $\hat{\mathbf{D}}_{\text{sr,i}}$ and midpoint-offset $\hat{\mathbf{D}}_{\text{mh,i}}$ domains.

In the first step of the algorithm, we apply the transpose of the sampling matrix $\mathbf{M}^{\top}$, which inserts empty rows or columns to the acquired data that will be filled by interpolation. Therefore, the subsampled source-receiver grid $n_{\bar{r}} \times n_{\bar{s}}$ becomes the desired grid of densely sampled data $n_r \times n_s$. Because rank selection is more difficult for data in the time domain, which contains more variability as it encompasses all frequencies, we carry out the LR approximation required for statics correction as well as interpolation in the frequency domain as indicated by the second step. The LR-based statics estimation exploits the relationships among multiple frequency bands. By applying the estimated statics at low frequencies to the full-band data, the coherency of the higher frequencies improves because neighboring frequency bands share common statics. Consequently, the accuracy of their LR approximation increases. After partial statics correction, which renders incoherent energy coherent, the singular values decay increases. This allows for using a smaller rank for more accurate LR approximation to estimate a signal with less statics influence. To do so, the algorithm contains loops over the number of frequency bands n_{f_b} (step 4) and rank scales n_l (step 5). For each frequency slice in a certain frequency band (step 6), we calculate the singular values (step 7), followed by LR approximation to obtain LR approximated data $\tilde{\mathbf{D}}_{\text{lr}}$ (step 8). After LR approximation of all the frequency slices in the current frequency band (step 9), we inverse Fourier transform (iFFT) the LR approximated data matrix $\mathbf{D}_{\text{lr}}$ and data with statics $\mathbf{D}_{\text{mh}}$ to the time domain (step 10), where we estimate the statics matrix $\mathbf{T}_{\text{mh}}$ (step 11) by calculating the lag that corresponds to the maximum crosscorrelation coefficient with

$$t_{\text{mh}} = \arg\max_t(\mathbf{d}_{\text{mh}}(t) \star \mathbf{d}_{\text{lr}}(t)), \tag{9}$$

where $\star$ indicates crosscorrelation, the notation (t) is used to show time dependence, and t_{mh} is the time shift of one trace estimated from data with statics $\mathbf{d}_{\text{mh}}$ and LR approximated data $\mathbf{d}_{\text{lr}}$. After estimating $\mathbf{T}_{\text{mh}}$ from all traces $\mathbf{D}_{\text{mh}}$ and $\mathbf{D}_{\text{lr}}$, they are used for statics correction of the full-band data (step 12) to estimate the statics-corrected data matrix $\hat{\mathbf{D}}_{\text{mh}}$, which for one trace is estimated as follows:

$$\hat{\mathbf{d}}_{\text{mh}} = \mathbf{d}_{\text{mh}}(t + t_{\text{mh}}). \tag{10}$$

The subsampled data with statics are then updated to be the statics-corrected data (step 13). After partial statics correction, the subsampled data are ready for interpolation as they exhibit higher coherency. After the multirank-scale iterations (step 14), we solve equation 6 to interpolate the statics-corrected frequency slices that belong to the current frequency band, as indicated in steps 15 and 16, where k is the rank used for interpolation.

Because the LR structure destruction is not only a result of the statics but also of subsampling (Figure 2), LR approximation of subsampled data cannot preserve data without the influence of statics. We need more accurate LR approximation that can preserve the statics-free data, which requires data with higher coherency. Because randomized subsampling decreases coherency, we use the interpolated frequencies, which now exhibit faster singular values decay, to allow for more accurate LR approximation. We repeat the statics estimation and correction, but using the interpolated data matrices (steps 17–20). The estimated statics $\mathbf{T}_{\text{mh,i}}$ are then used for statics correction of the interpolated and subsampled (to be used for statics estimation at the next iteration and to be interpolated) data $\hat{\mathbf{D}}_{\text{mh,i}}$ and $\mathbf{D}_{\text{mh}}$, respectively. Therefore, statics correction improves interpolation and vice versa. Consequently, reconstruction with the proposed joint scheme provides better results with less noise compared with the stepwise one (Figure 4). Finally, the data can be transformed to the source-receiver and time domains

Algorithm 1. Joint rank-reduction-based interpolation and statics estimation.

