Application of Multi-channel Wiener Filters to the Suppression of Ambient Seismic Noise in Passive Seismic Arrays

J. Wang[1], *F. Tilmann*[1], *R. S. White*[1], *H. Soosalu*[1] *and P. Bordoni*[2]

1. Bullard Laboratories, University of Cambridge, Madingley Road, CB3 0EZ, UK;

2. Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy.

For submission to The Leading Edge.

Corresponding author: jingbo@esc.cam.ac.uk

We are concerned with the detection and location of small seismic events, such as can be encountered in monitoring hydro-fracturing with surface sensors. Ambient seismic noise is the main problem in detection of weak seismic phases from these events, particularly as the sites of interest are often within or near producing fields. Bandpass filtering and stacking are the most widely used technique for enhancing the signal-to-noise ratio (SNR) in passive seismic experiments, but they are of limited value when both noise and signal share the same frequency band. Seismic arrays can be used to reduce the unwanted noise (e.g. traffic noise, pumping noise, scattering ground roll) by delay-and-sum techniques (also called beamforming) or by frequency-wavenumber filtering. Beamforming maximizes the array response for the assumed direction and slowness of the signal. Whereas in some situations it can be highly effective, and the azimuth and slowness of the signal can be determined by a grid search approach, it is vulnerable to contamination by side-lobe energy, particularly for broadband signals and noise (Rost and Thomas, 2002). Frequency-wavenumber filtering can be very effective but requires regularly spaced arrays and implicitly assumes plane wave propagation. Both methods perform poorly when the waveform changes significantly between stations of the array, as might be caused, for example, by differences in site response.

In this article, we present a multi-channel Wiener filtering technique, which allows the removal of coherent noise from three-component 2D arrays without making *a priori* assumptions about the mode of propagation (e.g., no plane-wave assumption is required for the noise field). We test the effectiveness of this filter with two case studies. In the first

case, we add synthetic signals of varying strengths to actual noise data recorded with a hexagonal array during hydro-fracturing within a producing oilfield in Wyoming, USA. Using this test we are able to provide estimates of the smallest event detectable with the filtered data, and compare the results with conventional techniques, such as stacking. The second test case is a dense, small-aperture 2D seismic array of 95 stations placed within an area of 130 m×56 m on a landslide deposit in the Northern Apennines, Italy. Numerous micro-earthquakes have been recorded with this array, whose faint P-phases serve as an ideal dataset for testing filtering techniques.

Using the two case studies, we discuss the effectiveness of the multi-channel Wiener filter on SNR improvement, and show that including horizontal components into the analysis increases the SNR improvement more than using only vertical components.

Theory

The basic principle is to use the noise on a number of reference traces to predict the noise on the primary channel, and then to subtract the predicted noise from the actual data. The transfer functions between the reference and primary channels are estimated by data-adaptive multi-channel filters in the frequency domain, similar to the approach taken by Özbek (2000), whose algorithm operated in the time domain. Specifically, we seek to minimize the difference between the predicted and actual data of the primary channel in a least-square sense, i.e.,

$$min \sum_{k=1}^{N} |T_{k,i} A_k - A_i|^2 \quad \text{k}\neq\text{i} \tag{1}$$

where N is the trace number, $T_{k,i}(k = 1 \ldots N(k \neq i))$ is the transfer function between the primary channel, A_i, and reference channels, A_k, and all quantities are understood to be functions of frequency. The solution to equation (1) is

$$\begin{pmatrix} A_1A_1^* & \cdots & A_{i-1}A_1^* & A_{i+1}A_1^* & \cdots & A_NA_1^* \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ A_1A_{i-1}^* & \cdots & A_{i-1}A_{i-1}^* & A_{i+1}A_{i-1}^* & \cdots & A_NA_{i-1}^* \\ A_1A_{i+1}^* & \cdots & A_{i-1}A_{i+1}^* & A_{i+1}A_{i+1}^* & \cdots & A_NA_{i+1}^* \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ A_1A_N^* & \cdots & A_{i-1}A_N^* & A_{i+1}A_N^* & \cdots & A_NA_N^* \end{pmatrix} \cdot \begin{pmatrix} T_{1,i} \\ \vdots \\ T_{i-1,i} \\ T_{i+1,i} \\ \vdots \\ T_{N,i} \end{pmatrix} = \begin{pmatrix} A_iA_1^* \\ \vdots \\ A_iA_{i-1}^* \\ A_iA_{i+1}^* \\ \vdots \\ A_iA_N^* \end{pmatrix} \quad (2)$$

where normal matrix elements AA^* are the cross-correlations of the complex spectra averaged over multiple windows, with each window tapered by a Bartlett window. The vector on the right-hand side is made up of the cross-correlations between the primary and reference channels. The filtered output trace of the primary channel becomes

$$A_i' = A_i - \sum_{k=1}^{N} T_{k,i} A_k \quad \text{k}\neq\text{i} \quad (3)$$

i.e., the difference between the predicted and the actual data, where the transfer function $T_{k,i}$ is estimated according to (2). This filtering process can be repeated for all traces by making each channel a primary channel in turn. Finally, the filtered traces are stacked. The strength of the filter is controlled by the window length and the number of windows.

