Seismic Trace Characterization for Shear Wave Velocity Assessment

Introduction

BCE has invested considerable resources into the development of techniques and algorithms to characterize acquired Downhole Seismic Testing (DST) data sets for shear wave velocity assessment, and to use that characterization as a guide for the processing of those data sets to calculate interval velocities.

The characterization is based on various independent parameters of the acquired DST data at a particular depth. Currently five parameters are considered:

- Parameter 1: the linearity estimates (LIN) from polarization analysis. The LIN trace metric quantifies the correlation between X, Y and Z axis responses.
- Parameter 2: the Cross Correlation Coefficient (CCC) of the full waveforms at the particular depth and the preceding depth. The CCC trace metric gives an indication of the similarity between the two waves being correlated when deriving relative arrival times.
- Parameter 3: the Signal Shape Parameter (SSP). The SSP trace metric quantifies the deviation of the shape of the frequency spectrum from an ideal bell shape
- Parameter 4: the Peak Symmetry Differential (PSD) trace metric facilitates the identification of traces whose peak source wave responses have been significantly skewed due to measurement noise or source wave reflection interference.
- Parameter 5: Signal to Noise Ratio (SNR). The SNR trace metric is solely provided to quantify what portion of the spectral content of the recorded seismogram resides within the desired source frequency spectrum irrespective of source wave distortions such as near-field effects, reflections, refractions, and "dirty sources".

For these five parameters the part of the data sets that will be used and filtered is shown in the table below:

Parameter	Part of trace that is reviewed	Applied Signal Filtering
LIN	Largest peak/trough ± 30 ms	200 Hz low pass
CCC	Largest peak/trough -30ms for upper trace Largest peak/trough +30ms for lower trace	200 Hz low pass
SSP	Entire trace	200 Hz low pass
PSD	Largest peak/trough ± 2 crossovers	200 Hz low pass
SNR	Largest peak/trough ± 2 crossovers	None / 200 Hz low pass

This Technical Note introduces the parameters, while Technical Note 21 provides a guide for the recommend post data processing and seismic signal processing based upon these parameters.

It should be noted that the characterization process may be updated as more data set are reviewed.

Parameter 1: LIN

The linearity or rectilinearity values can be obtained from hodograms, i.e. by plotting the X, Y and Z axis amplitudes against one another and fitting least squares best fit lines. Since hodograms with linearity values nearing 1.0 identify seismic recordings that have highly correlated responses on the X, Y and Z axes and strong directionality, the interval velocities calculated from such recordings are likely to be accurate. Hodograms with lower linearity values on the other hand indicate lower signal-to-noise ratios or SNRs (whether due to poor source generation, near-field waves, ambient noise that is not easily filtered out or source wave reflections) and thus the resulting interval velocity values are likely to be less accurate.

To provide for a more accurate quality assessment of the recorded data minimal digital frequency filtering should be applied to the raw data, while more refined and aggressive digital frequency can be applied during the actual data analysis to determine the interval velocities.

The X, Y and Z responses in Figure 1 below from a SH source (after applying a low pass frequency filter of 200 Hz) have clearly high SNRs values: the peaks and troughs on the X and Y axis line up, and there are minimal recordings on the Z axis as would be expected for this kind of source.

