Streaming orthogonal prediction filter in t-x domain for random noise attenuation[a]

[a]Published in Geophysics, 83, F41-F48, (2018)

Yang Liu, Bingxiu Li†*

ABSTRACT

In seismic exploration there are many sources of random noise, for example, scattering from a complex surface. Prediction filters (PFs) have been widely used for random noise attenuation, but these typically assume that the seismic signal is stationary. Seismic signals are fundamentally nonstationary. Stationary PFs fail in the presence of nonstationary events, even if the data are cut into overlapping windows ("patching"). We propose an adaptive PF method based on streaming and orthogonalization for random noise attenuation in the t-x domain. Instead of using patching or regularization, the streaming orthogonal prediction filter (SOPF) takes full advantage of the streaming method, which generates the signal value as each new noisy data value arrives. The streaming signal-and-noise orthogonalization further improves the signal recovery ability of the SOPF. The streaming characteristic makes the proposed method faster than iterative approaches. In comparison with f-x deconvolution and f-x regularized nonstationary autoregression (RNA), we tested the feasibility of the proposed method in attenuating random noise on two synthetic datasets. Field data examples confirmed that the t-x SOPF had a reasonable denoising ability in practice.

INTRODUCTION

Random noise in seismic data comes from many sources, such as wind motion and poorly planted geophones. Prediction filters (PFs) have been applied in seismic data processing for decades, and have proved their effectiveness for random noise attenuation. The PF has different coefficients from prediction-error filters (PEFs), which include extra causal time-prediction coefficients. The different prediction filtering methods, varying from f-x deconvolution (Canales, 1984) to t-x prediction filters (Claerbout, 1992; Abma and Claerbout, 1995), play an important role in random noise attenuation. However, seismic signals are fundamentally nonstationary, and stationary PFs/PEFs still fail in the presence of nonstationary events even if filtering can be done either by "patching" or breaking data into overlapping windows. Different regularization methods (Crawley, 1999; Curry, 2003; Fomel, 2009; Liu and Chen, 2013; Liu et al., 2015) help PFs/PEFs estimate the nonstationary coefficients corresponding to the underdetermined autoregression problems.

Most of the nonstationary PFs/PEFs use iterative or recursive approaches to calculate their coefficients. This leads to high computational costs, especially in the storage of variable coefficients (Ruan et al., 2015). Recently, a streaming PEF (Fomel and Claerbout, 2016) was proposed to solve this problem. This method updates the PEF coefficients incrementally as new data arrive. This method reduces the computational cost of the streaming PEF to a single convolution. Moreover, the exact inversion of the streaming PEF makes missing data interpolation straightforward.

In this paper, we propose an adaptive PF method based on streaming and orthogonalization (Chen and Fomel, 2015) to attenuate random noise in nonstationary seismic data. The proposed method is able to characterize the nonstationarity on both time and space axes. The streaming element makes the proposed method a convenient and fast denoising approach. The application of orthogonalization further strengthens its ability in random noise attenuation. Numerical tests using synthetic and field data demonstrate the effectiveness of the proposed SOPF method.

THEORY

In the t-x domain, seismic data always display native nonstationarity. However, one can assume that the new prediction filter $\mathbf{a}$ stays close to the previous filter on time axis $\bar{\mathbf{a}}_t$ and the previous filter on space axis $\bar{\mathbf{a}}_x$. The autoregression equation of the t-x streaming PF can be written in a shortened block-matrix notation (Fomel and Claerbout, 2016) as

$$\begin{bmatrix} \mathbf{d}^T \\ \lambda_t \mathbf{I} \\ \lambda_x \mathbf{I} \end{bmatrix} \mathbf{a} \approx \begin{bmatrix} -d(t,x) \\ \lambda_t \bar{\mathbf{a}}_t \\ \lambda_x \bar{\mathbf{a}}_x \end{bmatrix}, \tag{1}$$

where $\mathbf{I}$ is the identity matrix and $d(t,x)$ is the given data sample at the point (t,x). λ_t and λ_x are the scale parameters controlling the filter variability on the two axes.

Consider a 2D noncausal prediction filter with 20 prediction coefficients $a_{i,j}$:

$$\begin{matrix} a_{-2,-2}(t,x) & a_{-2,-1}(t,x) & \cdot & a_{-2,1}(t,x) & a_{-2,2}(t,x) \\ a_{-1,-2}(t,x) & a_{-1,-1}(t,x) & \cdot & a_{-1,1}(t,x) & a_{-1,2}(t,x) \\ a_{0,-2}(t,x) & a_{0,-1}(t,x) & 1(t,x) & a_{0,1}(t,x) & a_{0,2}(t,x), \\ a_{1,-2}(t,x) & a_{1,-1}(t,x) & \cdot & a_{1,1}(t,x) & a_{1,2}(t,x) \\ a_{2,-2}(t,x) & a_{2,-1}(t,x) & \cdot & a_{2,1}(t,x) & a_{2,2}(t,x) \end{matrix} \tag{2}$$

where "$\cdot$" denotes zero, and the data vector $\mathbf{d}$ and the filter vector $\mathbf{a}$ in equation 1 are shown as follows:

$$\mathbf{d} = \begin{bmatrix} d_{-M,-N}(t,x) \\ d_{-M+1,-N}(t,x) \\ \cdots \\ d_{M-1,N}(t,x) \\ d_{M,N}(t,x) \end{bmatrix}, \mathbf{a} = \begin{bmatrix} a_{-M,-N}(t,x) \\ a_{-M+1,-N}(t,x) \\ \cdots \\ a_{M-1,N}(t,x) \\ a_{M,N}(t,x) \end{bmatrix}, \tag{3}$$

where $d_{i,j}(t,x)$ represents the translation of $d(t,x)$ in both time and space directions with time shift i and space shift j. $2M+1$ and $2N$ are the temporal and spatial lengths of the prediction filter, e.g., M=2 and N=2 in equation 2.

