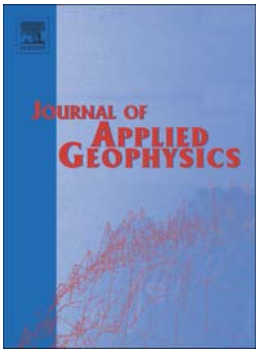

Adaptive noise estimation and suppression for improving microseismic event detection

S. Mostafa Mousavi, Charles A. Langston

PII: S0926-9851(16)30163-X
DOI: doi: 10.1016/j.jappgeo.2016.06.008
Reference: APPGEO 3009

To appear in: *Journal of Applied Geophysics*

Received date: 6 May 2016
Revised date: 6 June 2016
Accepted date: 26 June 2016

Please cite this article as: Mousavi, S. Mostafa, Langston, Charles A., Adaptive noise estimation and suppression for improving microseismic event detection, *Journal of Applied Geophysics* (2016), doi: 10.1016/j.jappgeo.2016.06.008

Adaptive Noise Estimation and Suppression for Improving Microseismic Event Detection

S. Mostafa Mousavi,

and

Charles A. Langston,

Running title: *Microseismic Noise Estimation and Denoising*

Center for Earthquake Research and Information (CERI)

University of Memphis,

3876 Central Ave., Suite 1

Memphis, TN 38152

smousavi@memphis.edu

clangstn@memphis.edu

To be submitted to

Journal of Applied Geophysics

June 2, 2016

ABSTRACT

Microseismic data recorded by surface arrays are often strongly contaminated by unwanted noise. This background noise makes the detection of small magnitude events difficult. A noise level estimation and noise reduction algorithm is presented for microseismic data analysis based upon minimally controlled recursive averaging and neighborhood shrinkage estimators. The method might not be compared with more sophisticated and computationally expensive denoising algorithm in terms of preserving detailed features of seismic signal. However, it is fast and data-driven and can be applied in real-time processing of continuous data for event detection purposes. Results from application of this algorithm to synthetic and real seismic data show that it holds a great promise for improving microseismic event detection.

1. Introduction

There has been a rapid growth in the application of microseismic monitoring using surface arrays primarily in monitoring of hydraulic fracturing and identification of induced seismic events (Maxwell, 2005). However, one of the major challenges in surface monitoring is distinguishing the noise from signal (Eisner et al, 2010). The signal-to-noise-ratio (SNR) of seismic data directly affects the reliability of the data and accuracy of source parameter estimation. Bandpass filtering is widely used in routine processing of the microseismic data to attenuate noise. Frequency filters are effective in attenuating noise outside of user-defined cut-off frequencies. However, filters are not usually effective for full attenuation of the noise without distortion of the original signal.

There are a number of more complicated denoising techniques that may be applied to the seismic data. These include principal component analysis (Hagen, 1982), f_x deconvolution (Canales, 1984), eigen image (Gulunay, 1986), Karhunen-Loève transform (Jones and Levy, 1987), time varying bandpass filters (Yilmaz, 1987), t_x prediction filtering (Abma and Claerbout, 1995), Optimum (Wiener) filters (Douglas A, 1997), band-pass, f_k, and kx_ky filtering (Yilmaz, 2001), artificial neural networks (Djarfour et al., 2008), Cadzow filtering (Trickett, 2008), S-transform (Askari and Siahkoohi, 2008), Fuzzy methods (Hashemi et al., 2008), singular spectrum analysis (Oropeza and Sacchi, 2011), sparse transform based denoising (Chen et al., 2016), mathematical morphology based denoising approach (Li et al., 2016), damped multichannel singular spectrum analysis (Huang et al., 2016), and the non-local means (NLM) algorithm (Bonar and Sacchi 2012). More effective approaches to enhance the signal are using array processing techniques such as f-k analysis (e.g., Naghizadeh 2011) or f-x filtering (Bekara and van der Baan, 2009; Naghizadeh and Sacchi, 2012; Chen and Ma, 2014). However, these methods require sufficient coherency in the arrival across the array. The other filter typically used are match filters (Eisner et al., 2008) which require an a-priori event with high SNR to act as a "master" or template event for the cross correlation with the continuous waveform. Recently Han and van der Baan (2015) used ensemble empirical mode decomposition and adaptive thresholding for the microseismic denoising. However, their method is not fully automated and has some parameters that need to be manually adjusted.

Seismic time series are nonstationary signals and, as a result, analysis in the time or

frequency domains alone may not provide adequate information. Hence, time-frequency methods that analyze frequency components and their amplitudes occurring as a function of time significantly increases the information that can be utilized in understanding a signal. Time-frequency transforms have been the basis of seismic denoising procedures used by Galiana-Merino et al., (2003), Pinnegar and Eaton (2003), Sobolev and Lyubushin (2006), Parolai (2009), Tselentis et al., (2012), and, Shuchong and Xun (2014).

This paper introduces an adaptive noise-estimation and denoising algorithm for seismic data based on the short time Fourier transform (STFT), an improved, minimally controlled recursive averaging estimator (Cohen, 2003), and neighboring thresholding in the STFT domain (Cai, and Silverman, 2001).

In the following section, we first describe the theory behind time-frequency denoising methods. Next, the noise-estimation and denoising approaches used in this study will be presented. At the end, the denoising algorithm will be tested by synthetic and real data examples.

2. Theoretical Background

Denoising methods for removing Gaussian random noise emerged after the pioneering works of Donoho and Johnstone (1994; 1995) who introduced nonlinear wavelet estimators. They showed that a simple thresholding in the wavelet domain results in nearly minimax error. Their well-known "VisuShrink" algorithm (Donoho and Johnston,

1994) thresholds wavelet coefficients one-by-one based on the universal threshold $\lambda_U = \sqrt{2\sigma^2 \ln(N)}$, in which $\mathcal{S}$ is the noise standard deviation and N is the length of the noisy signal). Time-frequency coefficients can be thresholded either by hard-thresholding or soft-shrinkage rules. In the hard-thresholding scheme, coefficients smaller than the threshold value λ_U are set to zero, while, in the soft-thresholding (or shrinkage) scheme larger coefficients are also shrunk toward zero.

In "SureShrink" of Donoho and Johnstone (1995) the same rules are used but threshold values are estimated for each level of decomposition through minimization of Stein's estimate of risk (Stein, 1981). Although improving performance compared to the previous denoising algorithms, these methods remained diagonal since threshold coefficients were solely based on their individual magnitudes. This can leave some isolated noise untouched. Cai (1999) studied non-diagonal estimators by proposing the "BlockJS" algorithm. In this method coefficients are divided into disjoint blocks and are thresholded group-by-group based on values of all coefficients in each block to increase the estimation precision by utilizing information about neighboring coefficients. The performance of denoising in this manner was improved by "NeighCoef" algorithm of Cai and Silverman (2001) where wavelet coefficients are shrunk one-by-one by an amount that depends on the coefficient and its neighboring coefficients inside overlapping blocks. However, block size and threshold level are kept constant in these algorithms. Block size and threshold level play important roles in the performance of a block thresholding estimator. Thus, Cai and Zhou (2009) proposed a data-driven approach to empirically select both the block size and threshold level in the "SureBlock" procedure by

minimizing Stein's unbiased risk estimator (SURE) (Stein, 1981).

Application of denoising methods based on block-thresholding on audio signals (Yu and Mallet, 2008), image denoising (Dengwen and Wengang, 2008), and seismic signals (Mousavi and Langston, 2016) show promising results for improving the SNR.