Input: $\bar{\mathbf{D}}_{\text{sr}} \in \mathbb{R}^{n_t \times n_{\bar{r}} \times n_{\bar{s}}}$, $\mathbf{k} \in \mathbb{N}^{n_f}$, $\mathbf{K} \in \mathbb{N}^{n_f \times n_l}$, $\mathbf{f_b} \in \mathbb{N}^{n_{f_b}-1}$
Output: $\hat{\mathbf{D}}_{\text{mh,i}} \in \mathbb{R}^{n_t \times n_h \times n_m}$, $\hat{\mathbf{D}}_{\text{sr,i}} \in \mathbb{R}^{n_t \times n_r \times n_s}$
1 Apply $\mathbf{M}^{\top}$ to each time slice of $\bar{\mathbf{D}}_{\text{sr}}$ to obtain $\mathbf{D}_{\text{sr}} \in \mathbb{R}^{n_t \times n_r \times n_s}$
2 Transform $\mathbf{D}_{\text{sr}}$ to $\mathbf{D}_{\text{mh}} \in \mathbb{R}^{n_t \times n_h \times n_m}$ with equation 1
3 Fast Fourier transform (FFT) $\mathbf{D}_{\text{mh}}$ to $\tilde{\mathbf{D}}_{\text{mh}} \in \mathbb{C}^{n_m \times n_h \times n_f}$
4 **for** $i \leftarrow 1$ **to** $n_{f_b} - 1$ **do**
5 **for** $l \leftarrow 1$ **to** n_l **do**
6 **for** $f \leftarrow \mathbf{f_b}^{(i)} + 1$ **to** $\mathbf{f_b}^{(i+1)}$ **do**
7 Calculate SVD for $\tilde{\mathbf{D}}_{\text{mh}}^{(f)}$ with equation 2
8 $\tilde{\mathbf{D}}_{\text{lr}}^{(f)} \leftarrow \sum_{j=1}^{\mathbf{K}^{(f,l)}} s^{(j)} \mathbf{u}^{(j)} \mathbf{v}^{(j)H}$
9 **if** $f = \mathbf{f_b}^{(i+1)}$ **then**
10 iFFT $\tilde{\mathbf{D}}_{\text{lr}}$ to $\mathbf{D}_{\text{lr}}$, iFFT $\tilde{\mathbf{D}}_{\text{mh}}$ to $\mathbf{D}_{\text{mh}}$
11 $\mathbf{T}_{\text{mh}} \leftarrow \mathcal{C}(\mathbf{D}_{\text{mh}}, \mathbf{D}_{\text{lr}})$ using equation 9
12 $\hat{\mathbf{D}}_{\text{mh}} \leftarrow \tau(\mathbf{D}_{\text{mh}}, \mathbf{T}_{\text{mh}})$ using equation 10
13 $\mathbf{D}_{\text{mh}} \leftarrow \hat{\mathbf{D}}_{\text{mh}}$, FFT $\mathbf{D}_{\text{mh}}$ to $\tilde{\mathbf{D}}_{\text{mh}}$
14 **if** $l = n_l$ **then**
15 **for** $f \leftarrow \mathbf{f_b}^{(i)} + 1$ **to** $\mathbf{f_b}^{(i+1)}$ **do**
16 $\hat{\tilde{\mathbf{D}}}_{\text{mh,i}}^{(f)} \leftarrow$ solution of equation 6: $\hat{\mathbf{b}}_{\text{mh}} \leftarrow \mathcal{M}_{\text{mh}}(\tilde{\mathbf{D}}_{\text{mh}}^{(f)})$, $\mathbf{X}_0 \leftarrow \tilde{\mathbf{D}}_{\text{lr}}^{(f)}$, $k \leftarrow \mathbf{k}^{(f)}$
17 Repeat steps 7 and 8 with $\hat{\tilde{\mathbf{D}}}_{\text{mh,i}}$ instead of $\tilde{\mathbf{D}}_{\text{mh}}$ to obtain $\tilde{\mathbf{D}}_{\text{lr,i}}$
18 Repeat steps 10 and 11 with $\tilde{\mathbf{D}}_{\text{lr,i}}$ and $\hat{\tilde{\mathbf{D}}}_{\text{mh,i}}$ instead of $\tilde{\mathbf{D}}_{\text{lr}}$ and $\tilde{\mathbf{D}}_{\text{mh}}$, respectively, to estimate $\mathbf{T}_{\text{mh,i}}$
19 Apply $\mathbf{T}_{\text{mh,i}}$ to $\hat{\mathbf{D}}_{\text{mh,i}}$ and $\mathbf{D}_{\text{mh}}$
20 Repeat step 13
21 Transform $\hat{\mathbf{D}}_{\text{mh,i}}$ from the midpoint-offset domain to $\hat{\mathbf{D}}_{\text{sr,i}}$ the source-receiver domain

DOI:10.1190/geo2023-0668.1

(steps 21 and 22). The output data are a collection of all the processed frequency bands. The time shifts, which are frequency-band and rank-scale dependent, also can be saved for analysis. We further analyze the results in the next section.

RESULTS

Using synthetic and field data, we further demonstrate the potential of our proposed reconstruction method.

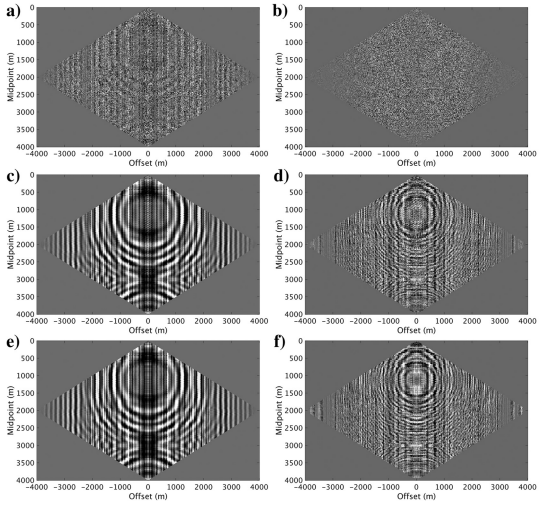

Figure 4. Midpoint-offset frequency slices after (a and b) interpolation without statics correction, (c and d) stepwise reconstruction, and (e and f) proposed reconstruction at (a, c, and e) 18 and (b, d, and f) 49 Hz.