The least-squares solution of equation 1 is

$$\mathbf{a} = (\mathbf{d}\mathbf{d}^T + \lambda^2\mathbf{I})^{-1}(-d(t,x)\mathbf{d} + \lambda^2\bar{\mathbf{a}}), \tag{4}$$

where

$$\lambda^2 = \lambda_t^2 + \lambda_x^2, \quad \bar{\mathbf{a}} = \frac{\lambda_t^2\bar{\mathbf{a}}_t + \lambda_x^2\bar{\mathbf{a}}_x}{\lambda_t^2 + \lambda_x^2}. \tag{5}$$

Fomel and Claerbout (2016) propose the Sherman-Morrison formula to directly transform the inverse matrix in equation 4 without iterations. The Sherman-Morrison formula is an analytic method for solving the inverse of a special matrix (Hager, 1989). If both matrices $\mathbf{A}$ and $\mathbf{I} - \mathbf{V}\mathbf{A}^{-1}\mathbf{U}$ are invertible, then $\mathbf{A} - \mathbf{U}\mathbf{V}$ is invertible and

$$(\mathbf{A} - \mathbf{U}\mathbf{V})^{-1} = \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{U}(\mathbf{I} - \mathbf{V}\mathbf{A}^{-1}\mathbf{U})^{-1}\mathbf{V}\mathbf{A}^{-1}. \tag{6}$$

In the special case where matrix $\mathbf{U}$ is a column vector $\mathbf{u}$ and matrix $\mathbf{V}$ is a row vector $\mathbf{v}$, equation 6 can be rewritten as

$$(\mathbf{A} - \mathbf{u}\mathbf{v})^{-1} = \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{u}(1 - \mathbf{v}\mathbf{A}^{-1}\mathbf{u})^{-1}\mathbf{v}\mathbf{A}^{-1}. \tag{7}$$

The direct derivation of the Sherman-Morrison formula (equation 7) is described in Appendix A.

Applying the Sherman-Morrison formula to equation 4, the t-x streaming PF coefficients and prediction error can be calculated as

$$\mathbf{a} = \bar{\mathbf{a}} - \left(\frac{d(t,x) + \mathbf{d}^T\bar{\mathbf{a}}}{\lambda^2 + \mathbf{d}^T\mathbf{d}}\right)\mathbf{d}, \tag{8}$$

and

$$r(t,x) = \frac{\lambda^2(d(t,x) + \mathbf{d}^T\bar{\mathbf{a}})}{\lambda^2 + \mathbf{d}^T\mathbf{d}}. \tag{9}$$

For seismic random noise attenuation, we assume the residual of prediction filtering $r(t,x)$ is the random noise at the point (t,x). For calculating 2D t-x streaming PFs, we need to store one previous time-neighboring PF, $\bar{\mathbf{a}}_t$, and one previous space-neighboring PF, $\bar{\mathbf{a}}_x$, both $\bar{\mathbf{a}}_t$ and $\bar{\mathbf{a}}_x$ will be used when the stream arrives at its adjacency.

One can compare a streaming PF with a stationary PF. The autoregression equation for a traditional PF takes the following form:

$$d(t,x) + \sum_{j=-N,j\neq 0}^{j=N}\sum_{i=-M}^{i=M} a_{i,j}d_{i,j}(t,x) \approx 0. \tag{10}$$

The least-squares solution of equation 10 at each point (t, x) is

$$\mathbf{a} = -d(t,x)(\mathbf{d}\mathbf{d}^T)^{-1}\mathbf{d}. \tag{11}$$

The matrix in equation 11 is similar to that in equation 4. The comparison of equation 4 and equation 11 indicates that the results of the streaming PFs become gradually more accurate as the scale parameter λ decreases. However, according to equation 9, a small λ may cause the residual $r(t, x)$ to also be small, which means that there is too much noise in the signal section. To solve this problem, we use a two-step strategy. First, we choose a relatively large λ to get a large residual $r(t, x)$. The first step amounts to an "over-filtering", which generates an approximately "clean" signal. Next, the signal leakage in the noise section can be extracted by applying signal-and-noise orthogonalization.

We derive the definition of the streaming orthogonalization weight (SOW) from the global orthogonalization weight (GOW) (Chen and Fomel, 2015). Assume that the leaking signal $\mathbf{s}_1$ has a linear correlation with the estimated signal section $\mathbf{s}_0$ in the first step,

$$\mathbf{s}_1 = \omega\mathbf{s}_0, \tag{12}$$

where ω is the GOW. The ideal signal $\mathbf{s}$ and noise $\mathbf{n}$ can then be estimated by

$$\mathbf{s} = \mathbf{s}_0 + \omega\mathbf{s}_0, \tag{13}$$

$$\mathbf{n} = \mathbf{n}_0 - \omega\mathbf{s}_0. \tag{14}$$

Under the assumption that the noise $\mathbf{n}$ is orthogonal to the signal $\mathbf{s}$,

$$\mathbf{n}^T\mathbf{s} = 0. \tag{15}$$

Substituting equation 13 and 14 into equation 15, one can get the GOW as

$$\omega = \frac{\mathbf{n}_0^T\mathbf{s}_0}{\mathbf{s}_0^T\mathbf{s}_0}. \tag{16}$$

To get the orthogonalization weight for each data value, one possible definition of the SOW is:

$$\omega(t,x) = \frac{s(t,x)n(t,x)}{s(t,x)^2} = \frac{n(t,x)}{s(t,x)}, \tag{17}$$

where $\omega(t, x)$ is the SOW for the data point $d(t, x)$. $s(t, x)$ and $n(t, x)$ represent the signal and noise values at the point (t, x), respectively. In the SOPF, $s(t, x)$ and $n(t, x)$ correspond to the predictable part $\mathbf{d}^T\mathbf{a}$ and the prediction residual $r(t, x)$ produced in the first step. Direct point-by-point division of the values may produce unstable values. One common method to solve this problem is the iterative least-squares method (Chen and Fomel, 2015). Here, we propose a streaming method to calculate the SOW.

Suppose that the SOW gets updated with each new data point $d(t,x)$. The new SOW, ω, should stay close to the previous time-neighboring SOW $\bar{\omega}_t$ and the previous space-neighboring SOW $\bar{\omega}_x$. Equation 17 can be rewritten as

$$\begin{bmatrix} s(t,x) \\ \gamma_t \\ \gamma_x \end{bmatrix} \omega \approx \begin{bmatrix} n(t,x) \\ \gamma_t \bar{\omega}_t \\ \gamma_x \bar{\omega}_x \end{bmatrix}, \tag{18}$$

where γ_t and γ_x are the scale parameters controlling the variability on the two axes.

