Seismic data interpolation without iteration using t-x-y streaming prediction filter with varying smoothness[a]

[a]Published in Geophysics, 87(1), V29-V38, (2022)

Yang Liu, Geng Wu*, Zhisheng Zheng**

ABSTRACT

Although there is an increase in the amount of seismic data acquired with wide-azimuth geometry, it is difficult to achieve regular data distributions in spatial directions owing to limitations imposed by the surface environment and economic factor. To address this issue, interpolation is an economical solution. The current state of the art methods for seismic data interpolation are iterative methods. However, iterative methods tend to incur high computational cost which restricts their application in cases of large, high-dimensional datasets. Hence, we developed a two-step non-iterative method to interpolate nonstationary seismic data based on streaming prediction filters (SPFs) with varying smoothness in the time-space domain; and we extended these filters to two spatial dimensions. Streaming computation, which is the kernel of the method, directly calculates the coefficients of nonstationary SPF in the overdetermined equation with local smoothness constraints. In addition to the traditional streaming prediction-error filter (PEF), we proposed a similarity matrix to improve the constraint condition where the smoothness characteristics of the adjacent filter coefficient change with the varying data. We also designed non-causal in space filters for interpolation by using several neighboring traces around the target traces to predict the signal; this was performed to obtain more accurate interpolated results than those from the causal in space version. Compared with Fourier Projection onto a Convex Sets (POCS) interpolation method, the proposed method has the advantages such as fast computational speed and nonstationary event reconstruction. The application of the proposed method on synthetic and nonstationary field data showed that it can successfully interpolate high-dimensional data with low computational cost and reasonable accuracy even in the presence of aliased and conflicting events.

INTRODUCTION

In seismic exploration, high-quality seismic data acquisition is a key ingredient in creating accurate subsurface interpretations, due to cost and access limitations, it is often impossible to achieve ideal surface sampling of sources and receivers where all spatial directions are well sampled. Therefore, there is often a need to regularize and interpolate the recorded seismic data at an early stage of the seismic processing workflow.

In past decades, several methods for seismic data interpolation have been proposed, and these include two major categories, which are based on the theories of wave dynamics and image analysis. According to the physical characteristics of seismic wave propagation, different types of integral continuous operators have proved their effectiveness for seismic data interpolation, such as shot continuation operators (Mazzucchelli et al., 1999) and offset continuation operators (Fomel, 2003). Interferometry is also used for interpolation of missing seismic data (Wang et al., 2009). Recently, new theories in signal processing such as compressive sensing (CS) and machine learning (ML) have shown great potential in reconstructing seismic traces. CS-based methods for data interpolation assume that seismic data obey sparsity when transformed to an appropriate domain, such as curvelet transform (Herrmann and Hennenfent, 2008; Naghizadeh and Sacchi, 2010; Yang and Gao, 2012; Shahidi et al., 2013), Radon transform (Jager et al., 2002; Shao et al., 2017), dreamlet transform (Wang et al., 2015), and seislet transform (Liu and Fomel, 2010; Liu et al., 2015). Moreover, application of machine learning in seismic exploration is a hot topic and it is expected to aid in interpolating missing seismic traces. Jia et al. (2018) used a Monte Carlo method for intelligent interpolation to reduce the cost of training sets. Wang et al. (2019) designed an eight-layer residual learning networks (ResNets) for regularly missing data reconstruction. Wang et al. (2020) explored a convolutional auto-encoder (CAE) method for interpolating irregularly sampled shot gathers by introducing transfer learning strategy. Machine learning methods depend on the characteristics of the training data, which can overcome the assumptions of linear events, sparsity, or low rank (Jia and Ma, 2017); however, the accuracy of the interpolated results is limited by the similarity of the characteristics between the training data and the processed data.

Prediction-based interpolation methods are important approaches for seismic data interpolation, and it involves both characteristics of seismic phase-shift operator and signal convolution operator. Prediction filters (PFs) or prediction-error filters (PEFs) can be implemented in the time-space or frequency-space domain. Spitz (1991) initially proposed f-x PFs for the interpolation of missing seismic data. Porsani (1999) improved Spitz's approach by introducing a half-step PF. Wang (2002) further extended prediction interpolation from the f-x domain to the f-x-y domain. Wang et al. (2007) also designed a localized f-x-y PF to interpolate 3D seismic data. Naghizadeh and Sacchi (2009) used exponentially weighted recursive least squares to calculate adaptive PFs in the f-x domain. Curry and Shan (2010) used multiples and frequency domain PEFs to interpolate missing data near offsets. Li et al. (2017) proposed a

multidimensional adaptive PEF to reconstruct seismic data in the frequency domain. Liu and Chen (2018) developed a prediction interpolation by using f-x regularized nonstationary autoregression (RNA), which can deal with the events that have space-varying dips. Zheng et al. (2019) developed a SPF in the f-x domain to interpolate missing traces, which reduces high computational cost by directly solving an inverse problem in the complex domain. Meanwhile, time-space PEFs were successfully applied to reconstruct datasets where the missing data might be regularly or irregularly represented. Claerbout (1992) first proposed missing-data restoration using PEFs in the t-x domain. Crawley et al. (1999) described a method for data interpolation with smoothly varying PEFs, which used "steering filters" to control the smoothness of the filters. Curry (2003) developed a nonstationary, multi-scale PEFs to interpolate irregularly-sampled data. Liu and Fomel (2011) restored decimated and randomly missing traces based on RNA in the time domain, which uses shaping regularization to control the smoothness of adaptive PEFs. Liu et al. (2018) proposed a 3D t-x-y multiscale multidirectional adaptive PEF to simultaneously reconstruct randomly and regularly missing data. Compared with f-x PF, a t-x PF could avoid the generation of false events in the presence of strong parallel events (Abma and Claerbout, 1995). This is because of the ability of t-x prediction, to control the length of the PFs in time. To reduce computational time and storage, Fomel and Claerbout (2016) proposed noniterative streaming PEFs to recover holes in 2D images.

In this paper, we proposed an SPF with varying smoothness in the time-space domain to reconstruct irregular and regular missing seismic traces; in this method, SPFs are extended from one to two spatial dimensions. The proposed method involves a two-step strategy (Claerbout, 1992; Crawley et al., 1999). In comparison with streaming PEFs (Fomel and Claerbout, 2016), we presented a similarity matrix to restrict the underdetermined least-squares problem of SPF, which enables regularization term to change with seismic data. We also designed a non-causal in space SPF to further improve the accuracy of interpolation and we compare its results with those from a causal in space filter. The proposed method shows the superiority of synchronous data reconstruction with irregular and regular missing seismic traces. Synthetic and field data tests demonstrate the effectiveness and efficiency of the proposed SPF method in reconstructing missing seismic data.

THEORY

Data interpolation can be cast as an inverse problem where the interpolated data can have minimum energy after specified filtering (Claerbout, 1992). A PEF can capture the inverse spectra of the data, thus a variety of PEFs have been used to find the missing data. Unlike the relationship between frequency-space PF and PEF, time-space PF has different coefficients from the corresponding PEF that involves causal time prediction coefficients along the column of the predicted data. The PEF creates the residual and the PF result is the data itself. Time-space PF only preserves spatial predictability in seismic data, therefore, it may provide more reasonable interpola-

tion results than time-space PEF, especially in field non-white noise environments. Data interpolation is commonly implemented as a two-step approach, which includes unknown PF estimation from the known data and missing data reconstruction from the calculated PF. Most adaptive PFs and PEFs based on iterative or recursive approaches are capable of handling the nonstationarity of seismic data, but iterations lead to high computation time and large storage requirements for variable coefficients. In this study, we propose a non-iterative, fast, adaptive PF that acts in the time-space domain.