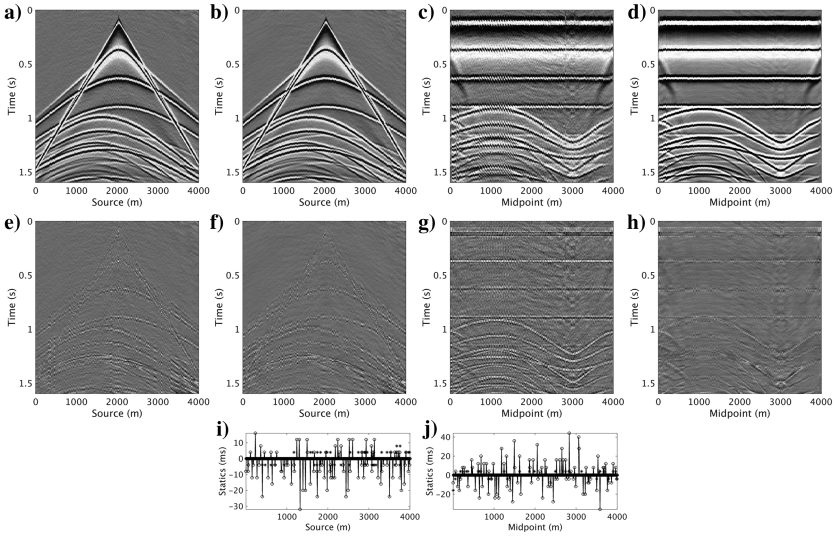

Figure 5. (a, b, e, and f) The CRG and (c, d, g, and h) COG after the (a and c) stepwise and (b and d) proposed reconstruction of 75% randomly subsampled data with statics along with (e and g) and (f and h) their corresponding residual amplitudes, respectively. (i and j) For the CRG displayed in (b) and the COG displayed in (d), respectively, the estimated statics with the proposed reconstruction plotted in solid lines with circle markers and their error plotted with asterisks, respectively.

Synthetic data

As shown in Figure 5a and 5c, imperfect statics correction results in poor reconstruction and large error (Figure 5e and 5g). In contrast, the proposed reconstruction, which simultaneously corrects for the NS effects and interpolates the data (Algorithm 1), provides improved results with less noise and higher accuracy, as shown in Figure 5b and 5d. Therefore, by using the interpolated frequencies, we can improve the estimated statics as iterations progress. This can be observed from the error of the estimated statics, which is minimal for the proposed reconstruction method (Figure 5i and 5j) compared with the stepwise one (Figure 3g and 3h). Consequently, the reconstruction amplitude error is lower for the former compared with the latter (see Figure 5g and 5h). The reconstruction error of each frequency slice relative to the statics-free densely sampled one shown in Figure 7 also confirms that the joint approach improves the reconstruction as its error is lower compared with the stepwise one. The same figure also shows that interpolation without statics correction leads to a large error, which makes statics estimation and correction after interpolation challenging.

Field data

To further examine the potential of the proposed joint reconstruction, we test its performance on field data. The field data set is affected by complex weathering layers, which include loose sand, fast carbonates, karsts, and gravel with surface elevation changes of approximately 100 m across the line (Al-Ali and Verschuur, 2006). The data also contain random noise and residual ground roll with different characteristics due to the different NS conditions. In this case, the NS on its own is challenging to data processing as can be seen from the gathers and stack, which contain poorly continuous reflectors affected by statics (Figures 8a, 9a, and 10a). Here, NMO correction is only used for display but not during the reconstruction process. In addition, the data set is acquired with 30 m source and receiver intervals on a crooked line, which we further decimate by removing 50% of the sources with uniform randomized subsampling (Figures 8b, 9b, and 10b). Such subsampling equates to 60 m periodic shot spacing on average, which is a frugal acquisition design. The minimum and maximum source-receiver offsets are 15 and 3585 m, respectively. Using few upholes, which are inadequate to compensate for the NS variations, we apply elevation statics correction and bring the sources and receivers to a flat datum. From the stack of the subsampled data (Figure 10a), it is obvious that interpretation is challenging, which calls for data reconstruction.

To reconstruct the data, we apply rank-minimization-based interpolation without prior statics correction, which results in the data and stacks shown in Figures 8c, 9c, and 10c. Interpolation in this situation provides an improvement compared with the subsampled data (Figures 8b, 9b, and 10b). However, by not accounting for rapid variations in the weathering layers, the interpolated data exhibit more noise compared with the original data (Figures 8a, 9a, and 10a). In addition, the events between 19 and 24 km and

0.4 and 0.6 s are distorted. To improve the results, we use the proposed reconstruction (Algorithm 1). The estimated statics for the displayed gathers are shown in Figure 8e, whereas the estimated statics for the whole data are shown in Figure 10e. The latter shows that larger statics are estimated from data between 19 and 24 km, which contain more complexity in the weathering layers compared with the data between 15 and 18 km (see Figure 10a). Reconstruction with our proposed method (Figures 8d, 9d, and 10d) reduces the noise compared with interpolation without statics correction. Moreover, it leads to higher power and more continuity of the events compared with the original data, as demonstrated by the stack section in Figure 10d. The improvement noticed on the stack between 19 and 24 km also coincides with the observation that most of the estimated statics (Figure 10e) belong to the same section.