In our previous study (Mousavi and Langston, 2016) we showed that an improved wavelet block-thresholding in a hybrid scheme removes random noise from seismic data while retaining the main features of the seismic signal. However, the proposed method relies on an estimation of signal arrival and is relatively slow for real- time applications. In this paper a noise suppression algorithm is developed based on automatic noise level estimation and neighboring block thresholding in the short time Fourier transform (STFT) domain. The proposed method is faster and can adaptively estimate the noise level in STFT domain. The optimal threshold and block size are automatically adjusted by minimizing the SURE.

METHOD

The seismic signal *s*, contaminated by independent random noise *d*, can be modeled by:

$$y[i] = s[i] + \sigma d[i], i = 1, \dots, N, \quad (1)$$

where, *y* is the recorded signal and σ is the noise level. STFT decomposes the signal y, into time-frequency atoms $Y[\ell, k]$ where ℓ and *k* are the time and frequency localization

indices:

$$Y_{k\ell} = \sum_{i=1}^{N} y[i]w[i - \ell u]\exp(i2\pi kn / K) \quad . \tag{2}$$

$w[i]$ is a time window of support size *K*, which is shifted with a step $u \leq K$. ℓ and *k* are, respectively, the time and frequency indices with $0 \leq \ell \leq N/u$ and $0 \leq k \leq K$. In this study we use the Hanning window with $u = K/2$.

The denoised signal is obtained by applying a time-frequency attenuation factor $a_{k\ell}$, over time-frequency atoms of the noisy signal $Y[\ell, k]$ to obtain an estimate of the Fourier coefficients of the seismic signal $\tilde{S}_{k\ell}$ that are as similar as possible to the $S_{k\ell}$:

$$\tilde{S}_{k\ell} = a_{k\ell} Y_{k\ell}. \tag{3}$$

Once the estimate of $\tilde{S}_{k\ell}$ for each Fourier coefficient has been obtained, the time domain signal can be recovered using the inverse short time Fourier transform.

2.1. Noise-Level Estimation

The estimate is usually done under the assumption that the noise-level is known. However, the noise-level is generally unknown in real cases and needs to be estimated from the observed data y. Noise level can be estimated based on the principle of minimum statistics (Martin, 2001). This approach assumes that the signal is stronger than

the noise, thus, if we track the minima of a noisy power spectrum with a sliding window we will have an estimate proportional to the background noise power. Following this assumption the variance s^2, of Fourier coefficients of the noise coefficients $D_{k\ell}$, can be estimated by averaging past spectral power values of noisy measurements using a smoothing parameter in the recursive averaging approach (Cohen 2001):

$$\tilde{\sigma}_d^2(k,\ell+1) = \alpha_d \tilde{\sigma}_d^2(k,\ell-1) + (1-\alpha_d)\left|D(k,\ell)\right|^2 \quad , \tag{4}$$

where, the $\tilde{\sigma}_d^2(k,\ell-1)$corresponds to the estimation of the noise variance for the previous frame and $\alpha_d (0 < \alpha_d < 1)$ is the smoothing parameter.

In the minima controlled recursive averaging of Cohen (2001) the recursive averaging is done using a time-varying frequency-dependent smoothing parameter that is adjusted by the probability of signal presence. To develop this, define two hypotheses $H_0(k,\ell)$ and $H_1(k,\ell)$that indicate, respectively, signal absence and presence in the *k*th frequency bin of the ℓth frame. In this technique, past spectral power values of noisy measurements during a period of signal absence are recursively averaged and the estimate is continued during signal presence:

$$\tilde{\sigma}_d^2(k,\ell+1) = \begin{cases} \alpha_d \tilde{\sigma}_d^2(k,\ell) + (1-\alpha_d)\left|Y(k,\ell)\right|^2 & if \quad H_0 \quad is \quad true \\ \tilde{\sigma}_d^2(k,\ell) & if \quad H_1 \quad is \quad true \end{cases} . \tag{5}$$

Considering the probabilities of the two hypotheses $P(H_0|Y)$ and $P(H_1|Y) = 1 - P(H_0|Y)$, it is possible to join the two estimates into one single average estimator as: (Cohen, 2003)

$$\tilde{\sigma}_d(k,\ell+1) = P(H_0(k,\ell)\,|\,Y(k,\ell))(\alpha_d\tilde{\sigma}_d^2(k,\ell) + (1-\alpha_d)\left|Y(k,\ell)\right|^2) + P(H_1(k,\ell)\,|\,Y(k,\ell))(\tilde{\sigma}_d^2(k,\ell))$$

. (6)

Equivalently, recursive averaging can be obtained by using:

$$\tilde{\sigma}_d^2(k,\ell) = \hat{\alpha}_d(k,\ell)\tilde{\sigma}_d^2(k,\ell) + (1-\hat{\alpha}_d(k,\ell))\left|Y(k,\ell)\right|^2 \quad , \qquad (7)$$

where $\hat{a}_d(k,\ell)$ is a time-varying frequency-dependent smoothing parameter that is adjusted by the probability of the signal presence estimated on noisy measurements. It is given by:

$$\hat{a}_d(k,\ell) = a_d + (1 - a_d)p(k,\ell) \, , \qquad (8)$$

where, $p(k,\ell) = P(H_1(k,\ell)\,|\,Y(k,\ell))$ is the conditional probability of the signal presence that is calculated using *a posteriori* and *a priori* SNRs by applying Baye's rule: (Cohen, 2003)

$$p(k,\ell) = \left(1 + \frac{q(k,\ell)}{1 - q(k,\ell)}(1 + x(k,\ell))\exp(-v(k,\ell))\right)^{-1} \quad , \qquad (9)$$

and where

$$\upsilon = \frac{\gamma\xi}{(1+\xi)} \quad , \tag{10}$$

and the *a posteriori* SNR $\gamma(k,\ell)$, and *a priori* SNR $\xi(k,\ell)$ are given respectively by:

$$\gamma(k,\ell) = \frac{|Y(k,\ell)|^2}{\sigma_d^2(k,\ell)} \quad . \tag{11}$$

$$\xi(k,\ell) = \frac{\sigma_s^2(k,\ell)}{\sigma_d^2(k,\ell)} \, . \tag{12}$$

The "decision–directed" approach of Ephraim and Malah (1984) is used to estimate the *a priori* SNR:

$$\tilde{\xi}(k,\ell) = \alpha G_{H_1}^2(k,\ell-1)\gamma(k,\ell-1) + (1-\alpha)\max\{\gamma(k,\ell),0\}, \tag{13}$$

where, α is a weighting factor that controls the tradeoff between noise reduction and signal distortion (Ephraim and Malah, 1984). G_{H_1} is the spectral gain factor (Ephraim and Malah, 1985).

$q(k,\ell) = P(H_0(k,\ell)|Y(k,\ell))$ is the *a priori* probability for signal absence which has

values ranging from 0 to 1 and can be estimated by the minima controlled estimation procedure proposed by Cohen (2003). This procedure basically consists of two iterations of smoothing and minimum tracking. In the first iteration a rough estimate of signal presence in each frequency band is made. Then, in the second iteration, the smoothing excludes relatively stronger signal components.

After noise estimation we normalize time-frequency coefficients using this estimation. The time–frequency plane is then segmented into disjoint rectangular macroblocks M_j, of length L_m in time and width W_m in frequency.