The least-squares solution of equation 18 is

$$\omega = \frac{s(t,x)n(t,x) + \gamma^2 \bar{\omega}}{s(t,x)^2 + \gamma^2}, \tag{19}$$

where

$$\gamma^2 = \gamma_t^2 + \gamma_x^2, \quad \bar{\omega} = \frac{\gamma_t^2 \bar{\omega}_t + \gamma_x^2 \bar{\omega}_x}{\gamma_t^2 + \gamma_x^2}. \tag{20}$$

Applying equation 13, one can get the denoised data value $\hat{s}(t,x)$ after the SOPF

$$\hat{s}(t,x) = s(t,x) + \frac{s(t,x)r(t,x) + \gamma^2 \bar{\omega}}{s(t,x)^2 + \gamma^2} s(t,x), \tag{21}$$

where $s(t,x)$ and $r(t,x)$ are the predictable signals and the prediction residuals calculated in the first step.

We implement the two-step strategy within the streaming method and obtain the denoised data value as each new noisy data value arrives. Table 1 compares the computational cost between f-x deconvolution, f-x regularized nonstationary autoregression (RNA) (Liu et al., 2012), and the proposed method. In general, the cost of SOPF is minimal.

Method	Filter storage	Cost
f-x deconvolution	$O(N_a)$	$O(N_a N_f N_x N_{iter})$
f-x RNA	$O(N_a N_f N_x)$	$O(N_a N_f N_x N_{iter})$
t-x SOPF	$O(N_a N_t)$	$O(N_a N_t N_x)$

Table 1: Rough cost comparison between f-x deconvolution, f-x RNA, and t-x SOPF. N_a is the filter size, N_f is the frequency size of the data in frequency domain, N_x is the spatial size of the data, N_t is the temporal size of the data, and N_{iter} is the number of iterations.

SYNTHETIC DATA TESTS

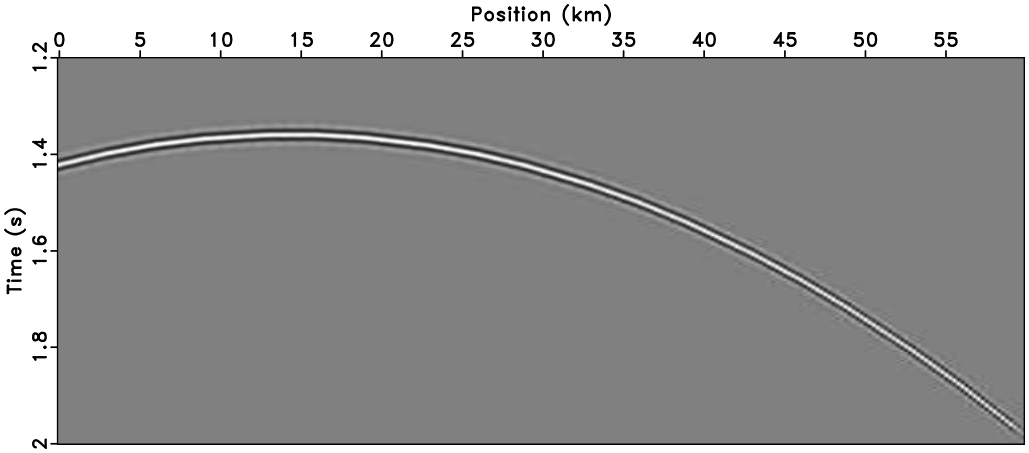
For comparison, we use f-x deconvolution, f-x RNA, and t-x SOPF to deal with random noise in two synthetic examples and one field dataset.

Curved model

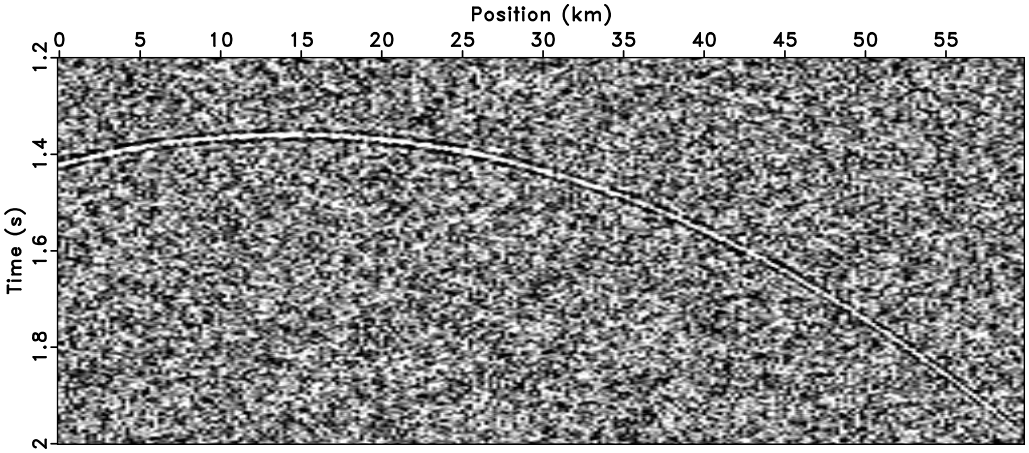
We use a synthetic example (Figure 1a) created by Raymond Abma (Liu and Fomel, 2011) to test the effectiveness of the proposed method in handling the nonstationarity. Figure 1b show data with uniformly distributed random noise. Figures 2a and 2b show the denoised result and the noise removed using f-x deconvolution, respectively. The data are divided into five patches with 40% overlap along the space axis. The f-x deconvolution creates artificial events and passes quite a lot of random noise. Nonstationary f-x RNA performs better. We set the filter length of the f-x RNA as 8 samples and the smoothing-radius size as 20 samples (in frequency) $\times$ 10-sample in space). Figure 3a shows that the f-x RNA passes less random noise than the f-x deconvolution. However, some artificial events still exist in the denoised result and there is signal energy leakage in the noise section (Figure 3b). Figure 4 is the result of processing by the t-x space-noncausal SOPF. The filter size is 7-sample (time) $\times$ 8-sample (space). The four scale parameters are 0.1 (λ_t), 0.08 (λ_x), 0.03 (γ_t), and 0.05 (γ_x), respectively. Figure 4a shows that the SOPF also introduces a few artifacts, but the artifacts follow a random distribution. Meanwhile, the difference (Figure 4b) between Figure 1b and Figure 4a indicates that the SOPF preserves signal better than the f-x RNA.