Table 1. Sensitivity of the proposed method to rank selection of the field data.

Case scenario	Change in interpolation ranks (%)	Change in statics estimation ranks (%)	Stack power (%)
1a	+25	0	+9
1b	0	+25	+17
1c	+25	+25	+9
2a	−25	0	+22
2b	0	−25	+21
2c	−25	−25	+19

The change in the interpolation and statics correction rank values is with respect to the chosen values stated in the "Rank selection" subsection of the "Discussion" section. The stack power column indicates the power of the estimated stack relative to the stack of densely sampled data.

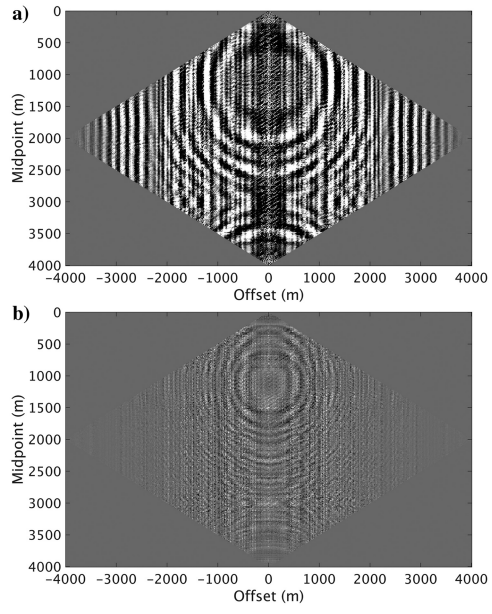

Figure 6. Interpolation with LR approximation: LR approximation after partial statics correction of Figure 2a and 2b, which we use as a starting solution $\mathbf{X}_0$ of equation 6 for interpolation. Frequency slices (real part) at (a) 18 and (b) 49 Hz.

DISCUSSION

In this section, we discuss the use of data windowing, rank selection with sensitivity analysis, computational efficiency, limitations, and extensions.

Data windowing

It is common for rank-based reconstruction techniques to rely on data windowing to ensure that the data have sufficient LR structure. Moreover, the number of different dips in each window can be used as a guide for rank selection (see Oropeza and Sacchi, 2011). However, selection of the windowing parameters is not trivial. In the presence of weathering layers, merging windows after statics correction can be a challenge as the separate windows may get assigned different statics that may result in mis-ties along the horizontal and vertical directions. This could be the case if windowing is used to reconstruct the field data example, where the section between 15 and 18 km exhibits low influence of the weathering layers compared with the one between 19 and 24 km (Figure 10). To avoid the nontrivial task of data windowing for statics correction and interpolation, we use the midpoint-offset-frequency domain, where the data can be well approximated by LR matrices. Therefore, we can successfully apply the proposed reconstruction techniques to the whole line without windowing, which is demonstrated by the synthetic and field data examples. Not using windows can result in large matrices. If that leads into computational limitations, which we do not suffer from for the shown examples, data windowing can become necessary.

Rank selection

The main parameters that need to be selected for Algorithm 1 are the ranks used for interpolation $\mathbf{k}$ of all the frequencies and those used for multiscale LR approximation $\mathbf{K}$. After partial statics correction, the singular values decay faster because the data exhibit higher coherency. Because we do not account for all the statics at the first iteration, there will remain incoherent energy, which is represented by the smaller singular values. Therefore, we can neglect singular values in the tail of the decay curve to capture the coherent energy, which we use for statics estimation and correction. The further the iterations progress, the faster the singular values will decay. Accordingly, we begin the iterations with a high-rank approximation and further decrease the rank as iterations progress. For interpolation, we need to preserve as much as possible of the signal, whereas for statics estimation, we only need a signal with less statics imprint that we use for

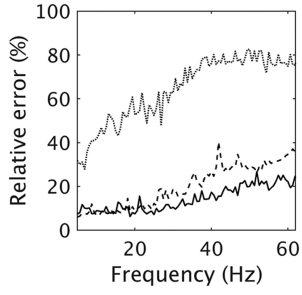

Figure 7. Error of reconstruction of the 75% randomly subsampled data with statics per frequency slice using interpolation (dotted curve), stepwise reconstruction (dashed curve), and proposed reconstruction (solid curve).

DOI:10.1190/geo2023-0668.1

crosscorrelation. The latter tends to be captured by the few largest singular values. Because amplitude preservation is essential after interpolation, the ranks used for interpolation are higher than those needed for statics estimation. Because the high frequencies contain more variability compared with the low ones, they usually require higher ranks. Therefore, we suggest to start with LR for low frequencies and linearly increase it with increasing frequency content for interpolation (Kumar et al., 2015) and statics estimation (Alfaraj et al., 2023). User analysis is necessary for the determination of the optimal ranks because complexity of the frequency slices also play a role.