2.2. Time-Frequency Denoising

Each Fourier coefficient is shrunk using an attenuation factor, $a_{k\ell}$, calculated based on the magnitudes of the neighboring coefficients in a square block B_i, with size $L_b \times L_b$ centered on the coefficient (Figure 1), where L_b is a positive odd number.

$$a_{k\ell} = (1 - \frac{\lambda^2}{S^2})_+ \quad . \tag{14}$$

$\lambda \geq 0$ is the threshold level and $S^2 = \sum_{k,\ell \in B_i} Y_{k\ell}^2$ is the L^2-energy of the noisy coefficients in the neighboring block B_i. $(g)_+$ is defined as: $(g)_+ = \max(g, 0)$.

Optimal block size L_b^* and threshold level λ^*, inside each macroblock are chosen by minimizing the mean squared error or expected risk $R = E\{\|s - \tilde{s}\|^2\}$. However, risk

cannot be calculated since, in reality, s is not known. So, following Cai and Zhou (2009) we estimate the risk by Stein's unbiased risk estimate (SURE) using the observed data.

Stein (1981) showed that given a multivariate normal observation $y \sim N(\mu, I)$, $\mu \in \mathbb{R}^p$, and a nearly arbitrary estimator $\tilde{\mu} = y + h(y)$ with $h: \mathbb{R}^p \to \mathbb{R}^p$ being weakly differentiable and $\nabla h = \sum_{i=1}^{L} (\frac{\partial}{\partial y_i}) h_i(y)$. If $E\left\{ \sum_{i=1}^{L} \left| (\frac{\partial}{\partial y_i}) h_i(y) \right| \right\} < \infty$, then the risk of estimator n for normalized coefficients by noise level ($s^2 = 1$) can be estimated by:

$$\tilde{R} = L + \left\| h(y) \right\|_2^2 + 2\nabla h(y). \tag{15}$$

Similarly over the block B_i, the estimate of the signal coefficients, $\tilde{S}$, can be written as:

$$\tilde{S}(k,\ell) = a(k,\ell) Y(k,\ell) = Y(k,\ell) + h(Y(k,\ell)). \tag{16}$$

According to (14) for the nth coefficient Y_n, we have:

$$h_n(Y_n) = \tilde{S}_n - Y_n = \begin{cases} -\dfrac{\lambda^2}{S^2} Y_n & (S_n > \lambda) \\ -Y_n & (S_n \le \lambda) \end{cases}. \tag{17}$$

Then the risk within one block can be estimated by:

$$SURE(Y, \lambda, L^2) = L^2 + \sum_{n=1}^{L^2} \left\| h_n(Y_n) \right\|_2^2 + 2\sum_{n=1}^{L^2} \frac{\partial h_n}{Y_n}, \tag{18}$$

where

$$\left\| h_n(Y_n) \right\|_2^2 = \begin{cases} \dfrac{\lambda^4}{S_n^4} Y_n^2 & (S_n > \lambda) \\ Y_n^2 & (S_n \le \lambda) \end{cases}, \tag{19}$$

and

$$\frac{\partial h_n}{\partial Y_n} = \begin{cases} -\lambda^2 \dfrac{S^2 - 2Y_n^2}{S_n^4} Y_n^2 & (S_n > \lambda) \\ -1 & (S_n \le \lambda) \end{cases}. \tag{20}$$

The optimal block size L_*^2 and threshold level λ_* then can be selected by:

$$(\lambda, L_*^2) = \arg\min SURE(Y, \lambda, L^2) . \tag{21}$$

2.3. Summary of the algorithm

The outline of the algorithm can be summarized as follows:

1. Transform the time-series data into the time–frequency atoms through the STFT (equation 2).

2. Estimate the noise level from STFT coefficients signal using recursive averaging

(equations 7) .

3. The optimal block size L* and threshold level λ* for coefficients are estimated by minimizing Stein' s unbiased risk estimate (equation 21).

4. Time-Frequency coefficients inside each block are modified using calculated attenuation factor (equation 14).

5. The final denoised signal is obtained via the inverse STFT of modified coefficients.

3. Results

We first apply the algorithm to noise-contaminated version of a synthetic seismic signal used in Mousavi and Langston (2016). In this way, the shape of the denoised signal and the onset time of the first pulse can be compared to the synthetic signal and results of the two denoising methods can be compared. Afterwards, the denoising algorithm will be applied to field seismic data.

3.1. Synthetic Test

A local synthetic seismogram (Figure 2) and its contaminated version with random noise are used for a synthetic test. Result of the denoising algorithm and the comparison of its performance with other methods is shown in Figure 3 and Table 1 respectively. Performance of denoising algorithm improves as SNR of the input increases. In Table 1 performance of the proposed method on the synthetic signal for the worst case (SNR=1.3) is compared to the bandpass filtering, hard- and soft-thresholding, and hybrid block thresholding (Mousavi and Langston, 2016). Proposed method does not achieve the performance of soft and hybrid block thresholdings. However, its performance is quite

close to hard thresholding and superior to the bandpass filtering. Comparing results with the hybrid approach in Mousavi and Langston, (2016), the accuracy of denoising in terms of preserving waveform shape and polarity of emergent arrivals, is lower. The root-mean-square-error (RMSE) between the denoised and synthetic signal is 0.06 and maximum cross-correlation coefficient is only 0.7. However, this method is an order of magnitude faster to implement and can be used for real time processing without knowing any pre-information about events such as arrival times and noise level.

The algorithm is successful in removing the background noise and improving the SNR significantly. However, the signal-leakage energy is relatively high. Data around the maximum amplitude are mostly preserved but the P and S coda parts are strongly attenuated. The polarity of the S phase is preserved (Figure 3). Compared to the noise estimation in the wavelet domain (Donoho and Johnston, 1994), recursive averaging's estimate is biased toward higher values. The overestimation of noise level can cause some part of the signal energy to be attenuated as well as the noise. However, this might not be a problem in some applications such as event detection where the focus of the denoising is more on improving the SNR. Moreover, the signal leakage can be moderated by either improving the automatic noise estimation to attenuate less signal energy or using approaches such as signal-and-noise orthogonalization (Chen and Fomel, 2015) to predict the attenuated signal energy and retrieve it back.

Sensitivity of the denoising algorithm to noise level has been presented in Figure 4. Performance of the method is very good (RMSE < 0.02) when SNR of input is greater than 5. Increase in denoising error is more severe for SNR of lower than 1.5. We found that the window size for the time-frequency transform and recursive averaging was very important in applying the algorithm. It is better to select a window size long enough to contain both signal and noise samples.

3.2. Field Seismic Data

Background noise is a challenging problem encountered in surface monitoring of microseismic events. We have applied the algorithm to two data sets. The first data set consists of a two-hour length of vertical record containing microseismic events induced during wastewater injection in central Arkansas (Figure 5). This seismogram was recorded by a broadband seismometer (Station CH2 at 35.3341N, -92.2982W) at the surface in 2011. Hypocentral distances are between 1.7 to 2.5 km and event sizes range between 0.32 to 0.78 Mw. The second data set includes borehole, near-surface, and surface records of microearthquakes associated with an underground collapse of a cavern in Bayou Corne, Louisiana, in 2013.

In the case of microearthquakes induced by wastewater injection (Figure 5), noise with higher and lower frequency content, compared to the signal frequency, was attenuated from the time-frequency representation of the data and SNR was increased. In addition to wide spread random noise, strong coherent noises exist at around 18, 27, and 52 Hz. These continuous noises are present in almost entire time span and removal of these

concentrated noises after denoising indicates that the proposed algorithm is capable of removing colored noises as well. Seven microseismic events are marked in denoised trace. Zoomed windows around each event before and after the denoising are presented in Figure 6. As you can see there coda energy was removed from seismograms after the denoising. There are some changes in the amplitudes after the denoising due to the superposition of noise and seismic signal and effects of the denoising by modifying time-frequency coefficients. However, overall waveform shapes and polarity remained the same.