Shot gather

The second synthetic model was created by Guochang Liu (Liu et al., 2012). The shot gather (Figure 5a) has four hyperbolic events and 501 traces. Figure 5b is the noisy data containing random noise. The denoised result from the f-x deconvolution (Figure 6a) still contains lots of random noise. There is also signal leakage in the removed noise (Figure 6b). For f-x RNA, we set the filter length as 10 samples and the smoothing radius as 3 samples (in frequency) $\times$ 20 samples (in space). Figure 7b shows that f-x RNA has a good denoising performance in the shot gather test. However, it creates some artificial events with weak energy, which are parallel with the events (Figure 7a). As for the t-x space-noncausal SOPF, the filter size is 11 samples (in time) $\times$ 6 samples (in space) and the scale parameters are 2.0 (λ_t), 0.8 (λ_x), 0.4 (γ_t), and 1.0 (γ_x), respectively. The comparison of Figure 7a and 8a illustrates that the SOPF has similar signal preservation ability to f-x RNA. Careful examination indicates that the shaping regularization (Fomel, 2007) in the f-x RNA has a more powerful smoothing effect than the streaming method. Figures 7b and 8b show that the SOPF also eliminates equivalent random noise compared with the f-x RNA. However, the proposed method is still removing some signal along with the noise; we can preserve more signal by selecting smaller values for λ in equation 9. Users can therefore decide whether to choose more noise attenuation or more signal preservation to meet their data processing objectives. The computational cost of the SOPF is much less than that of f-x RNA because no iteration is used in the proposed method. For further illustrating how filtering methods change the amplitudes of the original signal, we calculate the differences between the clean data (Figure 5a) and the denoised



(a)



(b)

Figure 1: Curved model (a) and noisy data (b).

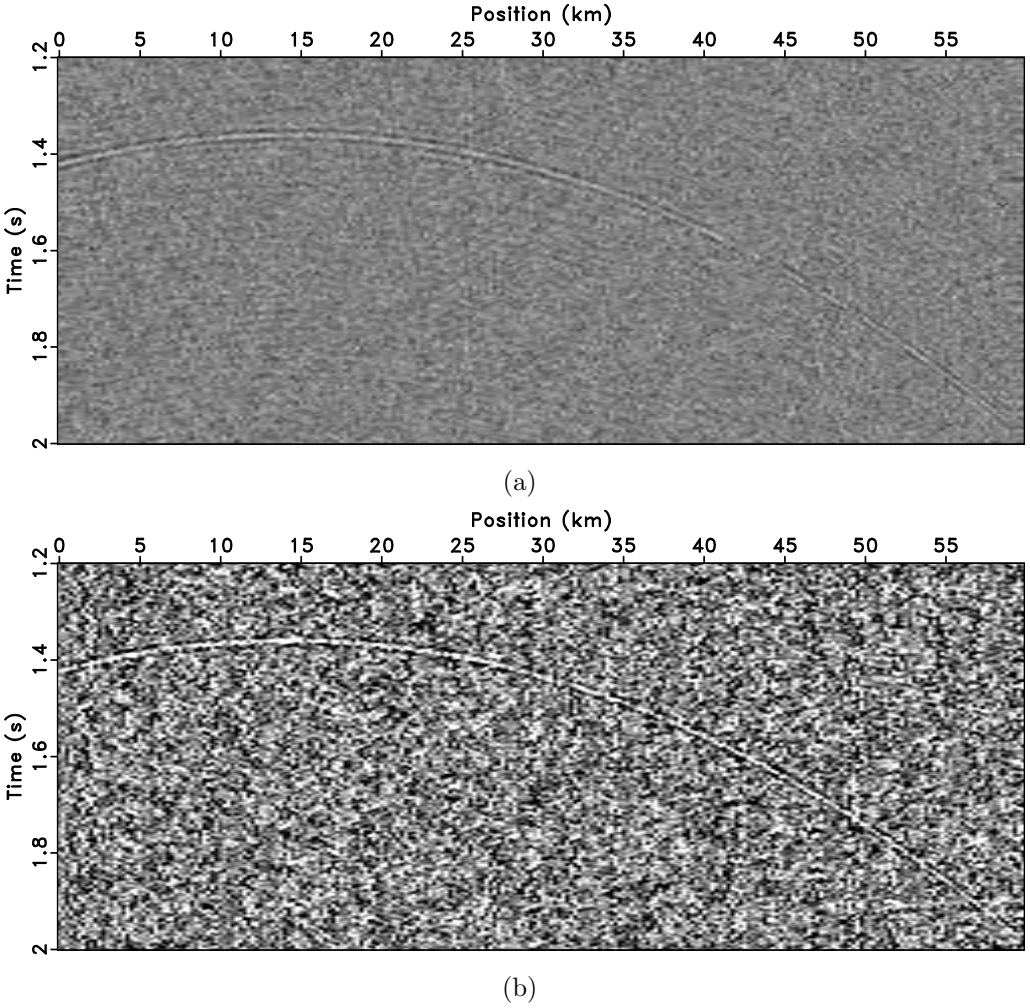
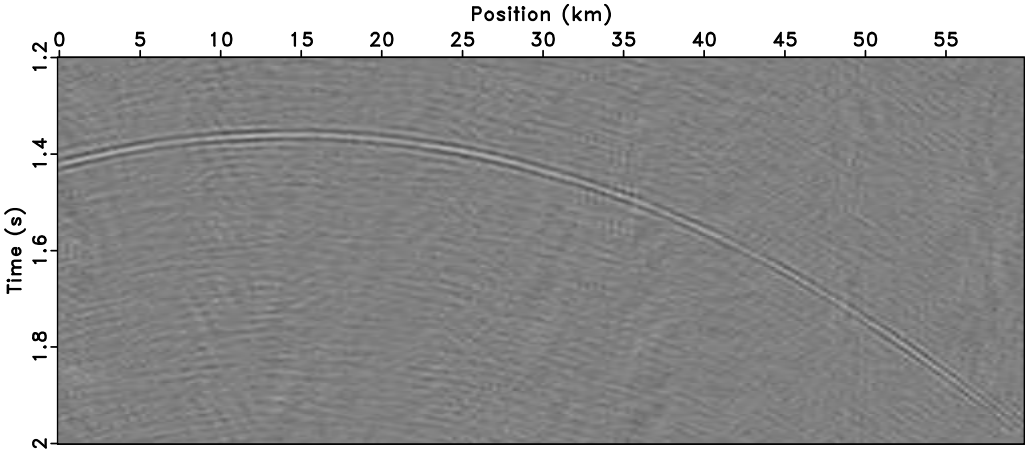
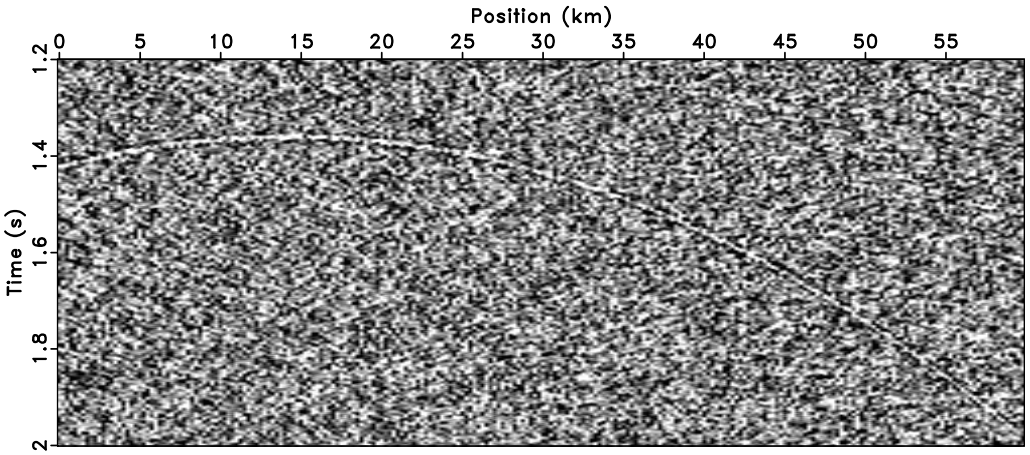