In addition, it is generally not possible to separate signal and noise *a priori*. To avoid tuning the transfer functions to the signal (rather than the noise), we calculate the transfer function based on the noise sequence before the expected signal arrival time. In detection mode, the transfer function would be updated in a rolling manner.

The implicit assumption of this filtering technique is that the effective wavenumber of the signal is different from the wavenumber of the noise, since otherwise the suppression of the noise will also suppress the signal. This assumption is likely to be reasonable in selected frequency bands as the noise is usually dominated by ground roll, whereas the desired signal will be a near-vertical arriving body wave with almost simultaneous arrival at all stations (or at least the arrival can be made simultaneous by appropriate time-shifting). However, when the wavelength of the ground roll is large compared to the

array dimensions at low frequencies, this technique will fail, and some signal suppression has to be accepted. A similar situation arises for quasi-linear arrays when the dominant propagation direction is perpendicular to the alignment of the array.

Furthermore, the potential effectiveness of an array for noise suppression is greatly dependent on the spatial distribution of the interfering noise (Backus et al., 1964). In other words, the coherency of the noise controls the predictability from the reference channels, which determine how much coherent noise will be suppressed. If seismic noise were completely random, then the output of the primary channel would be completely unpredictable. In that case, the transfer function has no effect for filtering, and the SNR improvement only results from stacking.

However, provided the noise is coherent, i.e., it originates from a small number of effective sources, it does not need to conform to any particular simple model: for example, it is not necessary to assume plane-wave propagation. Theoretically, even multiply scattered noise can be reduced effectively, if it originates from a well-defined and stationary source area. Likewise, strong site effects, with which many other array-based algorithms cannot cope effectively, are not expected to adversely affect the noise filtering, except in the final stacking step. Finally, no assumption is made about the array configuration, in particular the array does not need to be regularly spaced or circular, although—as we have just seen—particular configurations can be more or less effective.

Operating in the frequency domain rather than in the time domain has two advantages: 1. it is faster for computations with many channels due to the efficiency of Fast Fourier Transforms; 2. the solution is stablized by averaging over multiple windows and the strength of the filtering is controlled by the window number, unlike in the time domain approach (Özbek, 2000), where the strength of the filter is achieved by limiting the number of principal components of the singular value decomposition used in the solution, a parameter which needs to be chosen for each noise reduction problem. However, the necessity to average over multiple windows requires a longer time period for the determination of the transfer functions, making this implementation less suitable for rapidly changing noise environments (such as traffic moving in the vicinity of the array).

The algorithm does not specify what type of data the different channels contain. Initially, we use vertical data only, but later expand our approach to combine the horizontal and vertical data.

Application of the Wiener filter

Semi-synthetic example: oilfield noise environment

A passive surface seismic monitoring array, composed of 10 three-component Güralp 6TD seismometers (0.03 - 100 Hz), deployed in a hexagonal array, and 5 high-frequency (4.5 - 1000 Hz) seismometers, was installed in Wyoming during hydro-fracturing (Figure 1). No clear phases of microseismic events are visible because of the strong background noise generated by the ongoing production and drilling activity, as well as by pumping during hydro-fracturing. We take a 20-s-long sample of data acquired during pumping as an example to study how much the noise would have to be suppressed to allow detection of seismic hydro-fracture induced events. Based on assumptions about the likely depth of events and the average attenuation structure, we estimate that only events with $M \geq\sim 0$ would be marginally visible on a single sensor. From downhole data, we know that the largest hydro-fracture induced event had a magnitude of -1.8. It would thus require 36 dB ($20 \log_{10}(\frac{A_{M=0}}{A_{M=-1.8}})$) improvement for a single sensor to detect an event. Simple stacking would produce 10 dB ($10 \log_{10} N$) improvement for N = 9 channels (only the Güralp 6TD sensors, bad channels excluded) if the noise were completely incoherent. However, actual stacking only improved SNR by 3-6 dB, due to partially constructive noise interference. When the noise is coherent and distinct from the signal, we can do better.