Step 1: The t-x-y SPF estimation

Linear events with different constant dips can be predicted by a PF or an autoregression operator in the time-space domain, which is calculated to minimize the energy of the prediction error. Consider a 3D t-x-y PF $a_{i,j,k}$ to predict a given centered sample $d(t,x,y)$ of data:

$$\widehat{d}(t,x,y) = \sum_{i=-L}^{L} \sum_{\substack{j=-M \\ j\neq 0}}^{M} \sum_{\substack{k=-N \\ k\neq 0}}^{N} a_{i,j,k}(t,x,y) d_{i,j,k}(t,x,y), \tag{1}$$

where $d_{i,j,k}(t,x,y)$ represents the translation of $d(t,x,y)$ with time shifts i and space shifts j and k, nonstationary filter coefficients $a_{i,j,k}(t,x,y)$ change with time and space axes, and L, M, and N control the lengths of the filter along t, x, and y-axes, respectively.

In linear algebra notation, the filter coefficients $a_{i,j,k}$ are determined by minimizing the underdetermined least-squares problem:

$$\widehat{\mathbf{a}}(t,x,y) = \arg\min_{\mathbf{a}(t,x,y)} \parallel d(t,x,y) - \mathbf{d}(t,x,y)^T \mathbf{a}(t,x,y) \parallel_2^2, \tag{2}$$

where $\mathbf{a}(t,x,y)$ represents the vector of filter coefficients and $\mathbf{d}(t,x,y)$ represents the vector of data translations $d_{i,j,k}(t,x,y)$. For nonstationary situations, we can use different regularization term to constrain equation 2, such as global smoothness (Liu and Fomel, 2011). Sacchi and Naghizadeh (2009) introduced a local smoothness constraint to calculate the adaptive prediction filter. Fomel and Claerbout (2016) proposed the same constraint and solved the algebraic problem analytically with streaming computation, which demonstrated the same results as Sacchi and Naghizadeh's method. The local constraint is that the new filter $\mathbf{a}$ stays close to the prior neighboring filter $\bar{\mathbf{a}}$, $\xi\mathbf{a} \approx \xi\bar{\mathbf{a}}$, where ξ is a scale parameter. However, the regularization term occasionally fails in the presence of strong amplitude variation. Thus, we improved the constraint with varying smoothness. The SPF in the t-x-y domain was found by solving the least-squares problem:

$$\widehat{\mathbf{a}}(t,x,y) = \arg\min_{\mathbf{a}(t,x,y)} \parallel d(t,x,y) - \mathbf{d}(t,x,y)^T \mathbf{a}(t,x,y) \parallel_2^2 + \sum_{n=t,x,y} \xi_n^2 \parallel \mathbf{a}(t,x,y) - \mathbf{E}_n \bar{\mathbf{a}}_{\mathbf{n}}(t,x,y) \parallel_2^2, \tag{3}$$

where $\mathbf{E}_n$ is the similarity matrix, which controls the closeness between the adjacent filters. For the design of $\mathbf{E}_n$, we can use the data value and follow three principles:
1. Usage of PF to characterize the energy spectra of data; hence, both the adjacent data and the adjacent PFs are similar based on local plane wave assumption. Therefore, $\mathbf{E}_n$ should be close to identity matrix.
2. Data value is not be used alone in the expression of $\mathbf{E}_n$; otherwise, the calculation will be unstable because there exists large number of data with zero value in the missing seismic data.
3. The variation of data value can reasonably control the local smoothness of filter coefficients.

In this study, we designed the $\mathbf{E}_n$ based on the amplitude difference of the smoothed data:

$$\mathbf{E}_n = \begin{bmatrix} 1+\delta_n * (\tilde{d}_{n-1} - \tilde{d}_n) & 0 & \cdots & 0 \\ 0 & 1+\delta_n * (\tilde{d}_{n-2} - \tilde{d}_{n-1}) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1+\delta_n * (\tilde{d}_{n-i} - \tilde{d}_{n-i+1}) \end{bmatrix} \tag{4}$$

where δ_n is the sale factor and $\tilde{d}$ represent the smooth version of data that are less affected by random noise, e.g., the preprocessed data using Gaussian filter.

In a 3D case, the regularization term in equation 3 should include three directions:

$$\begin{cases} \xi_t \mathbf{a}(t,x,y) \approx \xi_t \mathbf{E}_t \mathbf{a}(t-1,x,y) \\ \xi_x \mathbf{a}(t,x,y) \approx \xi_x \mathbf{E}_x \mathbf{a}(t,x-1,y) \\ \xi_y \mathbf{a}(t,x,y) \approx \xi_y \mathbf{E}_y \mathbf{a}(t,x,y-1) \end{cases} \tag{5}$$

The least-squares solution of equation 3 is:

$$\mathbf{a}(t,x,y) = [\mathbf{d}(t,x,y)\mathbf{d}(t,x,y)^T + \xi^2 \mathbf{I}]^{-1} [d(t,x,y)\mathbf{d}(t,x,y) + \xi^2 \tilde{\mathbf{a}}(t,x,y)], \tag{6}$$

where

$$\begin{aligned} \tilde{\mathbf{a}}(t,x,y) &= \frac{\xi_t^2 \mathbf{E}_t \mathbf{a}(t-1,x,y) + \xi_x^2 \mathbf{E}_x \mathbf{a}(t,x-1,y) + \xi_y^2 \mathbf{E}_y \mathbf{a}(t,x,y-1)}{\xi^2}, \\ \xi^2 &= \xi_t^2 + \xi_x^2 + \xi_y^2, \end{aligned} \tag{7}$$

and $\mathbf{I}$ is the identity matrix. The regularization terms ξ_n should have the same order of magnitude as the data. From equation 7, we can consider $\xi_n^2\mathbf{E}_n$ as a whole term, which provides an adaptive smoothness for the nonstationary PF.

In equation 5, a stable update of SPF requires that the adjacent filter coefficients have the same order of magnitude, and the stable condition is based on the selection of the parameters δ_n and ξ_n. We can calculate the difference between the maximum and minimum values in the data, and δ_n is selected as the reciprocal of this difference to guarantee that $\mathbf{E}_n$ may be close to the identity matrix. Meanwhile, the parameter ξ_n should be chosen to the constant value between the minimum and maximum values of the data according to the smoothness level of the regularization.

The inverse matrix in equation 6 can be directly calculated without iterative conjugate-gradient method. Sherman-Morrison formula (Hager, 1989) provided an analytic solution for the inverse of a special matrix like $(\mathbf{A}-\mathbf{BC})^{-1}$, where matrix $\mathbf{B}$ is a column vector and matrix $\mathbf{C}$ is a row vector. If $\mathbf{A}$ and $\mathbf{I}-\mathbf{CA}^{-1}\mathbf{B}$ are invertible, the inverse matrix results in:

$$(\mathbf{A}-\mathbf{BC})^{-1}=\mathbf{A}^{-1}+\mathbf{A}^{-1}\mathbf{B}(\mathbf{I}-\mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1}. \tag{8}$$

In this paper, $\mathbf{A}=\xi^2\mathbf{I}$, $\mathbf{B}=-\mathbf{d}(t,x,y)$, and $\mathbf{C}=\mathbf{d}(t,x,y)^T$ in equation 8. After algebraic simplification, the filter coefficients arrive at the explicit solution as given below:

$$\mathbf{a}(t,x,y)=\bar{\mathbf{a}}(t,x,y)+\frac{d(t,x,y)-\mathbf{d}(t,x,y)^T\bar{\mathbf{a}}(t,x,y)}{\xi^2+\mathbf{d}(t,x,y)^T\mathbf{d}(t,x,y)}\mathbf{d}(t,x,y), \tag{9}$$

Step 2: Data interpolation with t-x-y SPF

The SPF error $r(t,x,y)$ can be expressed as follows:

$$r(t,x,y)=d(t,x,y)-\mathbf{d}(t,x,y)^T\mathbf{a}(t,x,y)=\xi^2\frac{d(t,x,y)-\mathbf{d}(t,x,y)^T\bar{\mathbf{a}}(t,x,y)}{\xi^2+\mathbf{d}(t,x,y)^T\mathbf{d}(t,x,y)}, \tag{10}$$

equation 10 shares the same form as the second term in the right hand side of equation 9. Substituting equation 10 into equation 9, we obtain equation 11:

$$\mathbf{a}(t,x,y)=\bar{\mathbf{a}}(t,x,y)+\frac{r(t,x,y)}{\xi^2}\mathbf{d}(t,x,y). \tag{11}$$