Figure 2: Denoised result (a) and noise removed (b) by the f-x deconvolution with patching.

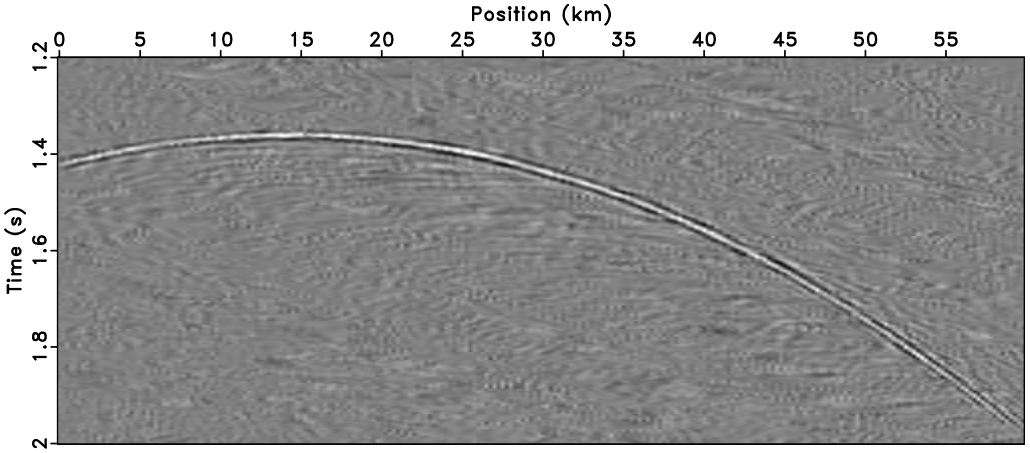


(a)

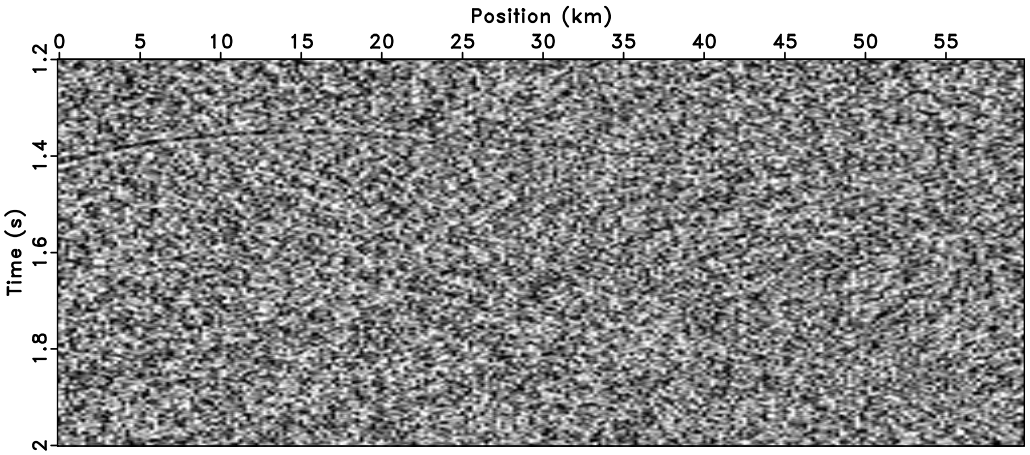


(b)

Figure 3: Denoised result by the nonstationary f-x RNA (a) and noise removed by the nonstationary f-x RNA (b).



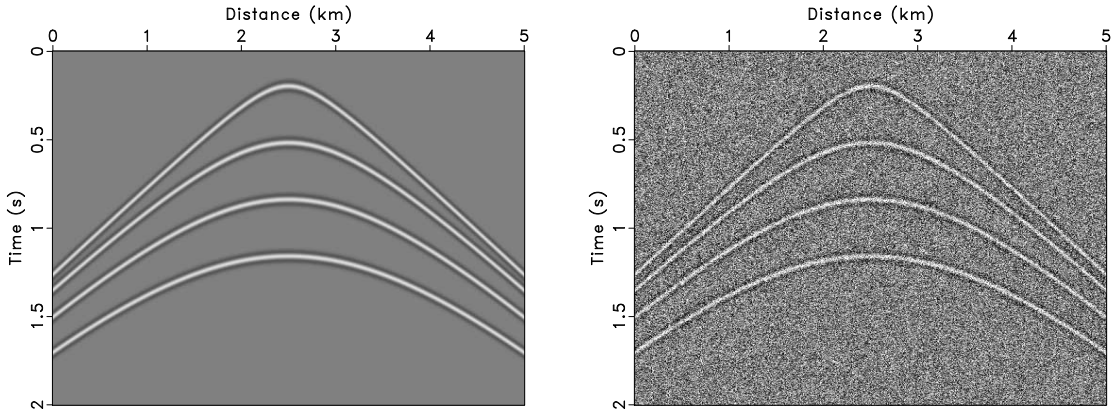
(a)



(b)