Apparent velocity measurements using slant-stack techniques suggest that the coherent noise is dominated by sources from certain directions. Thus, the coherency of the noise allows us to apply the frequency-dependent multi-channel Wiener filter technique. The detection threshold of the microseismicity is obtained by testing the Wiener filter on the semi-synthetic dataset.

We produced a semi-synthetic dataset of five cases by adding five signals (spikes band-

passed in 1-30 Hz) with different amplitudes to nine 20-s-long raw noise vertical-component data samples. We increased the amplitude ratio between the spike and noise (root mean square value) of $A_S/A_N = 1, 2, 3, 4, 5$, from top to bottom (Figure 2). A Butterworth filter, with limits 1 - 30 Hz, order = 3, is applied to the Wiener filtered outputs. The window length is 0.5 seconds, with 50% overlap between windows. i.e., 39 windows of the first half of the noise data (0-10 seconds) are used as noise references to generate the transfer functions, which are then averaged and applied to the second half of the semi-synthetic data (10-20 seconds). The filtered waveforms for the different A_S/A_N are shown in Figure 3, in which stacking and a single filtered trace WS05 before final stacking are shown for comparison. In Figure 3(d), the bottom three traces clearly show the signals with stronger amplitudes, but the fourth one, with $A_S/A_N = 2$, is only marginally visible. The top trace, where the SNR is even lower, is beyond the performance of the Wiener filter. The multi-channel Wiener filter is therefore expected to reduce the detection threshold by approximately 0.6 magnitude units. This is still short of what would be required to detect the downhole events in this particular experiment, but shows the efficiency of the method.

As we discussed before, the SNR improvement depends highly on the coherency of the noise. Since the coherency of the noise varies with frequency, it is necessary to obtain a precise measure of the SNR improvement capability in each frequency component using power spectral density (PSD) curves. Figure 4 shows that the Wiener filter works better at lower frequencies due to the array configuration, which results in good coherency below 10 Hz. The noise is reduced by up to $\sim$14 dB at 1-6 Hz, but the SNR is not dramatically improved because the signal is also reduced by 3-4 dB, resulting in an overall SNR improvement of $\sim$10 dB. Stacking only reduced the noise by 3-7 dB (less than expected due to the fact that the noise is coherent and stacks partly constructively).

Data example: Northern Italy

An array of 95 three-component Güralp 6TD seismometers was deployed on a landslide near the village of Cavola in the Northern Apennines, Italy, where intrinsic weakness of

the basement rocks causes a propensity for landslide initiation and reactivation (Bordoni and the CAVOLA Experiment Team, 2005). Station spacing in the grid was 10 m and 8 m, and the overall size of the array was 130 m×56 m (Figure 5(a)). In spite of the small station spacing, the noise on the different sensors shows marked variation dependent on the thickness of the landslide underneath each sensor. The multi-channel Wiener filter is thus tested in an environment where parametric approaches (e.g. plane wave) are difficult to apply. The dense array provides a flexible test-bed as different subsets of sensors can be used to probe the effectiveness of different array configurations. A local earthquake (epicentral distance 15.6 km, depth 29.8 km) with ambiguous P-phase arrivals was chosen as a test event for filtering. Raw data on vertical components of Line B is shown in Figure 5(b).

The P-phase visibility is significantly improved (Figure 6). The noise is suppressed by up to 22 dB over the optimally performing frequency band 8-15 Hz. The SNR is improved up to ∼20 dB, because of a small ∼2 dB signal suppression after the multi-channel Wiener filter has been applied. Using 7 traces for the Wiener filter generates much better results below 8 Hz than simply stacking 37 traces. Filtering using 7 traces is enough to reach the maximum noise attenuation (below 2 Hz), more traces do not improve the results. The advantage of using more stations for filtering becomes obvious with increasing frequencies, and using 37 traces for filtering is most effective at 8-15 Hz.