When a missing data is encountered, $r(t,x,y)$ can be assigned as zero, and equation 9 can be reduced to:

$$\mathbf{a}(t,x,y)=\bar{\mathbf{a}}(t,x,y). \tag{12}$$

Therefore, the data interpolation is also implemented in a streaming manner, where the missing data are reconstructed right after the unknown filter gets updated, and the interpolated data is shown as:

$$\widehat{d}(t,x,y)=\mathbf{d}(t,x,y)^T\mathbf{a}(t,x,y)=\mathbf{d}(t,x,y)^T\bar{\mathbf{a}}(t,x,y). \tag{13}$$

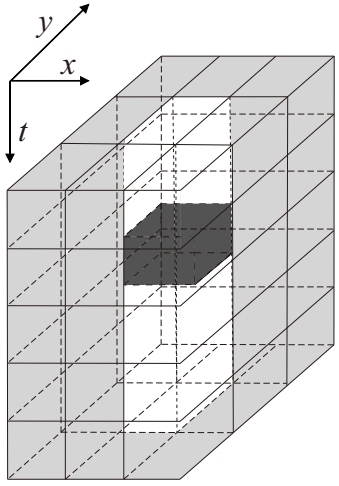
The field data always includes noise, hence the interpolated traces with noise is more realistic, where the prediction error $r(t, x, y)$ is set to a small random noise.

To use the available data for SPF estimation, we designed a 3D t-x-y non-causal in space SPF shown in Figure 1a, the light-gray grids represent prediction samples and the dark-gray ones exhibit target positions, whereas white grids represent unused samples. Non-causal in space SPF utilizes more adjoining traces around the target traces to predict signals, therefore, it can provide more accurate interpolated results than the causal in space version. The interpolation steps in the 2D case are illustrated schematically in Figure 1b. The black and white circles represent the known data and the missing data, respectively. Meanwhile, the dotted part is the prior non-causal SPF position, and dark-grey triangle is the target position. When the target position is known, the SPF coefficients $\mathbf{a}(t, x, y)$ can be obtained from equation 9. In streaming computation, we can use the time or space axis as the interpolation direction. The light-gray area in Figure 1b is the position where the SPF moves next, and the target trace becomes missing data. Further, spatial gaps are reconstructed according to equation 12 and 13. Note that the prior filter coefficients are required in this calculation. If the first target position is missing trace, e.g., marine data with near-offset missing, one may use the mirror data to initialize the coefficients of SPF in the space directions.

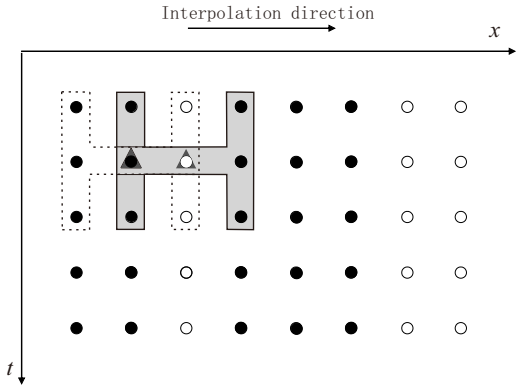
We also interpolated the results in the forward and backward spatial directions; adding the two results $\widehat{d}_{sum} = (\widehat{d}_{forw} + \widehat{d}_{back})/2$ can reduce the interpolated error caused by the directional properties of the streaming computation, where d_{forw} is the forward interpolated result and d_{back} is the backward one. The proposed method uses local varying smoothness of SPF to characterize time-space variation of nonstationary data, the analytical calculation of the inverse matrix in equation 6 avoids iteration, which results in superior computational speed. Table 1 compares the computational cost between 3D Fourier POCS (Abma and Kabir, 2006) and the 3D t-x-y SPF. The proposed method occupies less computational resources by reducing the cost to a single convolution.

Method	Cost	Filter storage
3D Fourier POCS	$O(N_t N_x N_y log(N_t N_x N_y) N_{iter})$	$O(N_f N_{k_x} N_{k_y})$
t-x-y SPF	$O(N_a N_t N_x N_y)$	$O(N_a N_x N_y)$

Table 1: Rough cost comparison between 3D Fourier POCS and t-x-y SPF. N_a is the filter size, N_{iter} is the number of iterations length, N_t is the data length in the time direction, N_x and N_y are the data length in the space directions, N_f is the data size along frequency axis, and N_{k_x} and N_{k_y} are the data size along wavenumber k_x and k_y axis, respectively.



(a)



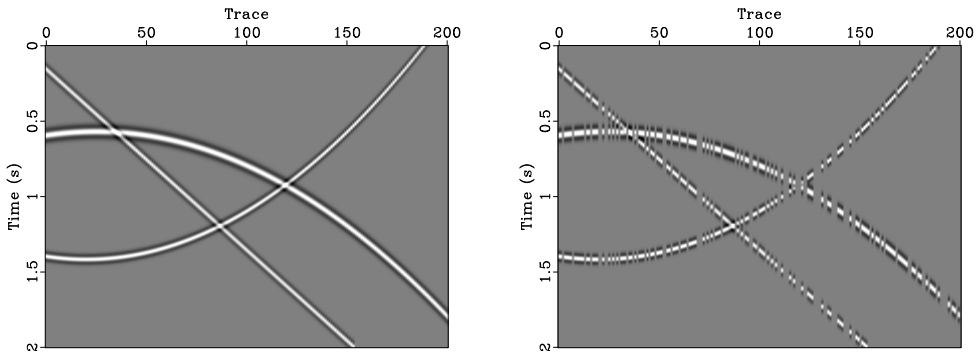
(b)

Figure 1: (a) Schematic illustration of a 3D non-causal in space SPF and (b) interpolation process.

SYNTHETIC DATA TESTS

2D nonstationary missing-trace interpolation test

We utilized a 2D synthetic model composed of nonstationary events with conflicting dips to test the effectiveness of the proposed method (Figure 2a). Two curved events and one dipping event show different amplitudes. We removed 40% of randomly selected traces specially at the intersecting position (Figure 2b). For comparison, we used a 2D Fourier POCS with 19×8 patches and a conventional streaming PEF (Fomel and Claerbout, 2016) to interpolate the missing traces (Figure 3a and Figure 3b). The size of each patch was set as 40×50 and the iteration number was selected to be 150 in the Fourier POCS method. The interpolated results using the SPF with causal and non-causal filters are shown in Figure 3c and Figure 3d, respectively. The difference between the reconstructed traces and the original traces is shown in Figure 4. For space non-causal filters, we designed a 2D t-x SPF with 25 (time) × 23 (space) coefficients and scalar parameters of $\xi_t = 0.05$, $\xi_x = 0.8$, $\delta_t = 0.5$, and $\delta_x = 0.5$ for each sample. The 2D Fourier POCS method with patching windows and SPF with causal filters produced similar results, except that more spatial aliasing was generated for the POCS method. The conventional streaming PEF has evident spatial aliasing at the cross position of the events and the lower right corner due to the helix transform. However, the proposed method reconstructed a reasonable result, where the interpolation errors were substantially reduced. The CPU times, for a single 2.10 GHZ CPU used in this study, were 2.68 s for the 2D Fourier POCS and 0.90 s for the proposed method.



(a) (b)

Figure 2: (a) Synthetic data with three events and (b) model with 40% randomly selected traces removed.