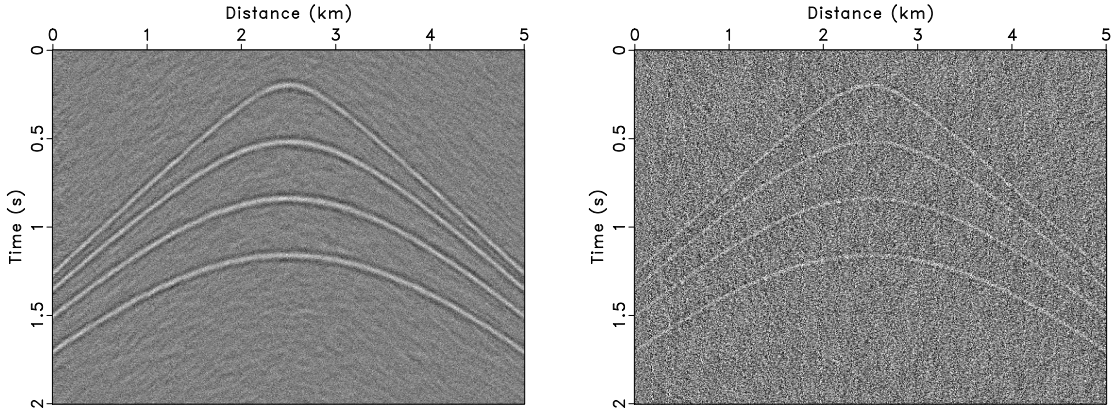
Figure 4: Denoised result by the *t-x* SOPF (a) and noise removed by the *t-x* SOPF (b).

results by the nonstationary f-x RNA (Figure 7a) or t-x SOPF (Figure 8a). The results are shown in Figure 9. The comparison shows the proposed method provides more reasonable signal preservation than the nonstationary f-x RNA.



(a) (b)

Figure 5: Shot gather (a) and noisy data (b).

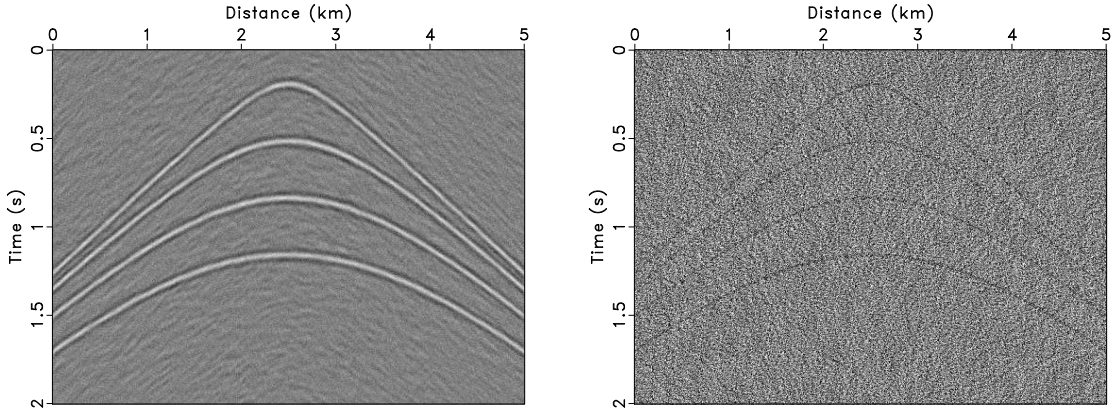


(a) (b)

Figure 6: Denoised result by the f-x deconvolution (a) and noise removed by the f-x deconvolution (b).

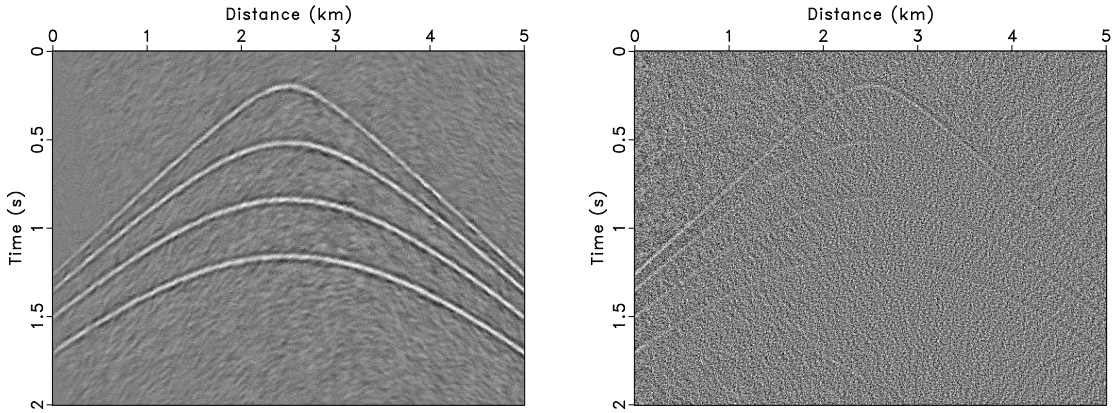
FIELD DATA TEST

For the field data test, we use a time-migrated dataset from Liu and Chen (2013). The input data shown in Figure 10 has 310 traces and 700 samples in the time direction with a sampling interval of 1 ms. Noise is mainly strong random noise caused by the surface conditions. Figure 11a and 11b are the denoised signal and the noise removed by the f-x deconvolution, respectively. These show that many signal events remain in the noise section. For further comparison, we apply f-x RNA to remove random noise (Figure 12a). The filter length of the f-x RNA is 12 samples and



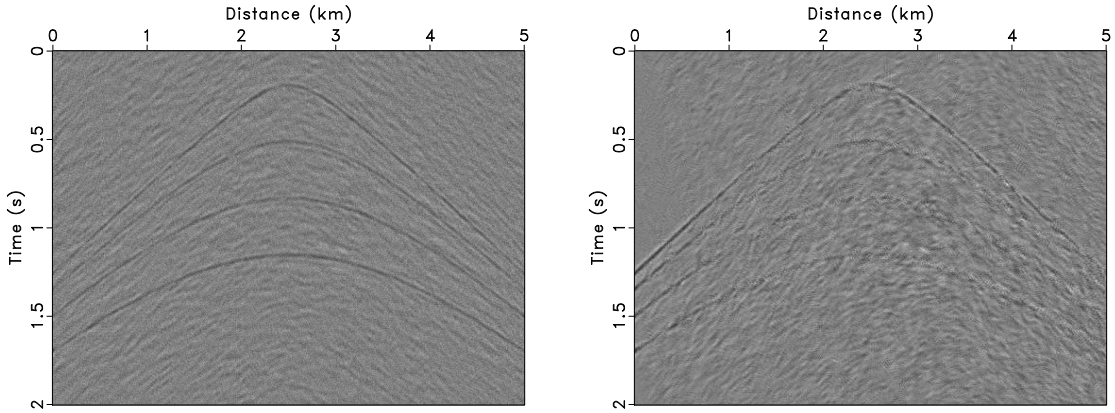
(a) (b)

Figure 7: Denoised result by the f-x RNA (a) and noise removed by the f-x RNA (b).