Wiener filter on three-component data

In this section, we apply the multi-channel Wiener filter to three-component data using the same window length for both the Wyoming and Cavola data. Cross-coupling noise between horizontal and vertical channels signifies that the two horizontal components can be taken as reference channels to reduce the noise on the vertical channel using the Wiener filter mechanism (Dahm et al., 2006). This can be done as a pre-processing step before applying the multi-channel filter on all of the individually filtered vertical channels. Assume there are N stations of three-component data, horizontal components (H1, H2) are taken as reference channels to reduce the noise on the vertical channel (V). There are three methods to make use of horizontal components in the pre-processing step, in which

channel V is treated separately first:

- Method I: H1 and H2 to filter V
- Method II: N*H1, N*H2 and (N-1)*V to filter V
- Method III: N*H1 and N*H2 to filter V

Semi-synthetic three-component tests using Wyoming data

For the three-component test using the Wyoming data, the actual horizontal noise data are left free of any synthetic signal; by restricting the synthetic signals to the vertical we are effectively modeling a vertically-propagating P wave which is approximately equivalent to the hydro-fracture monitoring scenario. We then apply the multi-channel Wiener filter using the same parameters as in previous tests. The filtered waveforms for the fourth spike test ($A_S/A_N = 2$) are shown in Figure 7(a). The three methods generate similar results.

It is discovered that the power spectral density of the results from the three methods are almost same below 10 Hz and method III is slightly more effective above 10 Hz. Hence, in Figure 7(b), we compare the filtered outputs from method III to the results of the vertical-only filtering. It is obvious that three-component data performs better in general comparing with Figure 4, especially above 10 Hz. Although the noise filtered by the three-component dataset is suppressed by an additional 3-7 dB compared to the vertical-only test, the filtered signal is preserved at the same level. This indicates that using horizontal components as reference channels preserves the signal better than using neighbouring vertical components. We conclude that three-component data enlarges the effective frequency range for filtering up to 20 Hz with the overall SNR improvement of 11-13 dB and preserves the target signals much better.

Real three-component tests using Cavola data

In this part, we use both Line B (7 three-component sensors) and Lines A-F (37 three-component sensors, bad channels are excluded) (see Figure 5) as test cases to apply the three methods outlined above.

In Figure 8(a), the results of vertical-only filtering are compared with those of three-component filtering using the three methods. All three methods using three-component data improve the SNR by an additional 6-8 dB (see Figure 8(b)). Method I using 37 three-components is no more effective than using only 37 vertical components. This is an indication that the predictability of the horizontal reference channels has been saturated, and no more SNR improvement will be achieved, even with more reference channels. So we also consider the SNR improvement with respect to the number of stations. We find that using three-component data is generally better than using three times the number of vertical components. For example, 7 three-component data channels generate better filtered results than using even 37 vertical component data channels. In particular, using horizontal channels produces better suppression of low-frequency noise than using neighbouring vertical channels. Additional simulations showed that the number of vertical channels reaches the maximum ability to improve SNR at about 18 stations, whilst there is no benefit at all from using horizontal components from more than 8 stations (equivalent to 24 channels). This means that three-component data are more suitable for the Wiener filter than only vertical components. Method II is the most robust processing route, which suppresses noise up to 30 dB at 1-15 Hz, and the overall SNR improvement reaches to 19 dB at 1-10 Hz. Method II appears to outperform Method III in this case, maybe because operating with all the traces simultaneously allows the algorithm to detect subtle correlations that are missed if subsets of traces are treated in a multi-step process as in the other two methods.

Conclusions

A multi-channel Wiener filter has been implemented in the frequency domain and its effectiveness is evaluated in the two different noise environments of an oilfield in Wyoming,

USA, and on a landslide in a rural environment in Italy. By making semi-synthetic tests using the noise data recorded in the oilfield, we show that the SNR in the frequency band 1-6 Hz can be improved by up to ∼10 dB using a 9-element array of vertical traces. An overall SNR improvement of 11-13 dB can be achieved in the frequency band 1-20 Hz using three-component data. The denser array in Italy shows the effectiveness of the multi-channel Wiener filter on a real dataset, which improves the SNR by ∼14 dB using 7 vertical traces, and by up to 20 dB using 7 three-component stations. However, in this experiment, the effectiveness of using three-components data does not improve when using 8 stations or more.

Acknowledgments

We thank Schlumberger Cambridge Research for providing funding for this project and for the hydro-fracture surface monitoring experiment. However, the views expressed here are those of the authors, who are solely responsible for any errors. For the hydro-fracture surface data we thank the Schlumberger office in Rock Springs, Wyoming, for help with logistics and deployment, BP for permission to deploy seismometers on one of their fields, Anna Horleston and Sharif Aboelnaga for assistance in the field, and SEIS-UK for the loan of the seismometers. The Cavola data were acquired by the Istituto Nazionale di Geofisica e Vulcanologia, Italy. We thank the Cavola Experiment Team, particularly John Haines (University of Cambridge), Giuliano Milana, Giuseppe Di Giulio and Fabrizio Cara (Istituto Nazionale di Geofisica e Vulcanologia). Ed Kragh and Everhard Muyzert provided helpful advice. Dept. Earth Sciences contribution No. ES9031.