Second data set consists of record of microseismic events recorded by three receivers installed at the bottom, and middle of a single borehole located at 30.0134° N, 91.1439° W and one broadband instrument at the surface near the borehole (located at 30.0087° N, 91.1398° W almost 1 km south east of the borehole) at Bayou Corne, Louisiana. In this study we have just used 7.5 minutes long vertical component seismograms recorded by each instrument on 1 November 2013 to verify performance of the proposed algorithm (Figure 7). As we go from bottom of the borehole to the surface (Figure 7a-1 to c-1) noise level is increased. Seven events associated with the underground collapse of a cavern near the borehole are clearly observable on recorded seismogram by a 2-Hz geophone at the bottom (~287 m deep) of the borehole (Figure 7a-1). However, these events are not clear on the data recorded by 3C broadband sensors in the same borehole near the surface (at ~ 190 m deep)(Figure 7b-1) and at the surface (1 km south east of the borehole) (Figure 7c-1). SNR for the near surface data improved after band-pass filtering between 2 and 20 Hz (Figure 7b-2). This is because most of the noise in near surface has lower-

frequency compared to microseismic events. Hence, spectral filtering helps to reveal most of the events that were covered by background noise. However, filtering does not improve SNR of surface data (Figure 7c-2) because of the presence of some high-frequency noise within the frequency range of the seismic events. In the right panel data are presented after denoising by our proposed method. Denoising is successful in removing the noise and significantly improving the SNR in both near surface and surface data (Figure 7b-3 and c-3). However, two small events at 150 and 300 seconds are missing in the surface data after denoising. One reason for this is that the surface sensor is located further from two other sensors and possibly the source of these two events. The one kilometer distance between this surface sensor and the borehole sensors can be enough to attenuated the energy of these smaller events. However, revealing microearthquakes that were buried under background noise has special importance for microseismic detection, which is a challenging problem in surface monitoring of microseismic events.

4. Discussion and conclusion

Typical microseismic data are characterized by low signal-to-noise ratios (SNR) and highly non-stationary noise and typically rely on array or cross-correlation based approaches to enhance the SNR prior to processing. In this article we have provided an alternative approach of increasing the SNR of the microseismic data recorded on a single channel by filtering the data. The proposed algorithm is based on neighboring thresholding in the short time Fourier transform domain. An automatic noise level estimation based on minimum controlled recursive averaging is embedded in the

algorithm. Using the short time Fourier transform and automatic noise estimation method makes the method fast and data-driven. The performance of proposed algorithm was tested on synthetic and real seismic data.

Denoising microseismic data using the proposed method does not need coherent arrivals in an array, a master event with high SNR, or parameters that need to be tuned manually. It is adaptive to data type and can effectively attenuate the background noise in an automatic procedure. The proposed method is a powerful denoising approach well suited for real time applications and can significantly improve SNR. Denoising the recorded data as proposed here will significantly lower the detection threshold and allow small seismic events to be detected more easily. Time-frequency denoising can be combined with detection and phase picker algorithms in the time-frequency domain such as those proposed by Karamzadeh et al, (2013), Bogiatzis and Ishii, (2015) or Mousavi et al, (2016). This will be of interest particularly to the denoising of microseismic data acquired using surface arrays that can show strong spatial noise variability and to seismic hazard studies which are typically acquired using 10's or fewer seismometers. However, the proposed method can have other applications, such as attenuation estimations or seismic hazard studies. Denoising the signal can assemble the spectral content more precisely and decrease the uncertainties in Q estimation (McNamara et al., 2012; Mousavi et al., 2014), improving the ability of models to match observations. This method has potential to be used in seismic hazard studies in urban areas where strong spatial noise variability exists (e.g. Mousavi et al., 2011).

Acknowledgment

We thank Paul Ogwari for providing the information about microseismic events. This study was supported by the Air Force Research Laboratory under Contract #FA9453-16-C-0015. We thank editor-in-chief of journal of applied geophysics, Jianghai Xia, Yangkang Chen, and anonymous reviewer for their insightful remarks and constructive comments.

References

Askari, R., Siahkoohi, H.R., 2008. Ground roll attenuation using the S and x-f-k transforms. Geophys Prospect 56, 105-114, doi: 10.1111/j.1365-2478.2007.00659.x.

Astudillo R.F., Orglmeister, R., 2013. Computing MMSE Estimates and Residual Uncertainty Directly in the Feature Domain of ASR using STFT Domain Speech Distortion Models, IEEE Trans. Speech Audio Process 21, 1023–1034.

Bekara, M., van der Baan, M. 2009. Random and coherent noise attenuation by empirical mode decomposition. Geophysics 74**,** V89– V98, doi: 10.1190/1.3157244.

Bogiatzis P., Ishii, M., 2015. Continuous Wavelet Decomposition Algorithms for Automatic Detection of Compressional- and Shear-Wave Arrival Times. Bull. Seism. Soc. Am. **105,** 1628–1641, doi: 10.1785/0120140267.

Bonar D., Sacchi, M., 2012. Denoising seismic data using the nonlocal means algorithm.

Geophysics 77, A5–A8.

Cai T., 1999. Adaptive wavelet estimation: A block thresholding and oracle inequality approach. Ann. Statist. 27, 898–924.

Cai T., Silverman, B.W., 2001. Incorporation information on neighboring coefficients into wavelet estimation. Sankhya 63, 127–148.

Cai T., Zhou, H., 2009. A Data-driven block thresholding approach to wavelet estimation. Annals of Statistics **37**, 569-595. doi: 10.1214/07-AOS538.

Canales, L.L., 1984. Random noise reduction. SEG Technical Program Expanded Abstracts **3,** 525–527.

Chen Y., Ma, J., Fomel, S., 2016. Double-sparsity dictionary for seismic noise attenuation. Geophysics 81, V17-V30.

Chen Y., Fomel, S., 2015. Random Noise Attenuation using Local Signal-and-Noise Orthogonalization. Geophysics **80**, WD1-WD9. doi: 10.1190/GEO2014-0227.1

Chen, Y., Ma, J., 2014. Random noise attenuation by **f-x** empirical mode decomposition predictive filtering. Geophysics **79,** V81–V91, doi: 10.1190/geo2013-0080.1.

Cohen I., Berdugo, B., 2001. Speech enhancement for non-stationary noise environments.

Signal Processing **81**, 2403-2418.

Cohen I., 2003. Noise spectrum estimation in adverse environments: Improved minima controlled recursive averaging, IEEE Trans. Speech Audio Process **11**, 466–475.

Dengwen Z., Wengang, C., 2008. Image denoising with an optimal threshold and neighbouring window. Pattern Recognition Letters **29**, 1694-C1697.

Djarfour, N., Aïfa, T., Baddari, K., Mihoubi, A., Ferahtia, J., 2008. Application of feedback connection artificial neural network to seismic data filtering. Comptes Rendus Geoscience **340,** 335–344. doi: 10.1016/j.crte .2008.03.003.

Donoho, D., Johnstone, I.M., 1994. Ideal spatial adaptation by wavelet shrinkage. Biometrika **81**, 425–455.

Donoho, D.L., Johnstone, I.M., 1995. Adapting to unknown smoothness via wavelet shrinkage. J. Amer. Statist. Assoc **90**, 1200–1224.