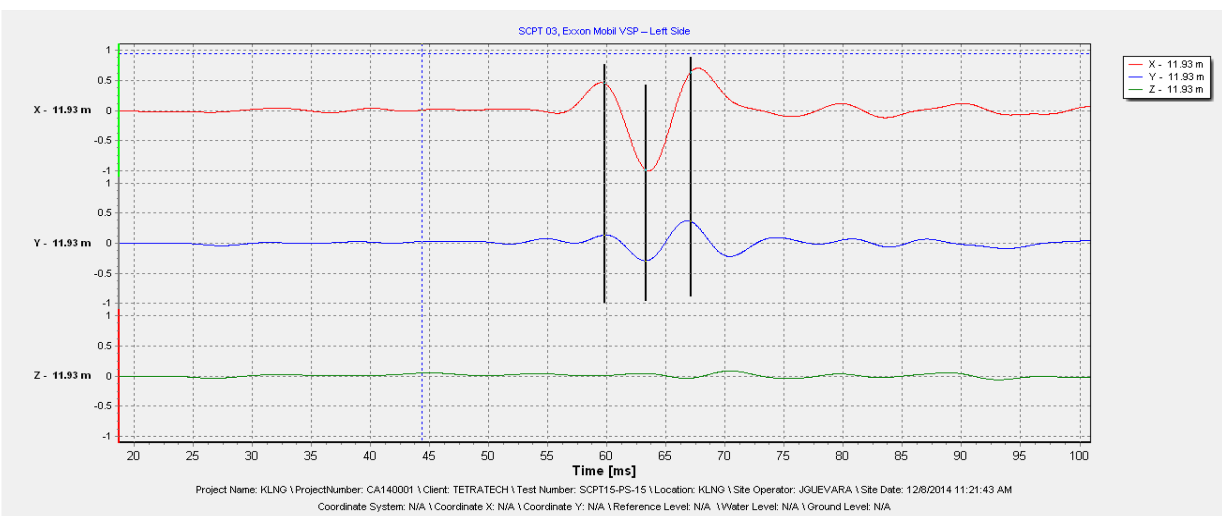

Figure 1: Correlated triaxial responses resulting in high linearity values. The source wave X and Y axis peaks and troughs are aligned

The hodogram plot for this triaxial recording is shown in Figure 2 below. In this hodogram the amplitudes of the X and Y axes recordings (dominant energy for a SH source) are plotted as green circles and the red line is the best fit straight line (with a calculated linearity of 0.92). This clearly reflects a good quality seismic source recording with a high correlation between the X and Y axis and high directionality along an axis with an azimuth of approximately 13º.

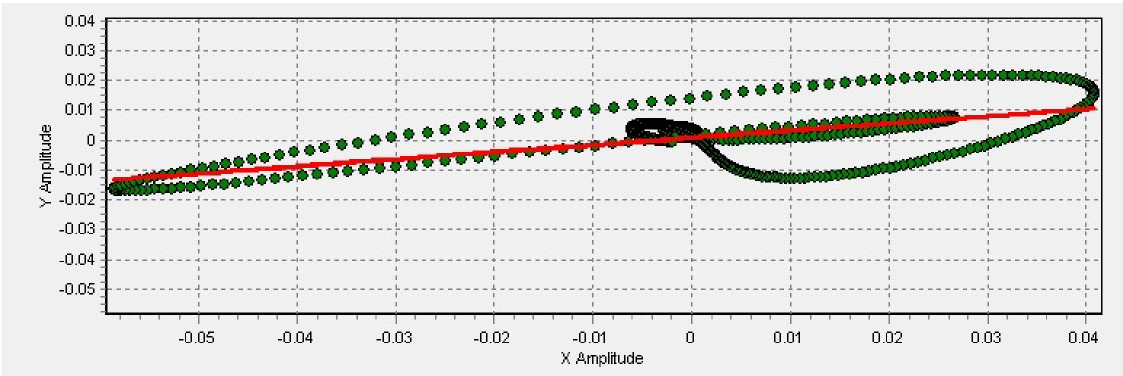

Figure 2: The hodogram plot for the triaxial responses illustrated in Fig. 1. The hodogram clearly reflects a good quality seismic source recording with a high correlation between the X and Y axis and high directionality along an axis with an azimuth of approximately 13º.

In the triaxial seismic trace recording in Figure 3 the peaks and troughs on the X, Y and Z do not line-up and the background noise (whether due to a poor source, vibrations within the testing vehicle upon impact of the SH wave sledge hammer and/or other causes) has frequency components similar to the source wave making the isolation of the source wave with frequency filters a challenging tasks.

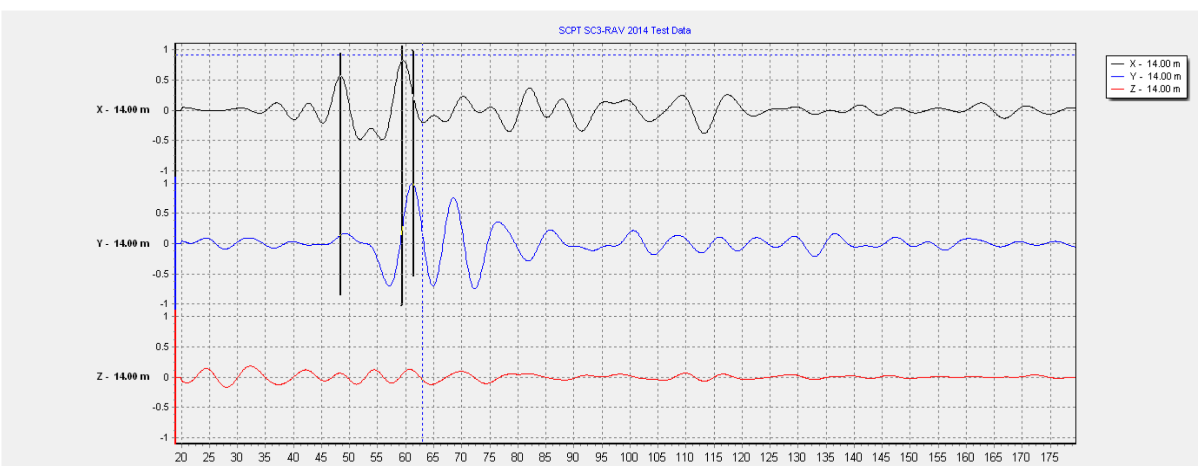

Figure 3: Poorly correlated triaxial responses resulting in low linearity values. The source wave X and Y axis peaks and troughs are not aligned.