References

Backus, M. M., Burg, J. P., Baldwin, D., and Bryan, E. (1964). Wide-band extraction of mantle p waves from ambient noise. *Geophysics*, 5:672–692.

Bordoni, P. and the CAVOLA Experiment Team (2005). Cavola experiment: a dense broadband seismic array on an active landslide. *Geophysical Research Abstracts*, 7:EGU05–A–07726.

Dahm, T., Tilmann, F., and Morgan, J. P. (2006). Seismic broadband ocean-bottom data and noise observed with free-fall stations: experiences from long-term deployments in the North Atlantic and the Tyrrhenian Sea. *Bull. Seism. Soc. Am.*, 96(2):647–664.

Özbek, A. (2000). Multichannel adaptive interference canceling. *SEG Technical Program Expanded Abstracts*, pages 2088–2091.

Rost, S. and Thomas, C. (2002). Array seismology: methods and applications. *Rev. Geophys.*, 40(3):1008.

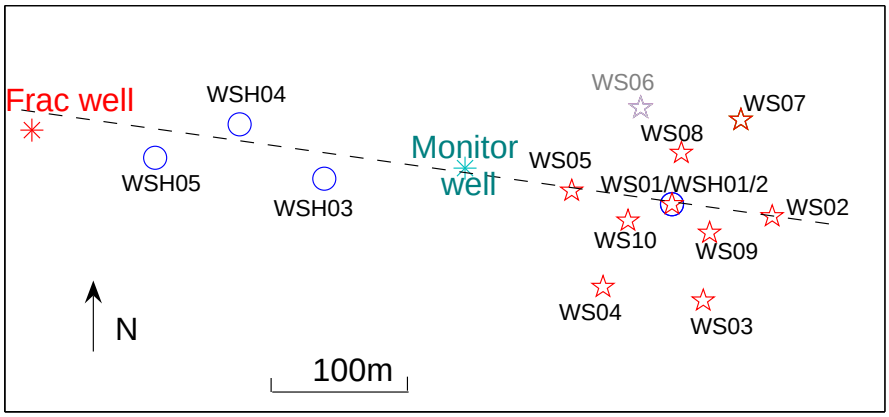

Figure 1: The station layout for the Wyoming array. The array consists of ten Güralp 6TD seismometers (stars) in a hexagon and five high-frequency seismometers (circles). Two high-frequency seismometers and one 6TD are co-located at the centre of the hexagon. The dashed line shows the central line of the network. Only the 6TD data are used in this paper. Seismometer WS06 (gray) was faulty and did not produce any data.

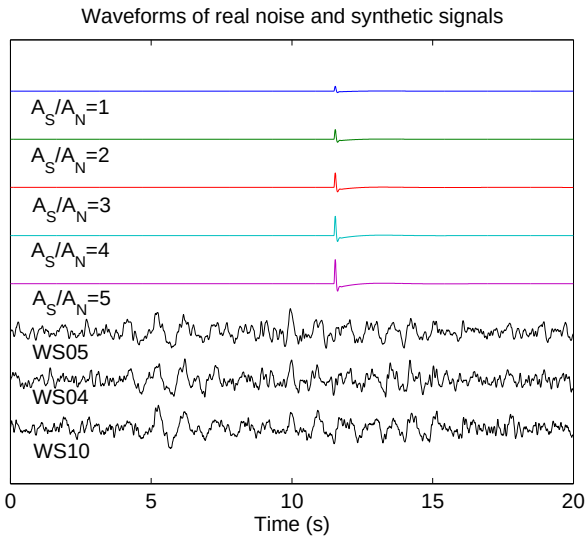

Figure 2: The bottom three traces are examples of band-passed (1 - 30 Hz) vertical component noise data at the stations WS05, WS04 and WS10. The noise is highly coherent between stations. The five spikes are synthetic signals added (around 12 s) to the noise for the semi-synthetic test, at the indicated A_S/A_N.