2D aliasing decimated-trace interpolation test

A benchmark example from Claerbout (2009) showed a strongly aliased gather. The number of space samples was set to 30. We used the two-step approach based on the

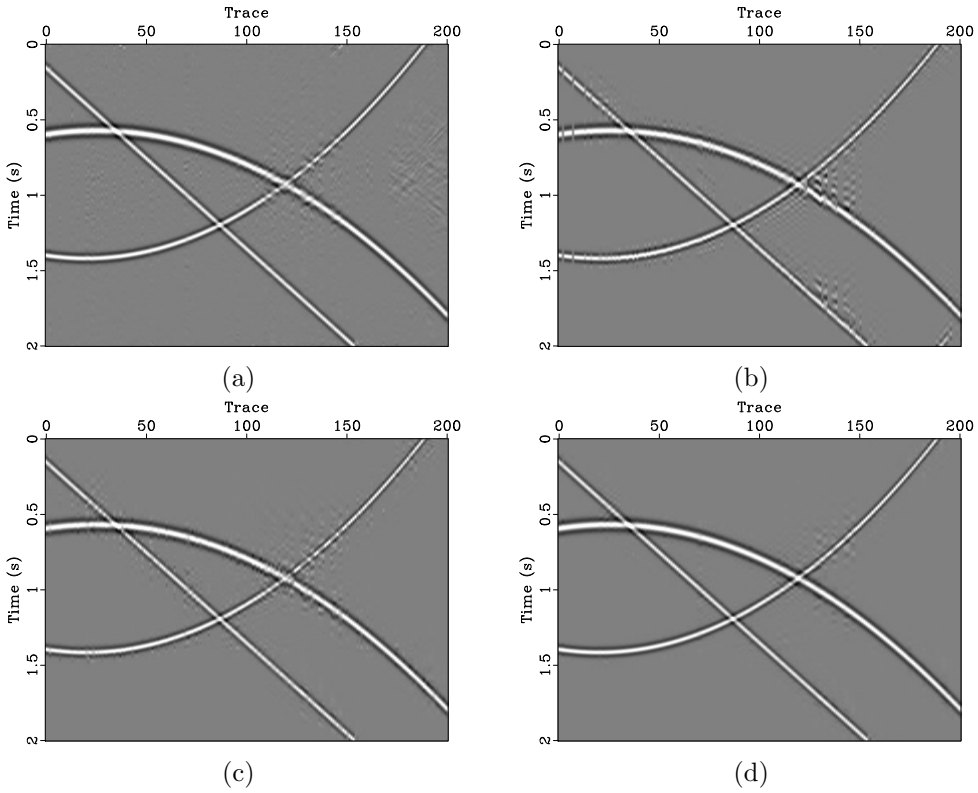


Figure 3: Reconstructed results by using different methods. (a) The 2D Fourier POCS, (b) the 2D streaming PEF, the 2D t-x SPF with (c) causal filter, and (d) non-causal filter.

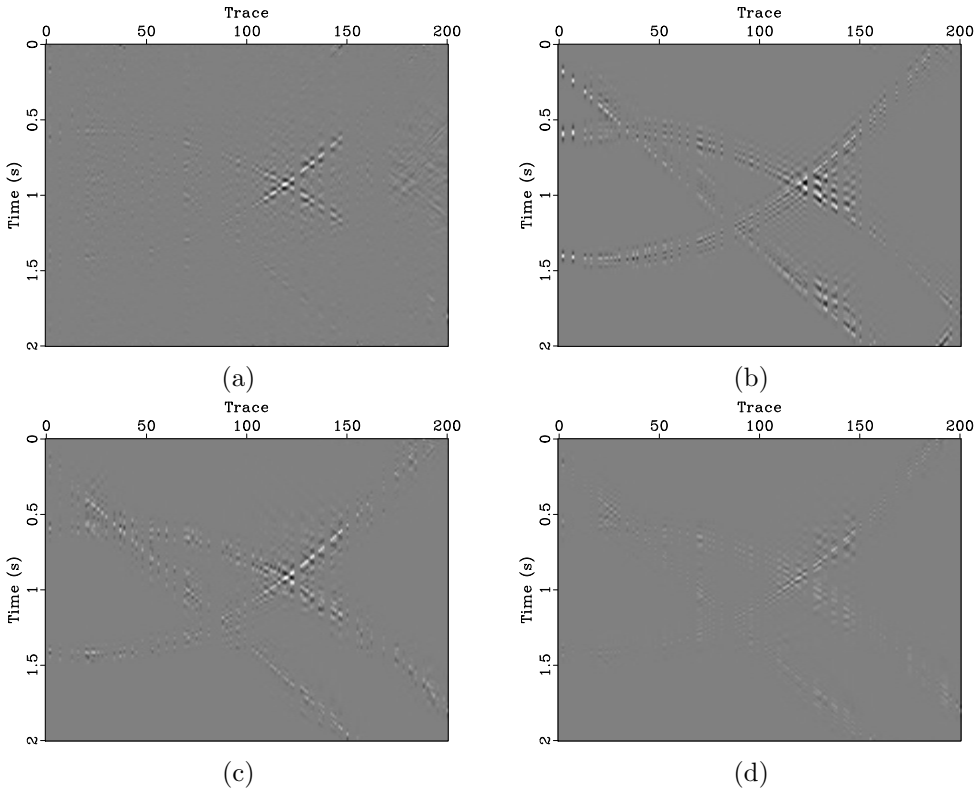
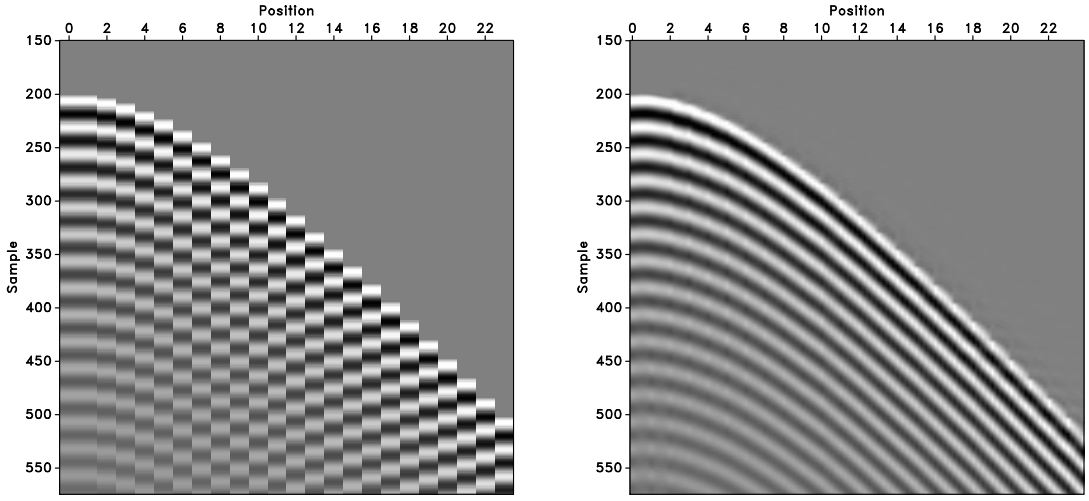


Figure 4: Interpolation errors by using different methods. (a) The 2D Fourier POCS, (b) the 2D streaming PEF, the 2D t-x SPF with (c) causal filter, and (d) non-causal filter.

t-x SPF to insert three additional traces between each of the adjointing input traces. We designed the SPF using 19 (time) $\times$ 11 (space) coefficients for each sample. The four scale parameters were 0.3 (ξ_t), 0.2 (ξ_x), 0.12 (δ_t), and 0.12 (δ_x). The proposed method effectively removed the spatial aliasing artifacts (Figure 5b). The SPF compared well with the plane-wave destruction (PWD) (Fomel, 2002) and adaptive PEF (Liu and Fomel, 2011), and showed higher efficiency in computational speed. The adaptive PEF methods were based on scale invariance for regular trace interpolation by interlacing the filter coefficients with zeros, however, the SPF methods cannot use the scale invariance because SPF is a local algorithm, which reconstructs decimated traces similar to missing traces. The CPU times, for single 2.10 GHZ CPU, were 1.18 s for the t-x SPF, 43.91 s for the PWD, and 10.71 s for the adaptive PEF.



(a) (b)

Figure 5: (a) Aliased synthetic model and (b) trace interpolation with the 2D t-x SPF. The interpolated data has four times more traces than the original model.