Eisner, L., Abbott, D., Barker, W., Lakings, J., Thornton, M., 2008. Noise suppression for detection and location of microseismic events using a matched filter. 78th Annual International Meeting, SEG, Expanded Abstracts, 1431–1435.

Eisner. L., Grechka, V., Williams-Stroud, S., 2010. Future of Microseismic Analysis:

Integration of Monitoring and Reservoir Simulation. AAPG Hedberg Conference, December 5-10, 2010, Austin, Texas.

Ephraim Y., Malah, D., 1984. Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator. IEEE Trans. Acoust. Speech. Signal Process. **32**, 1109–1121,

Galiana-Merino, J. J., Rosa-Herranz, J., Giner, J., Molina, S., Rotella, F., 2003. De-noising of short period seismograms by wavelet packet transform. Bull. Seismol. Soc. Am. **93**, 2554–2562.

Gulunay N., 1986. FXDECON and complex Wiener prediction filter: 56th Annual International Meeting, SEG, Expanded Abstracts, 279–281.

Hagen, D.C., 1982. The application of principal components analysis to seismic data sets: Geoexploration: International Journal of Mining and Technical Geophysics and Related Subjects **20**, 93–111. doi: 10.1016/ 0016-7142(82)90009-6.

Han, J., van der Baan, M., 2015. Microseismic and seismic denoising via ensemble empirical mode decomposition and adaptive thresholding. Geophysics **80**, 69-80.

Hashemi, H., Javaherian, A., Babuska, R., 2008. A semi-supervised method to detect seismic random noise with fuzzy GK clustering. Journal of Geophysics and Engineering

5, 457–468. doi: 10.1088/1742-2132/5/ 4/009.

Huang, W., R. Wang, Y. Chen, H. Li, and S. Gan, 2016. Damped multichannel singular spectrum analysis for 3D random noise attenuation, Geophysics, 81, 4, V261-V270.

Jones, I. F., Levy, S., 1987. Signal-to-noise ratio enhancement in multichannel seismic data via the Karhunen-Loève transform. Geophysical Prospecting **35,** 12–32, doi: 10.1111/j.1365-2478.1987.tb00800.x.

Karamzadeh, N., Javan Doloei, G., Reza, A., 2013. Automatic Earthquake Signal Onset Picking Based on the Continuous Wavelet Transform. IEEE Trans. Geosci. Remote. Sens. **51**, 2666–2674, doi: 10.1109/TGRS.2012.2213824.

Li H., Wang, R., Cao, S., Chen, Y., Huang, W., 2016. A method for low-frequency noise suppression based on mathematical morphology in microseismic monitoring. Geophysics 81, 3, V159-V167.

Martin R., 2001. Noise power spectral density estimation based on optimal smoothing and minimum statistics. IEEE Trans. on Speech and Audio Processing **9**, 504–512.

McNamara, D. E., M. Meremonte, J. Z. Maharrey, S.-L. Mildore, J. R. Al-tidore, D. Anglade, S. E. Hough, D. Given, H. Benz, L. Gee, and A. Frankel, 2012, Frequency-dependent seismic attenuation within the His- paniola Island region of the Caribbean sea:

Bulletin of the Seismological Society of America, 102, 773–782, doi: 10.1785/0120110137.

Mousavi, S.M., Langston, C.A., Horton, S.P., 2016. Automatic Microseismic Denoising and Onset Detection Using the Synchrosqueezed-Continuous Wavelet Transform. Geophysics **81**, 1-15, doi: 10.1190/GEO2015-0598.1.

Mousavi, S.M., Langston, C.A., 2016. Hybrid Seismic Denoising Using Higher-Order Statistics and Improved Wavelet Block Thresholding. Bulletin of Seismological Society of America **106**, doi: 10.1785/0120150345.

Mousavi, S. M., C. H. Cramer, and C. A. Langston, 2014, Average QLg, QSn, and observation of Lg blockage in the continental margin of Nova Scotia: Journal of Geophysical Research, 119, 7722–7744, doi: 10.1002/ 2014JB011237.

Mousavi, S. M., Omidvar B., Ghazban F., and Feyzi R., 2011, Quantitative risk analysis for earthquake-induced landslides—Emamzadeh Ali, Iran. Eng Geol 122:191–203. doi:10.1016/j.enggeo.2011.05.010.

Maxwell, S., 2005. A Brief Guide to Passive Seismic Monitoring. CSEG National Convention, 177–178.

Naghizadeh, M., 2011. Seismic Data Interpolation and Denoising in the Frequency-

Wavenumber Domain. Geophysics **77**, V71-V80. doi: 10.1190/GEO2011-0172.1.

Naghizadeh, M., Sacchi, M., 2012. Multicomponent f-x Seismic Random Noise Attenuation via Vector Autoregressive Operators. Geophysics **77**, V91-V99. doi: 10.1190/geo2011-0198.1.

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

Parolai S., 2009. Denoising of Seismograms Using the S Transform. Bull. Seismol. Soc. Am. **99,** 226–234. doi: 10.1785/0120080001.

Pinnegar C.R. Eaton, D.W., 2003. Application of the S transform to prestack noise attenuation filtering. J. Geophys. Res. **108**, 2422. doi: 10.1029/2002JB002258.

Sobolev, G., Lyubushin, A., 2006. Microseismic impulses as earthquake precursors: Izvestiya. Physics of the Solid Earth **42**, 721–733. doi: 10.1134/S1069351306090023.

Shuchong L., Xun, C., 2014. Seismic signals wavelet packet de-noising method based on improved threshold function and adaptive threshold. Computer Modeling and New Technologies **18**, 1291-1296.

Stein C., 1981. Estimation of the mean of a multivariate normal distribution. Ann. Statist

9, 1135–1151.

Trickett, S., 2008. F-xy cadzow noise suppression: SEG Technical Program Expanded Abstracts **27**, 2586–2590.

Tselentis G.A., Martakis, N., Paraskevopoulos, P., Lois., A., Sokos, E., 2012. Strategy for automated analysis of passive microseismic data based on S-transform, Otsu's thresholding, and higher order statistics. Geophysics **77**, KS43-KS54. doi: 10.1190/GEO2011-0301.1.

Yilmaz, Ö., 1987. Seismic Data Processing, 1st ed., Soc. of Explor. Geophys., Tulsa, Okla., 1987

Yilmaz, Ö., 2001. Seismic data analysis: processing, inversion, and interpretation of seismic data: Society of Exploration Geophysicists, 2 edition.

Yu Y., Mallat, S., Bacry, E., 2008. Audio Denoising by Time-Frequency Block Thresholding, IEEE Trans. on Signal Processing **56**, 1830-1839.