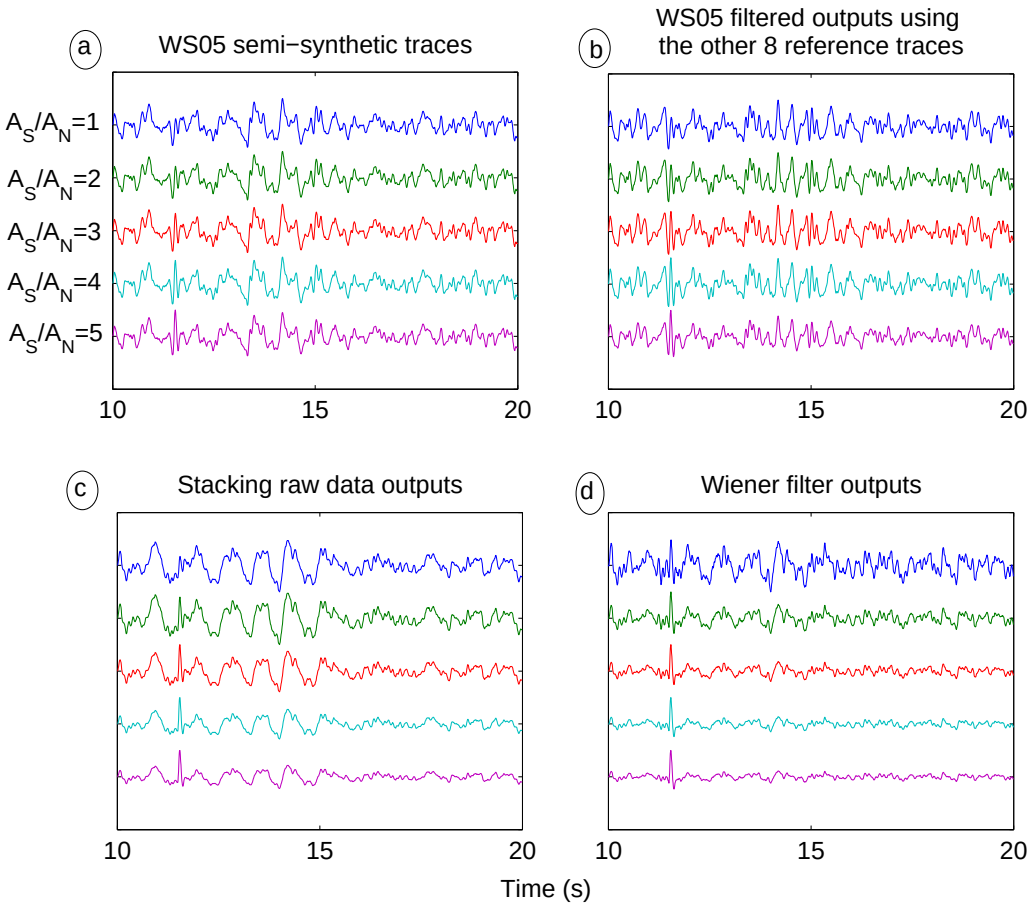

Figure 3: The normalized waveforms from five synthetic data examples with increasing A_S/A_N from top to bottom. (a) WS05 (noise plus signals) with varying strengths, examples of semi-synthetic traces. (b) The filtered trace WS05 before final stacking. (c) Conventional stack of all nine vertical channels. (d) The Wiener filtered results of all nine vertical channels.