3D synthetic data test

We created a 3D prestack dataset (Figure 6a) from a 2D slice out of the benchmark French model (French, 1974), and the data was subsampled by a factor of two in both offset and shot axes, which caused visible aliasing of dipping events. Furthermore, we removed 15% of randomly selected traces from the decimated data. The data interleaved with zero traces along the offset and shot directions is shown in Figure 6b. The challenge of this test was to account for nonstationarity, aliasing, both decimated and irregular missing traces, and computational cost. Figure 7a and Figure 7b display the interpolated result using 3D Fourier POCS and the conventional 3D t-x-y SPF, respectively. Notably, the Fourier POCS method can only recover randomly missing traces, and it fails in handling regularly missing traces. For the proposed 3D t-x-y SPF, the choices of the filter length were seven samples in time axis, nine samples in the offset axis, and three samples in the shot axis. We designed the scale parameters,

$\xi_t = 0.4$, $\xi_x = 0.5$, $\xi_y = 0.4$, and $\delta_t = \delta_x = \delta_y = 0.01$, to deal with the variability of events. The conventional 3D t-x-y SPF did not recover the decimated data well. However, the proposed method succeeded in interpolating irregular and regular missing traces simultaneously (Figure 7c), which produced reasonable results for curved events. The CPU times of the 3D Fourier POCS with 500 iterations and the 3D t-x-y SPF were 889.21 s and 33.72 s, respectively.

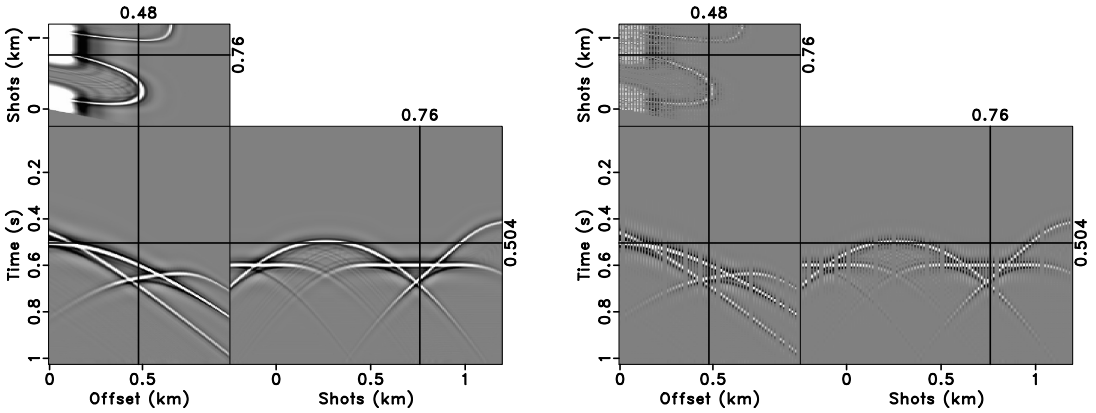


Figure 6: (a) 3D synthetic prestack data and (b) missing data interleaved with zero traces.

FIELD DATA EXAMPLE

To evaluate the performance of the t-x-y SPF interpolation method in 3D field conditions, we chose a set of marine shot gathers from a deep-water Gulf of Mexico survey (Fomel, 2002; Liu and Fomel, 2011). Figure 8a shows the complicated diffraction events caused by a salt body. We selected 35% traces of the input data by subsampling in the shot direction and removing 30% random traces (Figure 8b). For comparison, we used 3D Fourier POCS method and the conventional SPF to reconstruct the missing traces (Figure 9a and 9b, respectively). The Fourier POCS method also failed to interpolate the decimated traces and created some artificial events at the locations of the randomly-missing traces. The interpolated result could be partially improved by slicing data into patching windows. The conventional 3D t-x-y SPF also failed to recover the decimated data. Figure 9c shows that the proposed t-x-y SPF method produced better result, in which the missing gaps were recovered reasonable well, except for weaker amplitude in the common-offset sections. Figure 10 provides the f-k spectra corresponding to the original data and interpolated results with the Fourier POCS, the conventional 3D t-x-y SPF, and the proposed 3D t-x-y SPF, respectively. Figure 11 show the interpolation errors using these methods. The simultaneous occurrence of regular and irregular data missing is a challenge in the interpolation process. The proposed 3D t-x-y SPF method shows more reasonable results than the Fourier POCS and the conventional streaming PEF. Meanwhile, the

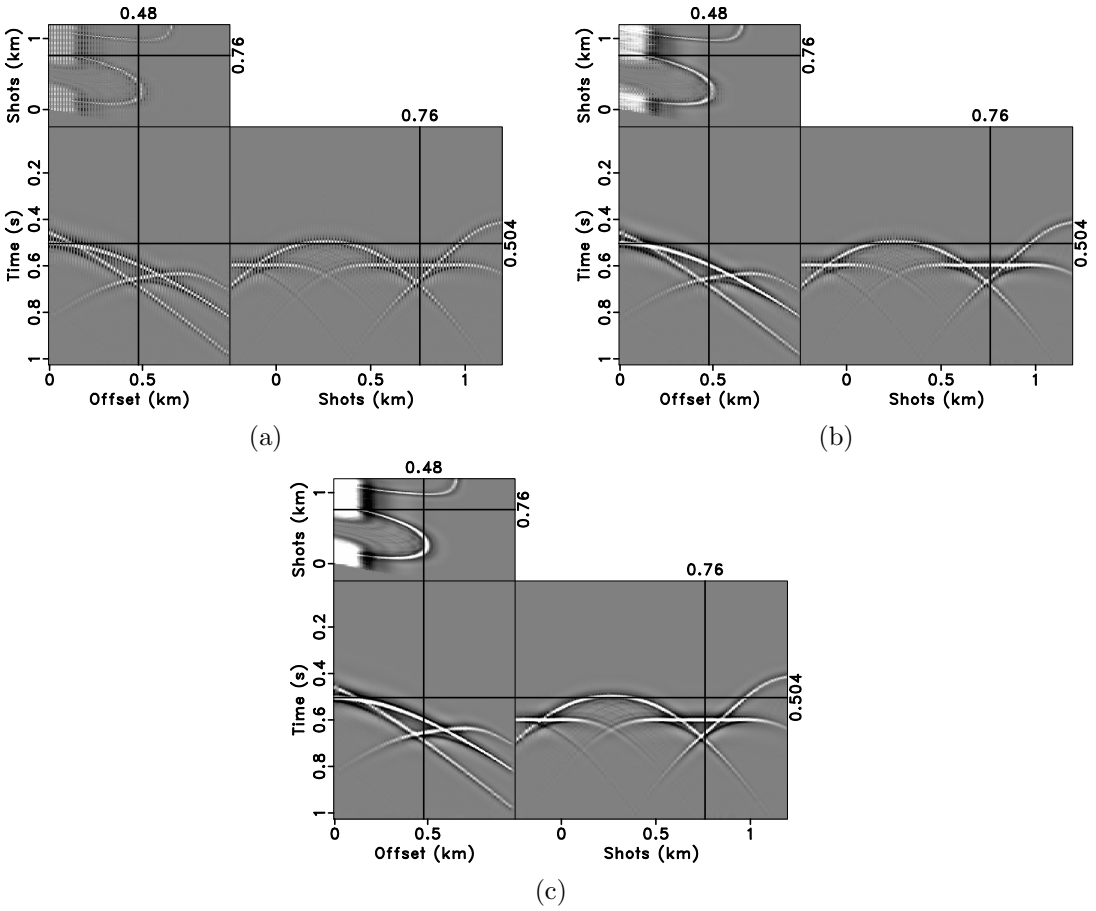


Figure 7: Reconstructed data volumes using different methods. (a) The 3D Fourier POCS, (b) the conventional 3D t-x-y SPF, and (c) the proposed 3D t-x-y SPF.

proposed algorithm is more efficient, and the CPU times for the 3D POCS with 500 iterations was 380.42 s whereas those of the 3D *t*-*x*-*y* SPF was 12.27 s.