(a) (b)

Figure 8: Denoised result by the t-x SOPF (a) and noise removed by the t-x SOPF (b).



(a) (b)

Figure 9: Comparison of the difference between Figure 5a and the corresponding denoised result using two different methods, the nonstationary f-x RNA (Figure 7a) (a) and the t-x SOPF (Figure 8a) (b).

smoothing radius size is 3 samples (in frequency) $\times$ 20 samples (in space). Figure 12b indicates that while the f-x RNA has fewer signal events in the noise section than the f-x deconvolution but some weak events still remain. Figure 13a shows that the proposed t-x space noncausal SOPF method can produce reasonable results, in which the continuity of events and geological structure are enhanced. The filter size of the SOPF is 6 samples (in time) $\times$ 10 samples (in space) and the scale parameters are 100.0 (λ_t), 1.0 (λ_x), 1.0 (γ_t), and 10.0 (γ_x), respectively. Compared with Figure 12b, the difference (Figure 13b) between Figure 10 and Figure 13a contains no obvious events, and the computational speed of the t-x SOPF is faster than the iterative f-x RNA methods. If one chooses smaller values for λ in equation 9, the noise component that is parallel to the signal may be partly kept, however, it reduces the damage to the signal.

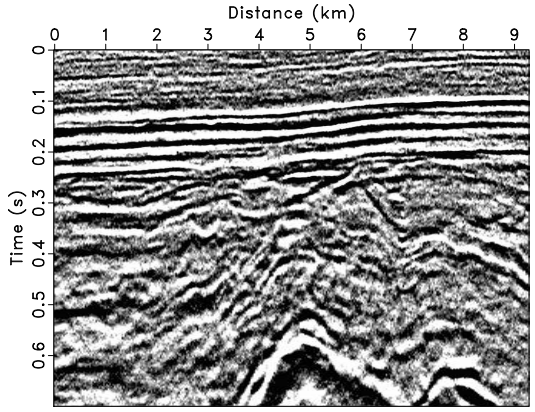
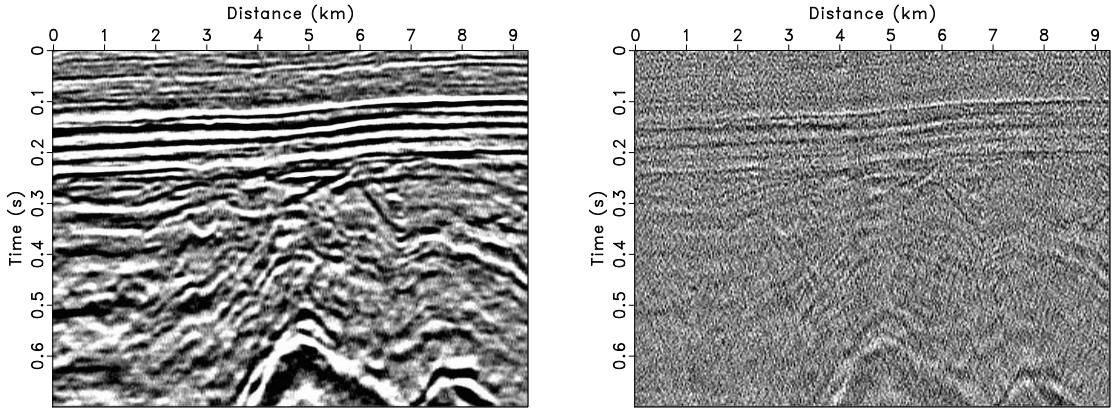
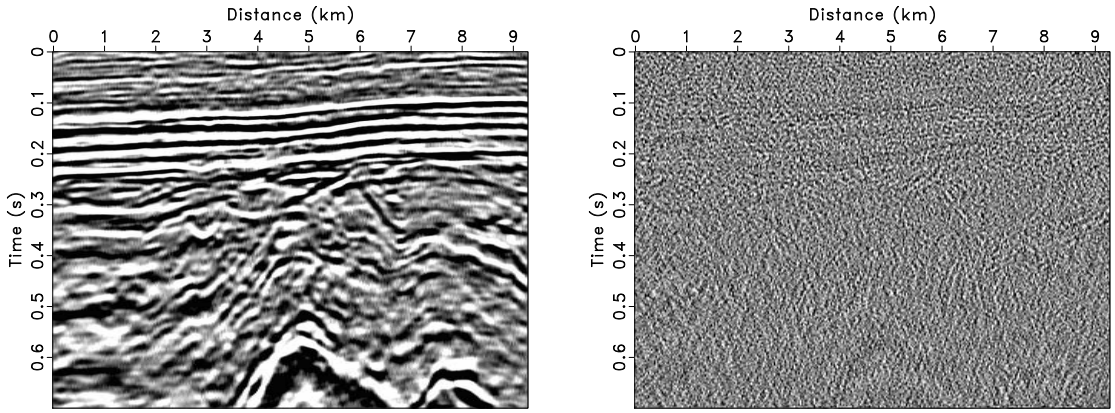


Figure 10: Field data.



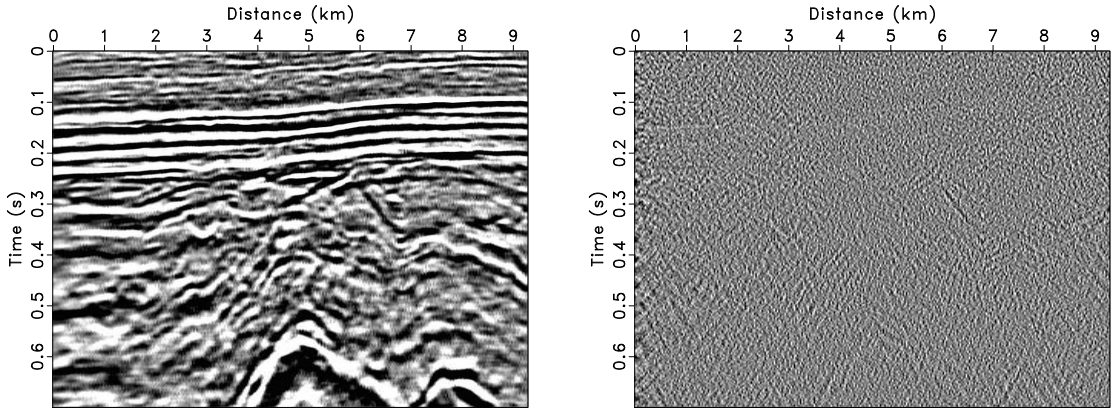
(a) (b)

Figure 11: Denoised result by the f-x deconvolution (a) and noise removed by the f-x deconvolution (b).