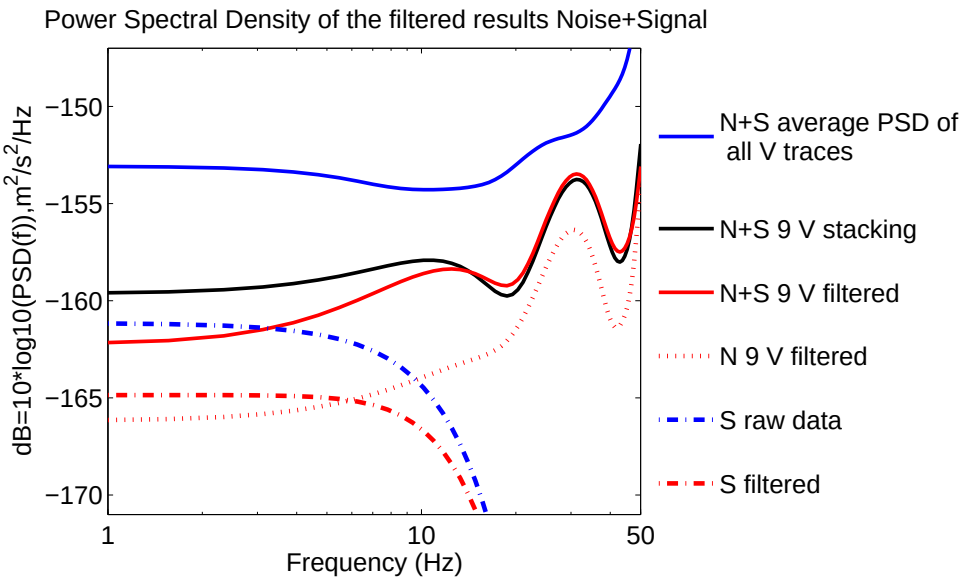

Figure 4: Comparison of the spectra of the Wiener-filtered output results from the spike test ($A_S/A_N = 2$). The transfer functions generated from noise data are applied on the raw noise data only, and the transfer functions generated from semi-synthetic data are applied on the synthetic signal only. In the labels of this and the following figures, we use "N" to represent *Noise*, "S" for *Signal* and "N+S" for *Noise+Signal*, "V" for vertical component filtering.

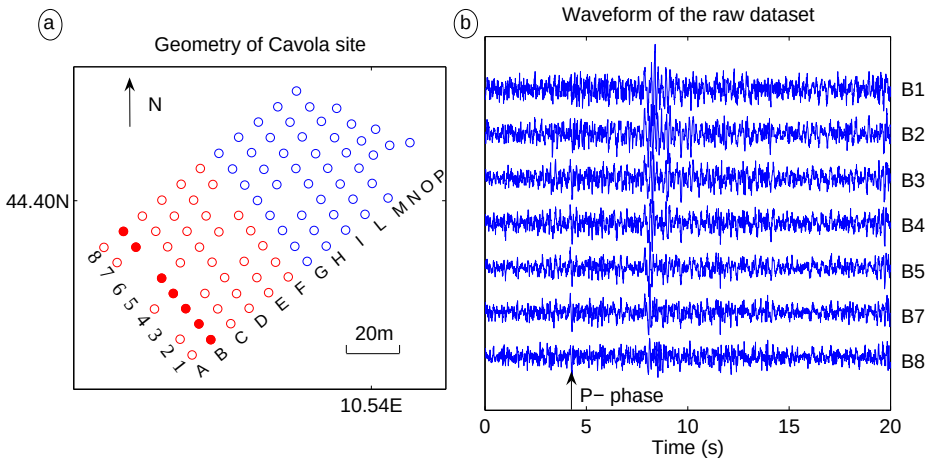

Figure 5: (a) The layout of the Cavola seismic array. Traces from Line B (7 sites with red filled circles) and Lines A-F (37 sites with all red symbols, bad channels are excluded) are used for filtering. (b) Examples of waveforms of Line B (red filled circles, vertical component), where the onset of the P- phase is ambiguous.

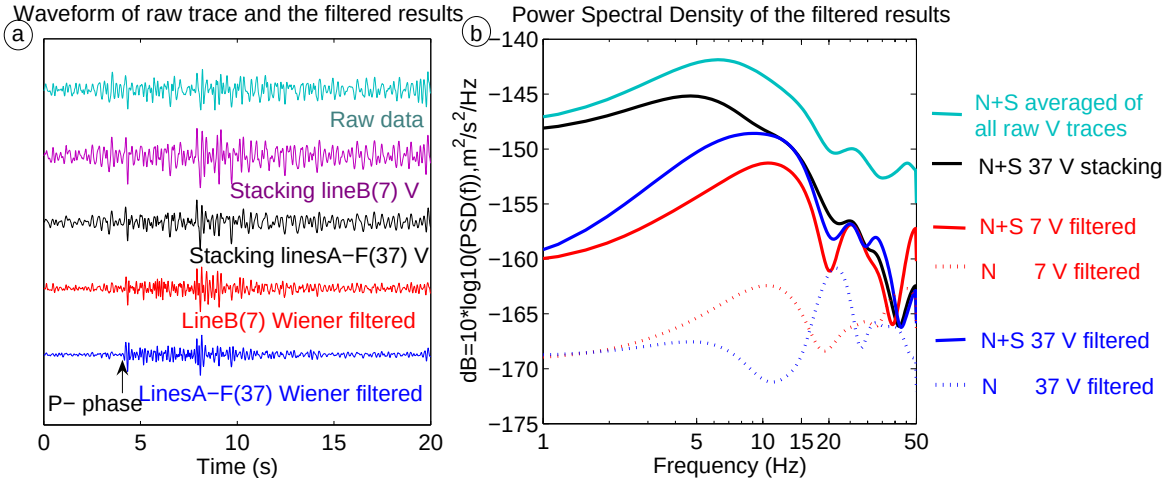

Figure 6: (a) The onset of the P- phase is obvious after filtering. The raw trace (station B2 in Figure5(b)) on top acts as an amplitude reference. The purple and black traces are results of stacking 7 and 37 traces, respectively. The red and blue traces are filtered waveforms from 7 seismometers (Line B) and 37 seismometers (Lines A-F), respectively (see Figure 5). (b) Power spectral density plots for evaluating SNR improvement.