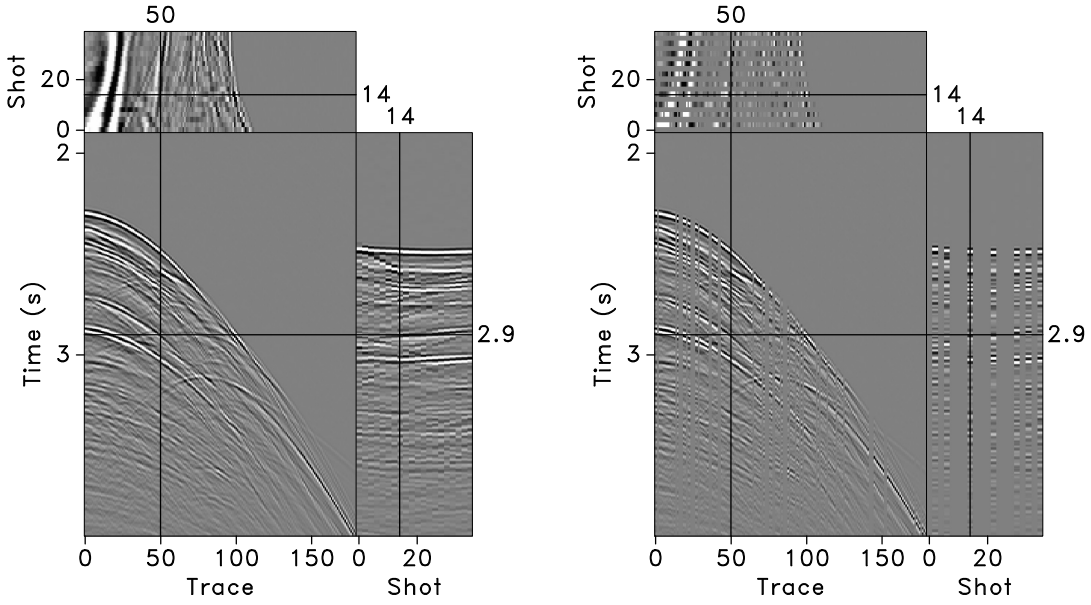


Figure 8: (a) A 3D field dataset and (b) data after subsampling in the shot direction and 30% randomly selected traces removed.

DISCUSSION

The extension of SPF to higher dimensions is straightforward, hence more space constrains have to be applied to the algorithm. For three spatial dimensions, we use spatial axes x, y and z, and the new filter coefficients by changing equation 3 show as given below:

$$\begin{aligned}\widehat{\mathbf{a}}(t,x,y,z) = \arg\min_{\mathbf{a}(t,x,y,z)} \parallel d(t,x,y,z) - \mathbf{d}(t,x,y,z)^T\mathbf{a}(t,x,y,z) \parallel_2^2 + \\ \sum_{n=t,x,y,z} \xi_n^2 \parallel \mathbf{a}(t,x,y,z) - \mathbf{E}_n\bar{\mathbf{a}}_{\mathbf{n}}(t,x,y,z) \parallel_2^2 .\end{aligned} \tag{14}$$

The least-squares solution of equation 14 is

$$\mathbf{a}(t,x,y,z) = [\mathbf{d}(t,x,y,z)\mathbf{d}(t,x,y,z)^T + \xi^2\mathbf{I}]^{-1}[d(t,x,y,z)\mathbf{d}(t,x,y,z) + \xi^2\bar{\mathbf{a}}(t,x,y,z)], \tag{15}$$

where

$$\begin{aligned}\bar{\mathbf{a}}(t,x,y,z) = \frac{\xi_t^2\mathbf{E}_t\mathbf{a}(t-1,x,y,z) + \xi_x^2\mathbf{E}_x\mathbf{a}(t,x-1,y,z) + \xi_y^2\mathbf{E}_y\mathbf{a}(t,x,y-1,z) + \xi_z^2\mathbf{E}_z\mathbf{a}(t,x,y,z-1)}{\xi^2}, \\ \xi^2 = \xi_t^2 + \xi_x^2 + \xi_y^2 + \xi_z^2.\end{aligned} \tag{16}$$

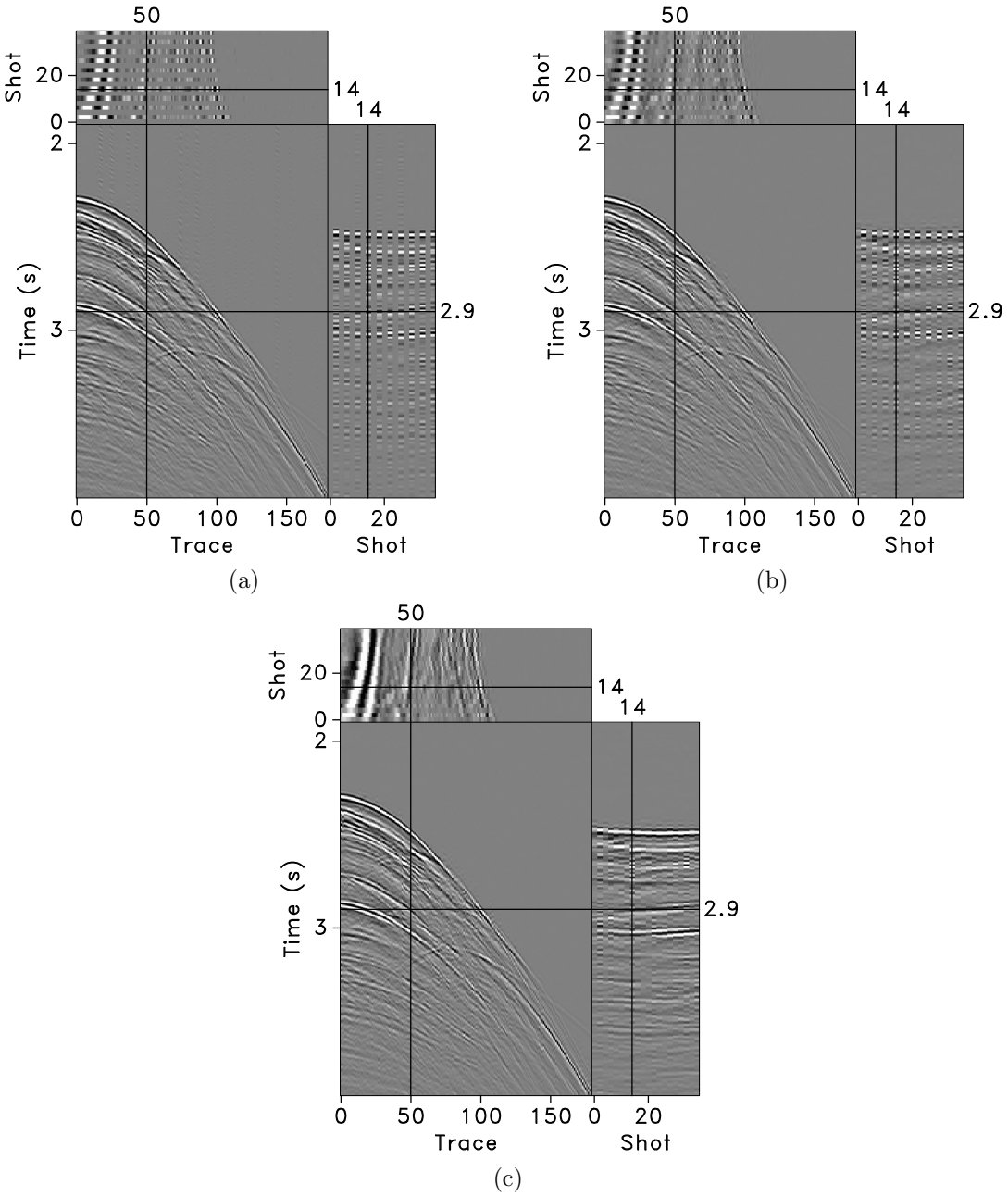


Figure 9: Interpolated results using different methods. (a) The 3D Fourier POCS, (b) the conventional 3D *t*-*x*-*y* SPF, and (c) the proposed 3D *t*-*x*-*y* SPF.