We choose the lowest and highest ranks that correspond to the lowest and highest frequencies of the synthetic data according to the preceding strategy. The rank values that we use for interpolation and LR approximation at the first, second, and third iterations are (15–75), (15–30), (5–15), and (3–5), respectively. We determine the rank values for the in-between frequencies by linearly increasing the rank between the lowest and highest ranks in each iteration. To quantify the influence of the rank selection, we use the stack power as a measure. The stack power increases by 202% and decreases by 13% compared with the densely sampled data with statics and statics-free data, respectively. Increasing the rank values by 50% leads to almost the same results ($\pm 0.5\%$), while decreasing the rank values by 50% results in 210% and -10% change in the stack power compared with the stacks of densely sampled data with statics and statics-free data, respectively. Therefore, the sensitivity of the joint reconstruction to the rank selection of this synthetic data is low.

For the field data, we follow the same strategy to select the rank values. The lowest and highest rank values that we use for interpolation and statics correction at the first, second, and third iterations are $(30-90), (5-15), (3-15)$, and $(3-5)$, respectively. These values lead to 19% increase in the stack power after applying our proposed joint reconstruction compared with the densely sampled data (compare Figure 10a and 10d). To examine the method's sensitivity for the rank selection, we test several scenarios outlined in Table 1. We notice that varying the rank values used for interpolation has higher impact on the stack power, surpassing the impact of altering the ranks used for static estimation. This conclusion is expected as the rank values used for interpolation explicitly estimate new data (phase and amplitude). In contrast, the LR-based statics estimation uses the rank values to estimate the time shifts that increase the coherency of the data. Moreover, the multiscale approach implemented for statics estimation alleviates the need of using accurate rank values. Despite the sensitivity of the method to the interpolation ranks, it remains low as increasing and decreasing the rank values by 25% (case scenarios 1a and 2a) leads to only 10% decrease and 3% increase in the stack power compared with the originally selected ranks.

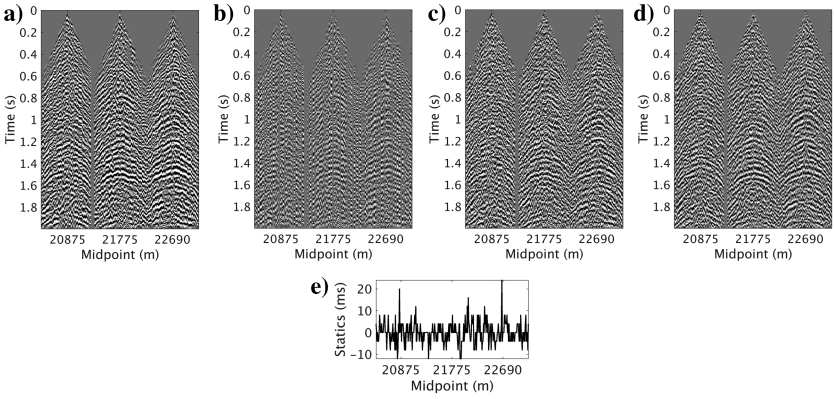

Figure 8. Three selected CMP gathers from the field data: (a) originally acquired data and data after (b) 50% randomized shots subsampling, (c) interpolation without statics correction, and (d) proposed reconstruction along with (e) the estimated statics of the displayed gathers.

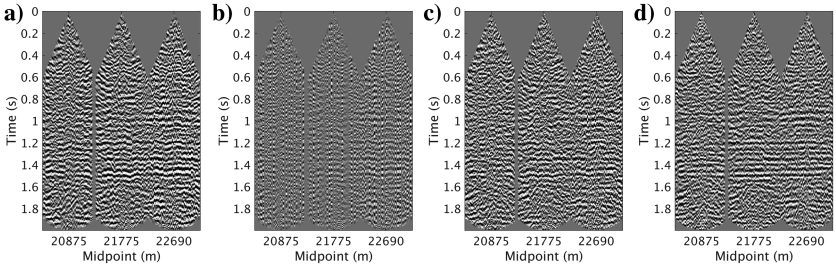

Figure 9. Three selected CMP gathers from the field data after NMO correction: (a) originally acquired data and data after (b) 50% randomized shots subsampling, (c) rank-minimization interpolation, and (d) our proposed reconstruction.

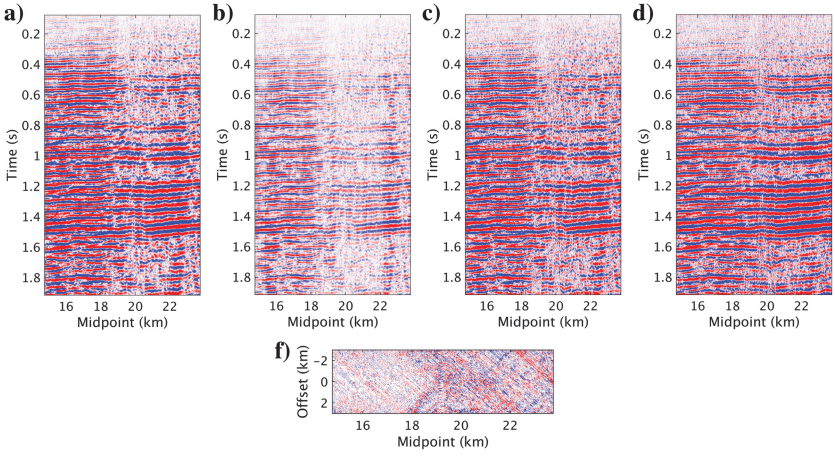

Figure 10. Field data stacks of the (a) original data and after (b) 50% randomized subsampling, (c) interpolation without statics correction, and (d) our proposed reconstruction along with (e) the estimated statics clipped to ± 12 ms.