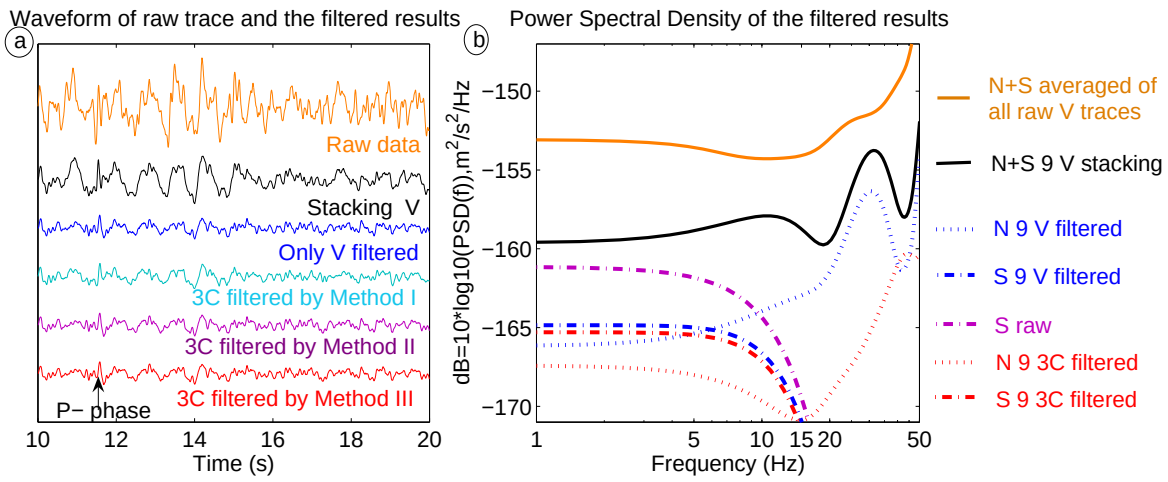

Figure 7: (a) Comparison of the filtered waveforms for the three methods using a semi-synthetic signal with ($A_S/A_N = 2$). In the labels of this and the following figures, we use "3C" for three-component filtering. The three methods generate similar results. (b) Power spectral density plot of semi-synthetic test results. The filtered effects of using only vertical component and three-component (method III) are compared by showing the results of only "N"oise and only "S"ignal.

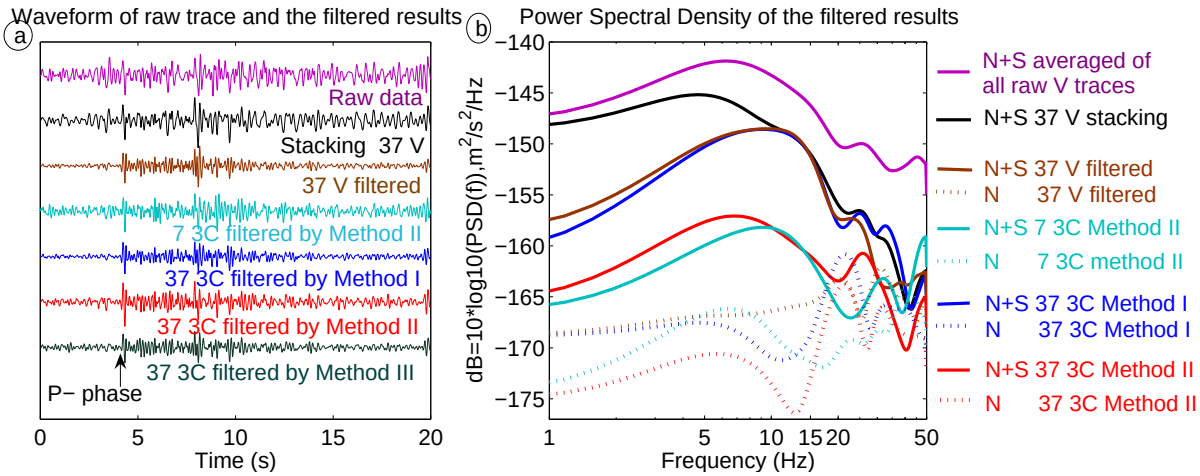

Figure 8: (a) Comparison of the filtered waveforms in various approaches, with the raw trace (top) as the reference. (b) Power spectral density plot of the filtered results.