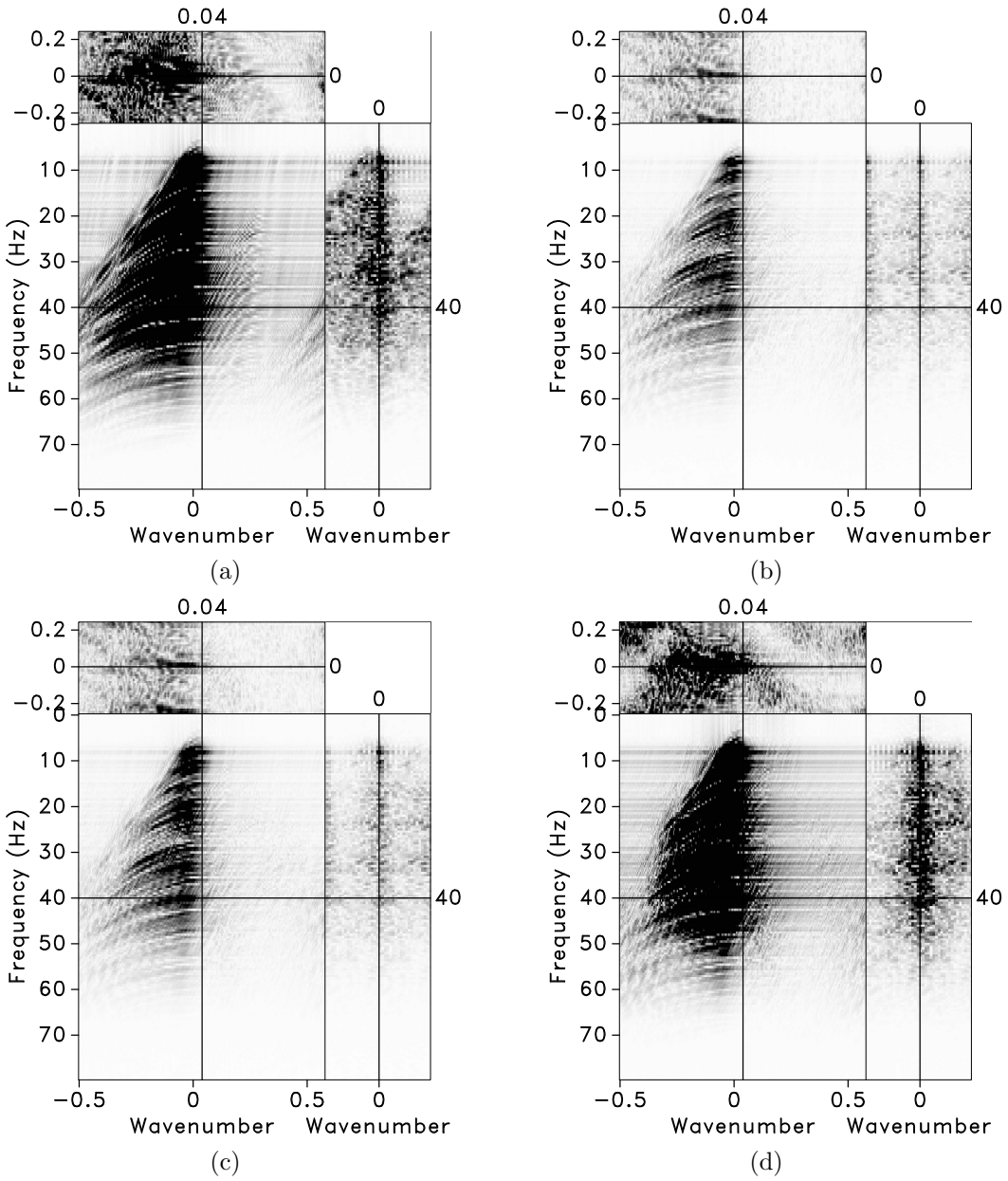


Figure 10: The f-k spectra for different data. (a) Original data (Figure 8a), (b) data in Figure 9a, (c) data in Figure 9b, and (d) data in Figure 9c.

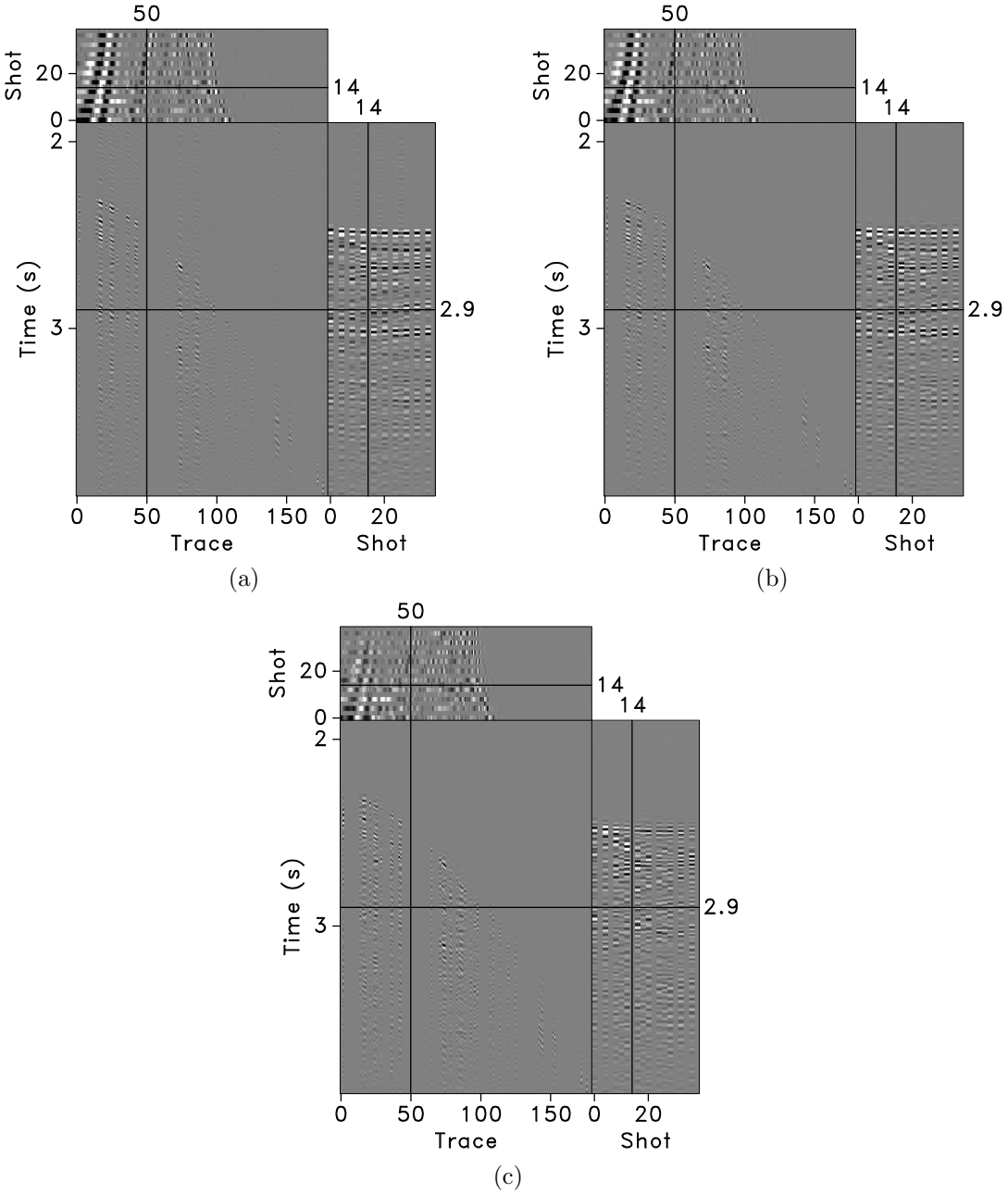


Figure 11: Interpolation errors using different methods. (a) The 3D Fourier POCS, (b) the conventional 3D *t*-*x*-*y* SPF, and (c) the proposed 3D *t*-*x*-*y* SPF.

Thus, the difference is the increased storage of the z-axis filter coefficients.

For low-amplitude events in field data, an AGC could be applied to the data before filter estimation to help ensure that low amplitude events are given equal attention in the SPF estimation. In practice, the prestack 3D traces could be described in 5D space (x_s, y_s, x_r, y_r, t), where (x_s, y_s) and (x_r, y_r) are the source and receiver coordinates. We can extract 3D seismic gathers from 5D space and use t-x-y SPF to interpolate the prestack 3D traces. Considering the spatial similarity of seismic events, it is recommended to implement the proposed method in the cmp-cmpline-offset-azimuth domain. Theoretically, any two spatial dimensions can be extracted from four spatial dimensions for interpolation; however, the accuracy of interpolated result depends on the number of missing traces and complexity in the 3D seismic gathers.