Computational efficiency

Other than obtaining improved reconstruction, the proposed method provides better efficiency compared with stepwise reconstruction. The

computational efficiency of the proposed reconstruction is related to the computational efficiency of the separate processes themselves. For LR-based statics estimation and correction, it is determined by the numbers of LR approximations $O(\min\{n_m^2 \times n_h, n_m \times n_h^2\})$ and crosscorrelations $O(n_t \times n_g)$, where n_g is the number of crosscorrelation lags, which turned out to be more efficient compared with residual statics estimation by stack power maximization (Ronen and Claerbout, 1985). Truncated SVD, where only a subset of the largest singular vectors and singular values are calculated, can speed up the LR approximation, which can be necessary for 3D large-scale seismic data. The efficiency of rank-minimization-based interpolation, which operates on the whole line, can be better than those that operate on separate gathers. To further enhance the computational efficiency, LR approximation and interpolation of frequency slices as well as the crosscorrelation of traces can be performed in parallel over multiple cores.

The joint reconstruction approach further improves the computational efficiency. As shown in Figure 6, LR approximation can interpolate the randomly subsampled data displayed in Figure 2a and 2b. However, this interpolation is poor as the approximation neglects data of importance. Nonetheless, it provides a good initial solution that we can use to reduce the number of required iterations to solve equation 3. Because LR approximation is a byproduct of LR-based statics estimation and correction, we obtain the initial guess cost-free with our proposed joint reconstruction. Moreover, because statics estimation and correction is performed in the midpoint-offset domain, we also modify the interpolation such that the input measured data is used in the same domain (equation 6). Therefore, we reduce the requirement to go back and forth between the acquisition and midpoint-offset domains. In addition, our model-independent approach requires less manual interaction compared with existing techniques, which usually need multiple iterations of velocity and statics estimation as they influence each other, followed by data interpolation of NMO-corrected data (Trad, 2009).

Limitations and extensions

Different NS conditions lead to different effects. Even for the field data example, the data between 15 and 18 km are less influenced by the NS weathering layers compared with the data between 18 and 24 km (Figure 10a). Therefore, interpolation of the former section can be achieved without the need of any statics correction (Figure 10c), which is not the case for the latter one that becomes improved with our proposed method (Figure 10d). Even though the data contain noise that we do not remove as part of the preconditioning, the proposed reconstruction does not suffer. Further improvement could be obtained with prior noise attenuation or by adding a noise reduction step as part of the joint reconstruction scheme because noise can influence the coherency of the data. In the used field data example, ground roll is aliased due to the coarse 30 m source and receiver intervals. Therefore, because it is challenging to interpolate, we remove it prior to data decimation with a simple frequency-wavenumber filter.

In terms of sampling schemes, we only consider randomized sampling. There are multiple ways to optimize the locations of sources and receivers, e.g., see Hennenfent and Herrmann (2008) and recent studies by Guo and Sacchi (2020) and Zhang et al. (2022). However, there are other limiting factors for land data acquisition design such as the NS properties, which needs to be considered. When the data are periodically subsampled, the proposed

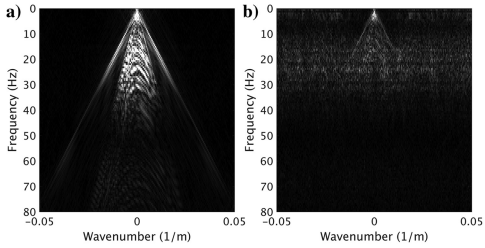

Figure 11. Frequency-wavenumber amplitude spectrum of a selected CRG from the synthetic data: (a) statics-free and (b) after adding statics and removing 50% of the shots at random.

method can be modified by replacing the LR-based interpolator in the midpoint-offset domain with another one that is more suitable for interpolation of periodically subsampled data. Similarly, other interpolation methods for randomly subsampled data, e.g., curvelet-, Fourier-, or Radon-based methods, can replace the used rank-minimization interpolation, if desired.

Interpolation and statics correction can be performed on separate gathers, e.g., by applying sparsity promotion on receiver gathers in the frequency-wavenumber domain because statics and randomized subsampling decrease the sparsity (Figure 11b) that densely sampled statics-free data exhibit (Figure 11a). As noticed, the low frequencies are less influenced by the subsampling and NS effects compared with the higher frequencies (Figure 11b). Therefore, a similar scheme to Algorithm 1 can be implemented, where the statics estimated from subsampled and interpolated low frequencies and applied to the full-band data allow for statics estimation and reconstruction of the higher frequencies as they improve their coherency. However, operating on the full line in the midpoint-offset domain allows us to exploit structural relationships between the different CMPs, which cannot be achieved using separate gathers. Interpolation of separate CMP gathers still allows us to exploit structural relationships up to a certain extent compared with separate shot or receiver gathers. However, the separate CMP gathers are not linked in this case. Therefore, there might be variability across the line, which is minimized with our proposed approach.