The hodogram plot for this recording is shown in Figure 4 below. The amplitudes of the X and Y axes recordings are plotted (green circles) and the red line is again the best fit straight line, but obviously with a much lower calculated linearity than in the first example (0.61 vs. 0.92).

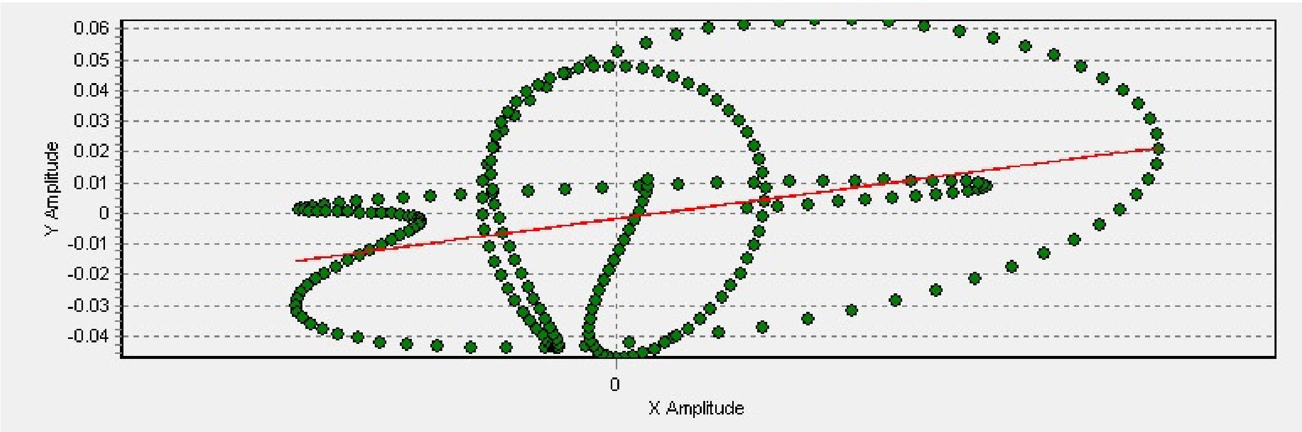

Figure 4: The hodogram plot for the triaxial responses illustrated in Fig. 3. The hodogram clearly reflects a poor quality seismic source recording with a low correlation between the X and Y axis and low directionality.

Parameter 2: CCC

The cross-correlation between two time or distance offset seismograms is given as (Gelb 1974)

$$\varphi_{xy}(\tau) = \sum_k X_k \, Y_{k+\tau} \tag{1}$$

where $\varphi_{xy}(\tau)$ is the cross-correlation function, Y_k the sampled data at distance 1 and at sample time k, X_k the sampled data at distance 2 at sample time k, and τ the time shift between the two sets of recorded waves (note: distance 2 > distance 1). The value of the time shift at the maximum cross-correlation value is assumed to be the relative travel time difference, Δt, for the source wave to travel the distance increment. This technique has several advantages over selecting time markers within the seismogram (Baziw 1993, 2002), among others the human bias associated with visually selecting a reference point or time marker is minimized.

Normalizing the cross-correlation of the zero mean seismic signals by their standard deviations gives the cross-correlation coefficient:

$$\rho_{xy}(\tau) = \sum_k X_k \, Y_{k+\tau} \Big/ \sqrt{\sum_k X_k^2} \sqrt{\sum_k Y_K^2} \tag{2}$$

The CCC between the two DST waves is typically used to assess the quality of the interval velocity estimate as this parameter gives an indication of the similarity between the two waves being correlated. While on its own the CCC has proven to be an unreliable indicator of the overall quality of a seismic trace (since it is highly dependent on the digital filter applied to the raw seismic signals), it is still a useful component of seismic trace characterization. As an STC parameter the CCC value is calculated on the full waveforms after applying polarization analysis.