CONCLUSIONS

In this study, we proposed a fast approach based on SPF for simultaneously reconstructing irregular and regular missing traces in the t-x-y domain. With the help of local smoothness constraints, we defined a streaming computation manner to calculate nonstationary PF without multiple iterations. The invertible characteristics of SPF has direct applications, such as for seismic data interpolation. The non-causal in space filter structure and the similarity matrix guaranteed the accuracy of the interpolation results. Compared with the Fourier POCS method, the t-x-y SPF can characterize reasonably nonstationary signal while avoiding artifacts that occur in the frequency domain method. Moreover, the proposed method is superior in terms of its computational cost. Experiments with synthetic examples and field data demonstrated that the t-x-y SPF is effective at efficiently recovering irregular and regular missing traces in nonstationary seismic data.

ACKNOWLEDGMENTS

We thank Sergey Fomel and Zhicheng Geng for inspiring discussion about streaming computation. We thank the editor-in-chief, Dr. John Etgen, the anonymous associate editor, and four anonymous reviewers for the helpful suggestions that improved the quality of the paper. This work is supported by National Natural Science Foundation of China (grant nos. 41974134 and 41774127). All results are producible in the Madagascar open-source software environment (Fomel et al., 2013).

REFERENCES

Abma, R., and J. F. Claerbout, 1995, Lateral prediction for noise attenuation by t-x and f-x techniques: Geophysics, **60**, no.6, 1887–1896.

Abma, R., and N. Kabir, 2006, 3D interpolation of irregular data with a POCS algorithm: Geophysics, **71**, no.6, E91–E97.

Claerbout, J. F., 1992, Earth soundings analysis: Processing versus inversion: Blackwell Scientific Publications.

———, 2009, Basic earth imaging: Stanford exploration project.

Crawley, S., R. Clapp, and J. F. Claerbout, 1999, Interpolation with smoothly nonstationary prediction-error filters: 69th Annual International Meeting, SEG, Expanded Abstracts, 1154–1157.

Curry, W., 2003, Interpolation of irregularly-sampled data with non-stationary, multiscale prediction-error filters: 73rd Annual International Meeting, SEG, Expanded Abstracts, 1913–1916.

Curry, W., and G. Shan, 2010, Interpolation of near offsets using multiples and prediction-error filters: Geophysics, **75**, no.6, WB153–WB164.

Fomel, S., 2002, Applications of plane-wave destruction filters: Geophysics, **67**, no.6, 1946–1960.

———, 2003, Seismic reflection data interpolation with differential offset and shot continuation: Geophysics, **68**, no.2, 733–744.

Fomel, S., and J. F. Claerbout, 2016, Streaming prediction-error filters: 86th Annual International Meeting, SEG, Expanded Abstracts, 4787–4791.

Fomel, S., P. Sava, I. Vlad, Y. Liu, and V. Bashkardin, 2013, Madagascar: open-source software project for multidimensional data analysis and reproducible computational experiments: Journal of Open Research Software, **1**, e8.

French, W. S., 1974, Two-dimensianal and three-dimensional migration of model-experiment reflection profiles: Geophysics, **39**, no.3, 265–277.

Hager, W. W., 1989, Updating the inverse of a matrix: SIAM Review, **31**, no.2, 221–239.

Herrmann, F., and G. Hennenfent, 2008, Non-parametric seismic data recovery with curvelet frames: Geophysics Journal International, **173**, no.1, 233–248.

Jager, C., T. Hertweck, and P. Hubral, 2002, The unified approach and its applications:wave-equation based trace interpolation: 72nd Annual International Meeting, SEG, Expanded Abstracts, 2178–2181.

Jia, Y., and J. Ma, 2017, What can machine learning do for seismic data processing? an interpolation application: Geophysics, **82**, no.3, V163–V177.

Jia, Y., S. Yu, and J. Ma, 2018, Intelligent interpolation by monte carlo machine learning: Geophysics, **83**, no.2, V83–V97.

Li, C., G. Liu, Z. Hao, S. Zu, F. Mi, and X. Chen, 2017, Multidimensional seismic data reconstruction using frequency-domain adaptive prediction-error filter: IEEE Transactions on Geoscience and Remote Sensing, **56**, no.4, 2328–2336.

Liu, G., and X. Chen, 2018, Seismic data interpolation using frequency domain complex nonstationary autoregression: Geophysical Prospecting, **66**, no.3, 478–497.

Liu, G., C. Li, Z. Guo, and Y. Rao, 2018, Irregularly sampled seismic data reconstruction using multiscale multidirectional adaptive prediction-error filter: IEEE Transactions on Geoscience and Remote Sensing, **57**, no.5, 2909–2919.

Liu, Y., and S. Fomel, 2010, OC-seislet: seislet transform construction with differential offset continuation: Geophysics, **75**, no.6, WB235–WB245.

———, 2011, Seismic data interpolation beyond aliasing using regularized nonstationary autoregression: Geophysics, **76**, no.5, V69–V77.

Liu, Y., S. Fomel, and C. Liu, 2015, Signal and noise separation in prestack seismic data using velocity-dependent seislet transform: Geophysics, **80**, no.3, WD117–WD128.

Mazzucchelli, P., F. Rocca, and U. Spagnolini, 1999, Regularizing land acquisition using shot continuation operators:effects on amplitudes: 69th Annual International Meeting, SEG, Expanded Abstracts, 1995–1998.

Naghizadeh, M., and M. D. Sacchi, 2009, f-x adaptive seismic-trace interpolation: Geophysics, **74**, no.1, V9–V16.

———, 2010, Beyond alias hierarchical scale curvelet interpolation of regularly and irregularly sampled seismic data: Geophysics, **75**, no.6, WB183–WB202.

Porsani, J. M., 1999, Seismic trace interpolation using half-step prediction filters: Geophysics, **64**, no.5, 1461–1467.

Sacchi, M. D., and M. Naghizadeh, 2009, Adaptive linear prediction filtering for random noise attenuation: 79th Annual International Meeting, SEG, Expanded Abstracts, 3347–3350.

Shahidi, R., G. Tang, J. Ma, and F. Herrmann, 2013, Application of randomized sampling schemes to curvelet-based sparsity-promoting seismic data recovery: Geophysical Prospecting, **61**, no.5, 973–997.

Shao, J., Y. Wang, Q. Xue, and X. Chang, 2017, Radon domain interferometric interpolation of sparse seismic data: 87th Annual International Meeting, SEG, Expanded Abstracts, 4343–4347.

Spitz, S., 1991, Seismic trace interpolation in the F-X domain: Geophysics, **56**, no.6, 785–794.

Wang, B., R. Wu, X. Chen, and J. Li, 2015, Simultaneous seismic data interpolation and denoising with a new adaptive method based on dreamlet transform: Geophysics Journal International, **201**, no.2, 1180–1192.

Wang, B., N. Zhang, W. Lu, and J. Wang, 2019, Deep-learning-based seismic data interpolation: A preliminary result: Geophysics, **84**, no.1,V 11–V20.

Wang, X., Y. Zheng, and M. Perz, 2007, 3D seismic data interpolation via spatially localized f-xy prediction: 77th Annual International Meeting, SEG, Expanded Abstracts, 2021–2025.

Wang, Y., 2002, Seismic trace interpolation in the f-x-y domain: Geophysics, **67**, no.4, 1232–1239.

Wang, Y., Y. Luo, and G. T. Schuster, 2009, Interferometric interpolation of missing seismic data: Geophysics, **74**, no.3, SI37–SI45.

Wang, Y., B. Wang, N. Tu, and J. Geng, 2020, Seismic trace interpolation for irregularly spatial sampled data using convolutional auto-encoder: Geophysics, **85**, no.2, V119–V130.

Yang, P., and J. Gao, 2012, Curvelet-based POCS interpolation of nonuniformly sampled seismic records: Journal of Applied Geophysics, **79**, 90–99.

Zheng, Z., Y. Liu, C. Liu, and G. Wu, 2019, Seismic data interpolation using streaming prediction filter in the f-x domain: 81st Annual International Conference and Exhibition, EAGE, Expanded Abstracts, 1–5.