Table 1

Root-mean-square error (RMSE), signal-to-noise ratio (SNR), and maximum correlation coefficients between denoised and original signal (CC) from the synthetic test using bandpass filtering between 5 and 20 Hz, Hard and Soft Thresholding (Donoho and Johnston, 1994), and Hybrid Block Thresholding (Mousavi and Langston, 2016),

Method	RMSE	CC	TFT	Elapsed Time (s)
Bandpass Filtering	0.063	0.683	FT	0.19
Hard Thresholding	0.061	0.796	CWT	0.50
Soft Thresholding	0.048	0.833	CWT	0.51
Hybrid Block Thresholding	0.027	0.935	CWT	9.23
Neighboring Thresholding	0.06	0.721	STFT	0.87

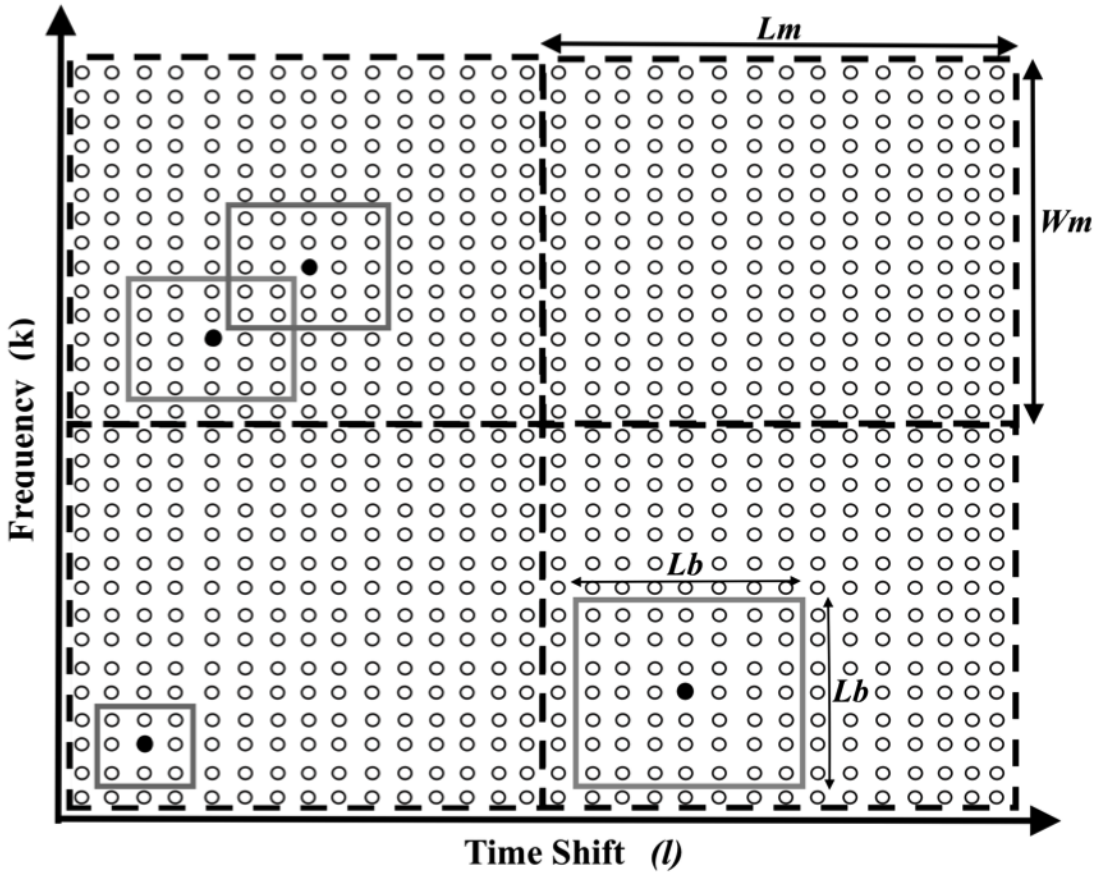

Figure 1. Schematic representation of partitioning time-frequency coefficients shown by open circles into macro blocks and neighboring blocks. Macro blocks are shown with the dashed lines where L_m and W_m are length and width of macroblocks respectively. Gray solid lines are representing blocks of neighboring coefficients with size $L_b \acute{} L_b$ used for calculation of the attenuation factor for the central coefficients shown by solid circles.

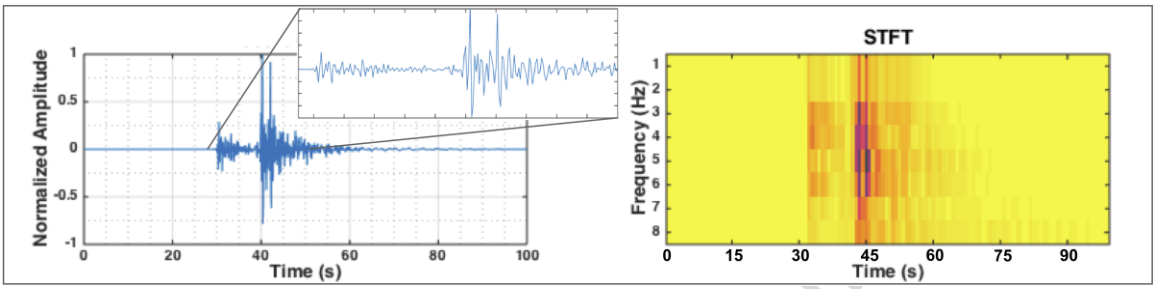

Figure 2. a) The synthetic seismogram and its associated STFT. Details on calculation of the synthetic signal are given in Mousavi and Langston (2015).