To extend the proposed reconstruction method to 3D data, Algorithm 1 can be modified by incorporating a suitable transform-revealing domain. Luckily, there are different options that can be investigated to map the 3D data into 2D matrices, which include the midpoint-offset domain along the x- and y-directions and the source and receiver coordinates along the x- and y-directions or the x- and y-coordinates of the sources and receivers. It is essential that the chosen domain reveals the necessary properties that allow for LR approximation and rank-minimization-based interpolation.

CONCLUSION

Complex NS conditions with rapid variations in properties of the weathering layers degrade the quality of subsurface models of interest. The addition of acquisition gaps, e.g., to reduce the acquisition costs, exacerbates this effect. Interpolation without accounting for the NS effects fails to provide satisfactory results. The conventional reconstruction approach with surface-consistent short-wavelength statics correction followed by interpolation requires distinction between primaries and multiples and NMO velocity estimation, which are challenging to perform with subsampled data influenced by the

DOI:10.1190/geo2023-0668.1

weathering layers. Stepwise reconstruction with model-independent LR-based statics correction followed by rank-minimization interpolation improves the reconstruction. However, it results in noisy and erroneous reconstruction due to imperfections in the estimated statics. These imperfections are due to the combined influence of weathering and subsampling, which reduce the coherency and destroy the LR structure. To overcome that, we propose to reconstruct the data with simultaneous rank-reduction-based statics correction and interpolation. With this approach, statics correction and interpolation improve each other's performance, as data with improved coherency from one are fed into the other. In addition, we increase the computational efficiency by avoiding repetitive transformations between the source-receiver and midpoint-offset domains and by using LR approximated data (cost-free product from the LR-based statics estimation) as initial solution to the rank-minimization optimization problem. The application of our proposed method to randomly subsampled synthetic data and a challenging field data set affected by complex weathering layers demonstrates its potential.

DATA AND MATERIALS AVAILABILITY

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

REFERENCES

Al-Ali, M., and D. Verschuur, 2006, An integrated method for resolving the seismic complex near-surface problem: Geophysical Prospecting, **54**, 739–750, doi: 10.1111/j.1365-2478.2006.00575.x.

Alfaraj, A. M., F. Herrmann, and R. Kumar, 2017, Reconstruction of S-waves from low-cost randomized and simultaneous acquisition by joint sparse inversion: 87th Annual International Meeting, SEG, Expanded Abstracts, 2533–2538, doi: 10.1190/segam2017-17681376.1.

Alfaraj, A. M., E. Verschuur, and F. J. Herrmann, 2021, Residual statics correction without NMO — A rank-based approach: First International Meeting for Applied Geoscience & Energy, SEG, Expanded Abstracts, 2565–2569, doi: 10.1190/segam2021-3583455.1.

Alfaraj, A. M., D. Verschuur, and F. J. Herrmann, 2023, Low-rank-based residual statics estimation and correction: Geophysics, **88**, no. 3, V215–V231, doi: 10.1190/geo2022-0415.1.

Aravkin, A. Y., J. V. Burke, and M. P. Friedlander, 2013, Variational properties of value functions: SIAM Journal on Optimization, **23**, 1689–1717, doi: 10.1137/120899157.

Aravkin, A. Y., R. Kumar, H. Mansour, B. Recht, and F. J. Herrmann, 2014, Fast methods for denoising matrix completion formulations, with applications to robust seismic data interpolation: SIAM Journal on Scientific Computing, **36**, S237–S266, doi: 10.1137/130919210.

Berg, E. V., and M. P. Friedlander, 2008, Probing the Pareto Frontier for basis pursuit solutions: SIAM Journal on Scientific Computing, **31**, 890–912, doi: 10.1137/080714488.

Candès, E. J., and M. B. Wakin, 2008, An introduction to compressive sampling: IEEE Signal Processing Magazine, **25**, 21–30, doi: 10.1109/MSP.2007.914731.

Cheng, J., and M. Sacchi, 2016, Fast and memory-efficient singular spectrum analysis for seismic data reconstruction and denoising: 86th Annual International Meeting, SEG, Expanded Abstracts, 4064–4068, doi: 10.1190/segam2016-13955076.1.

Cox, M., 1999, Static corrections for seismic reflection surveys: SEG.

Donoho, D. L., 2006, Compressed sensing: IEEE Transactions on Information Theory, **52**, 1289–1306, doi: 10.1109/TIT.2006.871582.

Gholami, A., 2013, Residual statics estimation by sparsity maximization: Geophysics, **78**, no. 1, V11–V19, doi: 10.1190/geo2012-0035.1.

Gholami, A., 2014, Phase retrieval through regularization for seismic problems: Geophysics, **79**, no. 5, V153–V164, doi: 10.1190/geo2013-0318.1.

Golub, G. H., and C. Reinsch, 1971, Singular value decomposition and least squares solutions, *in* F. L. Bauer, A. S. Householder, F. W. J. Olver, H. Rutishauser, K. Samelson, and E. Stiefel, eds., Linear algebra: Springer, 134–151.

Guo, Y., and M. D. Sacchi, 2020, Data-driven time-lapse acquisition design via optimal receiver-source placement and reconstruction: 90th Annual International Meeting, SEG, Expanded Abstracts, 66–70, doi: 10.1190/segam2020-3426928.1.