(a) (b)

Figure 12: Denoised result by the f-x RNA (a) and noise removed by the f-x RNA (b).



(a) (b)

Figure 13: Denoised result by the t-x SOPF (a) and noise removed by the t-x SOPF (b).

CONCLUSIONS

We introduce a fast approach to adaptive PF for random noise attenuation in the t-x domain. Our approach uses neighboring similarity of PF to handle time-space variations in nonstationary seismic data, which is based on elementary algebraic operations and a streaming method instead of an iterative strategy. Compared with f-x deconvolution and f-x RNA methods, t-x SOPF can capture a reasonably detailed signal while avoiding artifacts that occur in the frequency-domain method. Moreover, the t-x SOPF is superior in terms of its computational costs. Experiments with synthetic examples and field data demonstrate that the proposed method is effective at attenuating the random noise in nonstationary seismic data even in the presence of curved and conflicting events.

ACKNOWLEDGMENTS

The authors are grateful to Sergey Fomel for inspiring discussions. We thank the associate editor, Joseph Dellinger, and four anonymous reviewers for helpful suggestions, which improved the quality of the paper. This work is supported by National Natural Science Foundation of China (Grant No. 41774127 and 41522404) and National Key Research and Development Program of China (Grant No. 2016YFC0600505). All results are reproducible in the Madagascar open-source software environment (Fomel et al., 2013).

APPENDIX A: DIRECT DERIVATION OF THE SHERMAN-MORRISON FORMULA

The Sherman-Morrison formula can be derived directly by solving the linear problem as

$$(\mathbf{A}-\mathbf{uv})\mathbf{x}=\mathbf{b}, \tag{A-1}$$

where $\mathbf{u}$ is a column vector and $\mathbf{v}$ is a row vector. Assuming $\mathbf{A}^{-1}$ is already known, we pre-multiply equation A-1 by $\mathbf{A}^{-1}$ and denote $\mathbf{A}^{-1}\mathbf{u}=\mathbf{z}$ and $\mathbf{A}^{-1}\mathbf{b}=\mathbf{y}$ to give

$$\mathbf{x}-\mathbf{zvx}=\mathbf{y}. \tag{A-2}$$

Because $\mathbf{vx}$ is a scalar quantity, we can denote it as α. To find α, we pre-multiply the equation A-2 by $\mathbf{v}$ to give

$$\alpha-\mathbf{vz}\alpha=\mathbf{vy}. \tag{A-3}$$

Since $\mathbf{vz}$ and $\mathbf{vy}$ in the equation A-3 are scalars, one can easily solve for α to get

$$\alpha=\frac{\mathbf{vy}}{1-\mathbf{vz}}. \tag{A-4}$$

Thus the solution can be written as

$$\begin{aligned}\mathbf{x} &= \mathbf{y}+\alpha\mathbf{z} && \text{(A-5)}\\ &= \mathbf{A}^{-1}\mathbf{b}+\mathbf{A}^{-1}\mathbf{u}(1-\mathbf{vA}^{-1}\mathbf{u})^{-1}\mathbf{vA}^{-1}\mathbf{b} && \text{(A-6)}\\ &= [\mathbf{A}^{-1}+\mathbf{A}^{-1}\mathbf{u}(1-\mathbf{vA}^{-1}\mathbf{u})^{-1}\mathbf{vA}^{-1}]\mathbf{b}. && \text{(A-7)}\end{aligned}$$

We then obtain the Sherman-Morrison formula

$$(\mathbf{A}-\mathbf{uv})^{-1}=\mathbf{A}^{-1}+\mathbf{A}^{-1}\mathbf{u}(1-\mathbf{vA}^{-1}\mathbf{u})^{-1}\mathbf{vA}^{-1}. \tag{A-8}$$

REFERENCES

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

Canales, L. L., 1984, Random noise reduction: 54th Annual International Meeting, SEG, Expanded Abstracts, 525–527.

Chen, Y., and S. Fomel, 2015, Random noise attenuation using local signal-and-noise orthogonalization: Geophysics, **80**, WD1–WD9.

Claerbout, J., 1992, Processing versus inversion: Blackwell Scientific Publications.

Crawley, S., 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, multi-scale prediction-error filters: 73th Annual International Meeting, SEG, Expanded Abstracts, 1913–1916.

Fomel, S., 2007, Shaping regularization in geophysical-estimation problems: Geophysics, **72**, R29–R36.

———, 2009, Adaptive multiple subtraction using regularized nonstationary regression: Geophysics, **74**, V25–V33.

Fomel, S., and J. 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.

Hager, W. W., 1989, Updating the inverse of a matrix: SIAM Review, **31**, 221–239.

Liu, G. C., and X. H. Chen, 2013, Noncausal *f-x-y* regularized nonstationary prediction filtering for random noise attenuation on 3D seismic data: Journal of Applied Geophysics, **93**, 60–66.

Liu, G. C., X. H. Chen, J. Du, and K. L. Wu, 2012, Random noise attenuation using *f-x* regularized nonstationary autoregression: Geophysics, **77**, V61–V69.

Liu, Y., and S. Fomel, 2011, Seismic data interpolation beyond aliasing using regularized nonstationary autoregression: Geophysics, **76**, V69–V77.

Liu, Y., N. Liu, and C. Liu, 2015, Adaptive prediction filtering in *t-x-y* domain for random noise attenuation using regularized nonstationary autoregression: Geophysics, **80**, V13–V21.

Ruan, K., J. Jennings, E. Biondi, R. G. Clapp, S. A. Levin, and J. Claerbout, 2015, Industrial scale high-performance adaptive filtering with PEF applications: Stanford Exploration Project, **SEP-160**, 177–188.