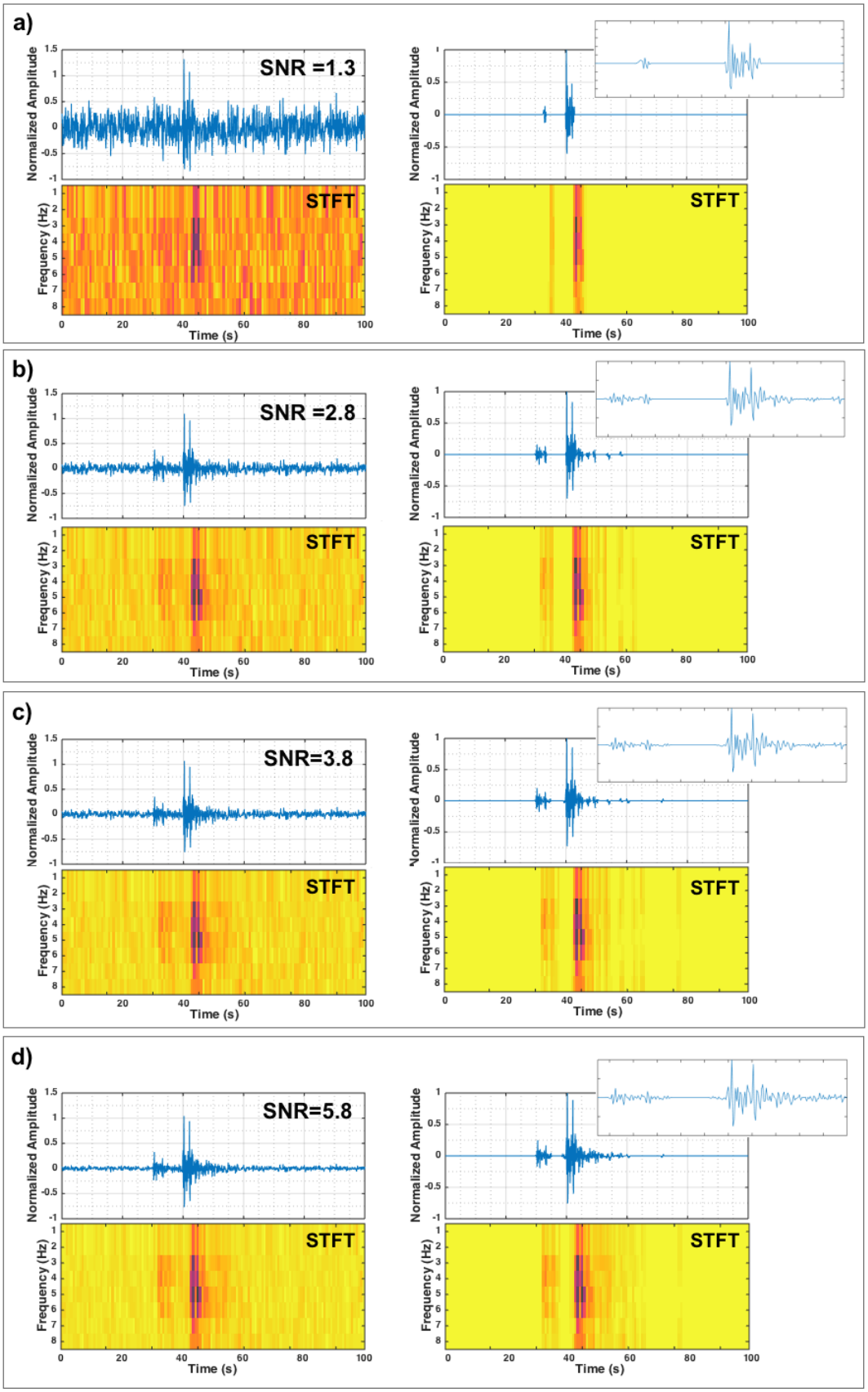

Figure 3. Left, synthetic signal plus random noise. Right, result of denoising algorithm and zoomed windows around P and S. Signal to noise ration of input data increases from top to bottom.

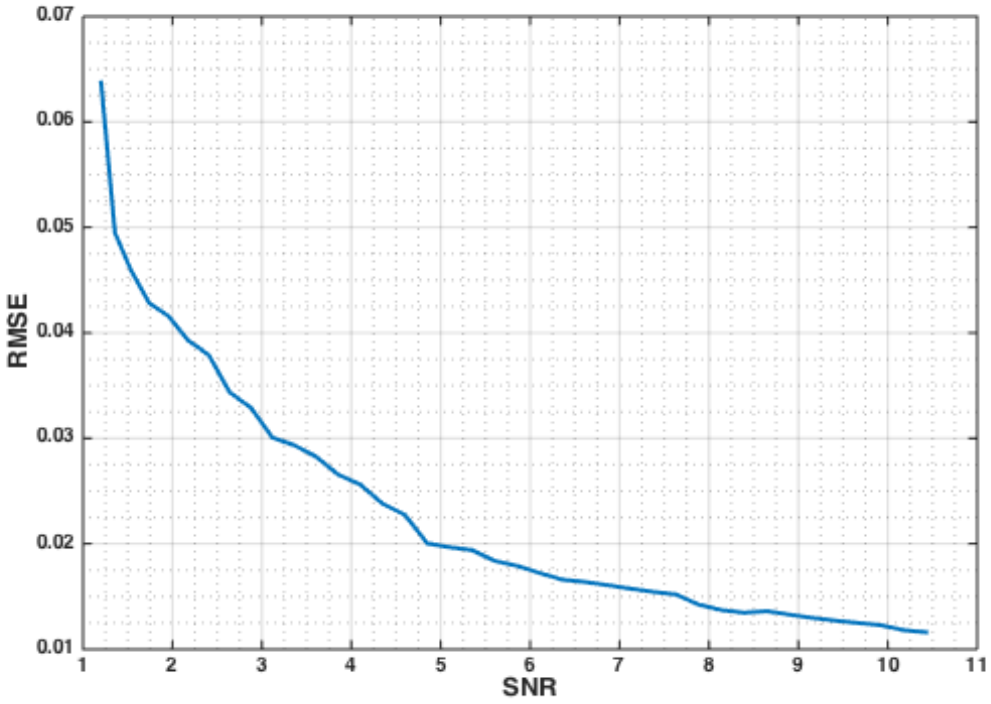

Figure 4. Sensitivity test of the proposed method to SNR of input.

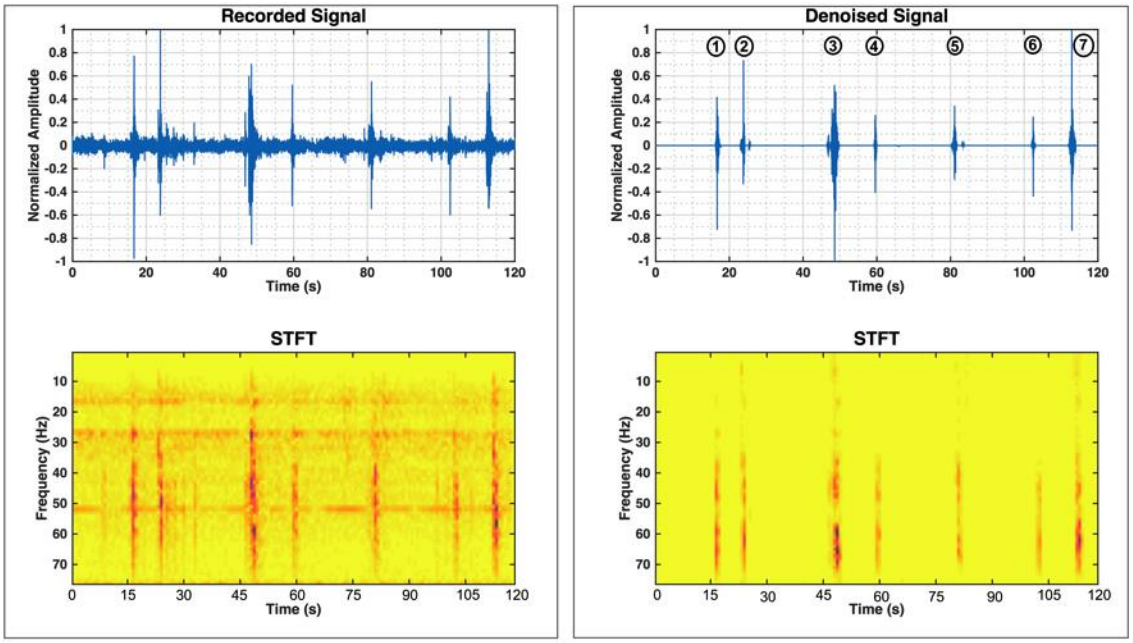

Figure 5. Denoising real seismic data. The left column shows presumably induced microseismic events due to wastewater injection in central Arkansas in 2010 recorded by a broadband seismometer at the surface. The right column shows the same trace and its STFT after denoising.