Hennenfent, G., and F. J. Herrmann, 2008, Simply denoise: Wavefield reconstruction via jittered undersampling: Geophysics, **73**, no. 3, V19–V28, doi: 10.1190/1.2841038.

Herrmann, F. J., M. P. Friedlander, and O. Yilmaz, 2012, Fighting the curse of dimensionality: Compressive sensing in exploration seismology: IEEE Signal Processing Magazine, **29**, 88–100, doi: 10.1109/MSP.2012.2185859.

Kabir, M. N., and D. Verschuur, 1995, Restoration of missing offsets by parabolic Radon transform: Geophysical Prospecting, **43**, 347–368, doi: 10.1111/j.1365-2478.1995.tb00257.x.

Keho, T. H., and P. G. Kelamis, 2012, Focus on land seismic technology: The near-surface challenge: The Leading Edge, **31**, 62–68, doi: 10.1190/1.3679329.

Kreimer, N., and M. D. Sacchi, 2012, A tensor higher-order singular value decomposition for prestack seismic data noise reduction and interpolation: Geophysics, **77**, no. 3, V113–V122, doi: 10.1190/geo2011-0399.1.

Kumar, R., C. D. Silva, O. Akalin, A. Y. Aravkin, H. Mansour, B. Recht, and F. J. Herrmann, 2015, Efficient matrix completion for seismic data reconstruction: Geophysics, **80**, no. 5, V97–V114, doi: 10.1190/geo2014-0369.1.

Mosher, C., C. Li, Y. Ji, F. D. Janiszewski, B. Bankhead, L. Williams, J. Hand, and J. Anderson, 2017, Compressive seismic imaging: Moving from research to production: 87th Annual International Meeting, SEG, Expanded Abstracts, 74–78, doi: 10.1190/segam2017-17679803.1.

Oropeza, V., and M. Sacchi, 2011, Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis: Geophysics, **76**, no. 3, V25–V32, doi: 10.1190/1.3552706.

Recht, B., M. Fazel, and P. Parrilo, 2010, Guaranteed minimum rank solutions to linear matrix equations via nuclear norm minimization: SIAM Review, **52**, 471–501, doi: 10.1137/070697835.

Recht, B., and C. Ré, 2013, Parallel stochastic gradient algorithms for large-scale matrix completion: Mathematical Programming Computation, **5**, 201–226, doi: 10.1007/s12532-013-0053-8.

Ronen, J., and J. F. Claerbout, 1985, Surface-consistent residual statics estimation by stack-power maximization: Geophysics, **50**, 2759–2767, doi: 10.1190/1.1441896.

Stanton, A., N. Kazemi, and M. D. Sacchi, 2013, Processing seismic data in the presence of residual statics: 83rd Annual International Meeting, SEG, Expanded Abstracts, 1838–1842, doi: 10.1190/segam2013-1453.1.

Taner, M. T., F. Koehler, and K. A. Alhilali, 1974, Estimation and correction of near-surface time anomalies: Geophysics, **39**, 441–463, doi: 10.1190/1.1440441.

Thorbecke, J. W., and D. Draganov, 2011, Finite-difference modeling experiments for seismic interferometry: Geophysics, **76**, no. 6, H1–H18, doi: 10.1190/geo2010-0039.1.

Trad, D., 2009, Five-dimensional interpolation: Recovering from acquisition constraints: Geophysics, **74**, no. 6, V123–V132, doi: 10.1190/1.3245216.

Trad, D. O., T. J. Ulrych, and M. D. Sacchi, 2002, Accurate interpolation with high-resolution time-variant Radon transforms: Geophysics, **67**, 644–656, doi: 10.1190/1.1468626.

Trickett, S., L. Burroughs, A. Milton, L. Walton, and R. Dack, 2010, Rank-reduction-based trace interpolation: 80th Annual International Meeting, SEG, Expanded Abstracts, 3829–3833, doi: 10.1190/1.3513645.

Trickett, S. R., 2003, F-xy eigenimage noise suppression: Geophysics, **68**, 751–759, doi: 10.1190/1.1567245.

Wiggins, R. A., K. L. Larner, and R. D. Wisecup, 1976, Residual statics analysis as a general linear inverse problem: Geophysics, **41**, 922–938, doi: 10.1190/1.1440672.

Yilmaz, O., 2001, Seismic data analysis: SEG.

Zhang, Y., M. Louboutin, A. Siahkoohi, Z. Yin, R. Kumar, and F. J. Herrmann, 2022, A simulation-free seismic survey design by maximizing the spectral gap: Second International Meeting for Applied Geoscience & Energy, SEG, Expanded Abstracts, 15–20, 10.1190/image2022-3751690.1.

Zhang, Y., S. Sharan, and F. J. Herrmann, 2019, High-frequency wavefield recovery with weighted matrix factorizations: 89th Annual International Meeting, SEG, Expanded Abstracts, 3959–3963, doi: 10.1190/segam2019-3215103.1.

Zwartjes, P. M., and M. D. Sacchi, 2007, Fourier reconstruction of nonuniformly sampled, aliased seismic data: Geophysics, **72**, no. 1, V21–V32, doi: 10.1190/1.2399442.

Biographies and photographs of the authors are not available.