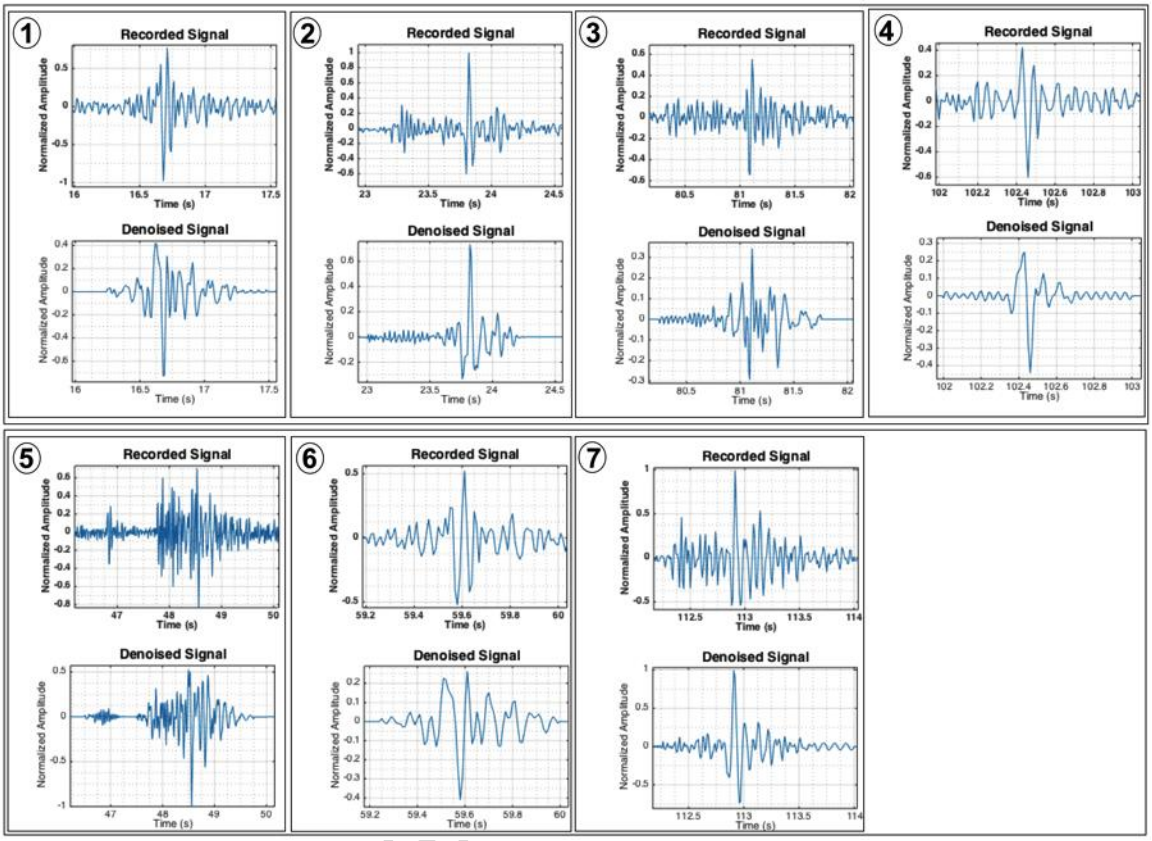

Figure 6. Zoomed windows around seven events marked on the figure 3 before (top plot in each panel) and after (bottom) the denoising.

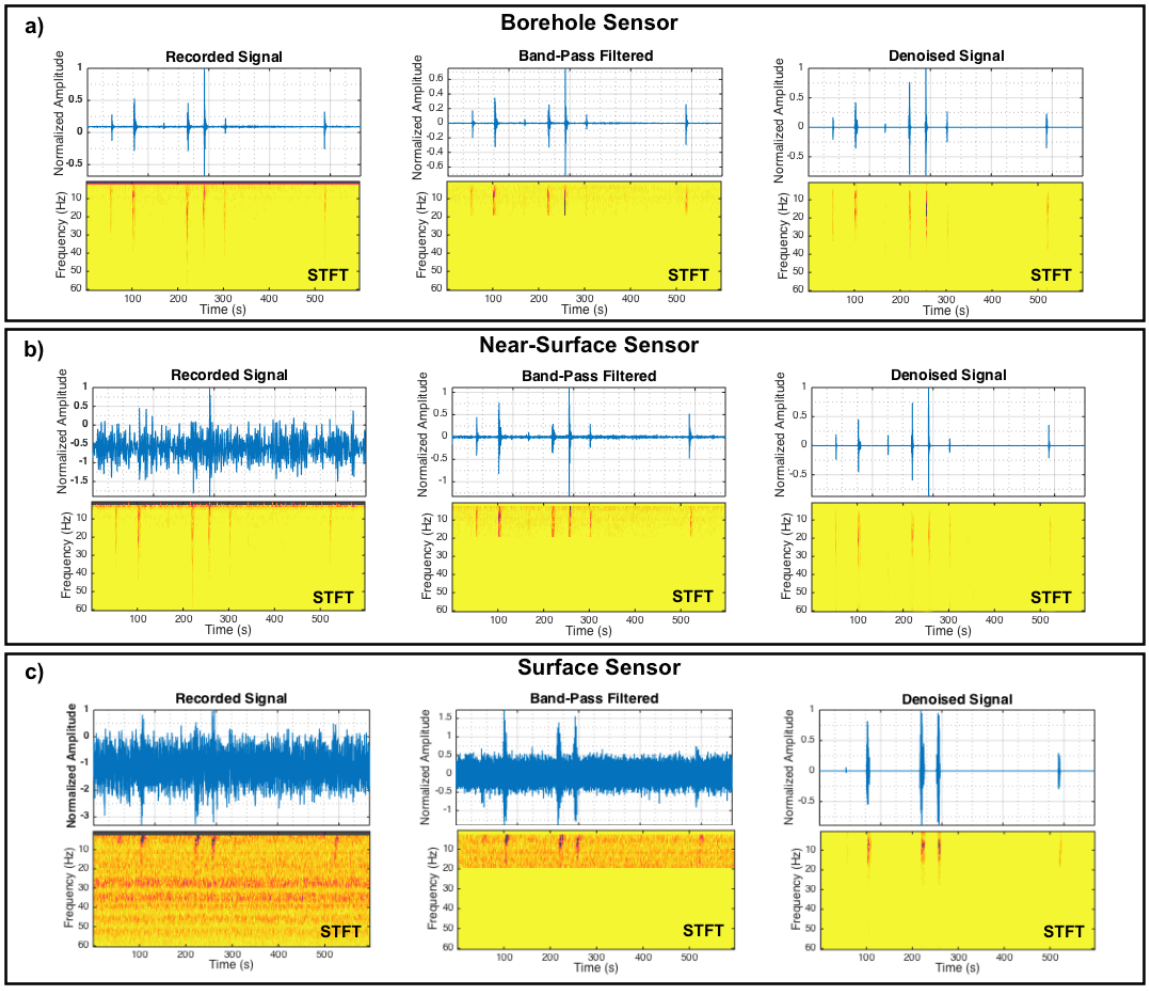

Figure 7. 7.5 minutes long vertical seismograms passively recorded (with 200 samples per second) in November 1st 2013 at Bayou Corne, Louisiana. a) A three-component 2-Hz geophone (LA17.01) at the bottom of a borehole (~287 m deep) located at 30.0134° N, 91.1439° W. b) A three-component broadband sensor (Trillium-compact) (LA17.02) at the top of the same borehole (~ 190 m deep). c) A three-component broadband sensor (Trillium-compact) (LA14) at the surface located at 30.0087° N, 91.1398° W (1 km south east of LA17). The left panel are raw data recorded in these stations and their short time Fourier transform (STFT). 7 microearthquakes induced by underground collapse of the cavern in the area are observable on the borehole data, while near-surface and surface

data are much noisier. Middle panel are the same traces after band-pass filtering between 2 and 20 Hz. As most of the noises in near-surface data (b) have lower-frequencies compared to microseismic events, spectral filtering helps revealing most of the events that were covered under background noise. However, filtering does not improve SNR at the surface data because of presence of some high-frequency noises within the frequency range of seismic events. In the right panel data are presented after denoing by proposed method of this study. Denoising is successful in removing the noise and significant improvement of the SNR.

Highlights

In this paper we have developed an automatic algorithm for noise level estimation and noise suppression of continuous seismic data based upon minimally controlled recursive averaging and neighborhood shrinkage estimators. The procedure is totally data driven, hence without any pre-adjustment can adopt all the parameters needed for the noise suppression from nature of the signal. Neighboring thresholding has been adjusted for STFT so it allows for the fast and effective noise suppression in the time-frequency domain. In this work our goal was more on developing a fast and automatic procedure to facilitate the microseismic event detection with high background noise level rather that the conventional seismic denoising where the focus is more on preserving detailed features of the signal.