Parameter 3: SSP

The probability density of a normal (or Gaussian) distribution is given as

$$f(x|\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad (3)$$

where μ denotes the mean or expectation of the distribution and σ denotes the standard deviation with variance σ^2. The area under the normal pdf curve is unity. Figure 5 illustrates example of normal pdfs for varying μ and σ^2 values. All the curves in Fig. 1 have the classical bell-shape.

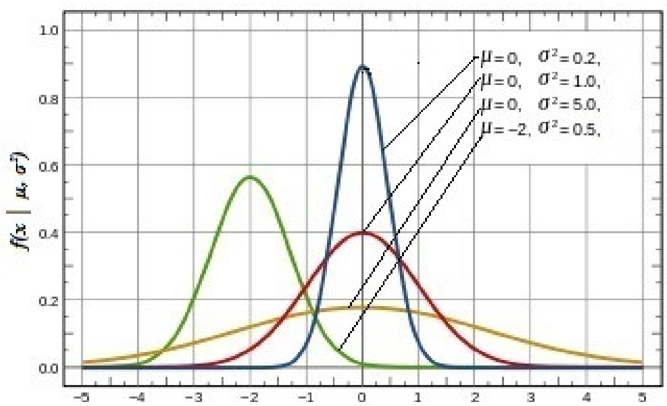

Figure 5: Example of normal pdfs for varying μ and σ2 values. (after, http://www.dplot.com/probability-scale.htm

Figure 6 illustrates a Berlage source wave (Baziw and Ulrych (2006), Baziw and Verbeek (2014)), which is commonly used within seismic signal processing for simulation purposes. The Berlage source wave is analytically defined as

$$w(t) = AH(t)t^n e^{-ht}\cos(2\pi f t + \emptyset) \qquad (4)$$

where *H(t)* is the Heaviside unit step function [$H(t) = 0$ for $t \le 0$ and $H(t) = 1$ for $t > 0$]. The amplitude modulation component is controlled by two factors: the exponential decay term *h* and the time exponent *n*. These parameters are considered to be nonnegative real constants. Figure 7 illustrates the frequency spectrum (solid black line) of the Berlage source wave shown in Fig. 6 with the normal pdf approximation shown as a dotted grey line, with μ = 69 Hz and σ = 32.5 . As is evident from Figure 7, the frequency spectrum of the simulated Berlage source wave closely matches that of a bell-shaped curve.

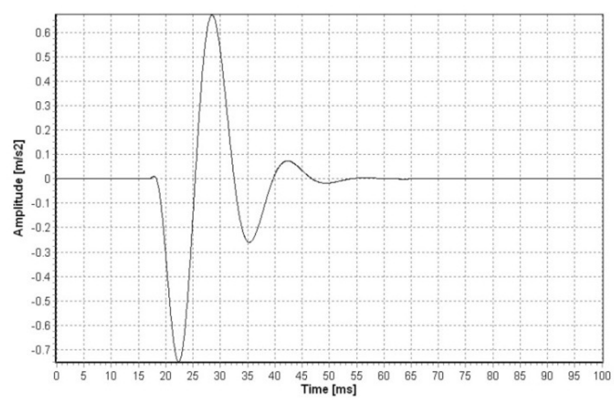

Figure 6: Berlage source wave with of f = 70 Hz, n = 2, h = 270 and ϕ = 40º specified.

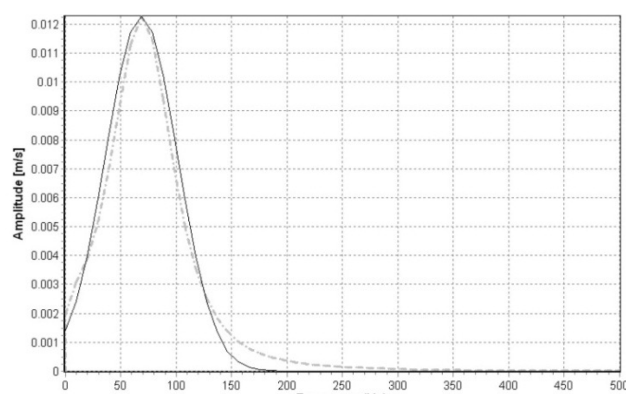

Figure 7: Frequency spectrum (solid black line) of Berlage source wave illustrated in Fig. 2 with the normal pdf approximation shown as a dotted grey line.

To determine the deviation of the source wave frequency spectrum from a desirable bell-shaped curve the following process is proposed:

1. Apply a digital zero-phase shift frequency filter to the entire seismic trace so that high frequency measurement noise is removed.
2. Calculate frequency spectrums for *X(t)* and *Y(t)* recordings, $S_X(f)$ and $S_Y(f)$, and determine which axis has the dominant frequency response axis (denote as *S(f)*).
3. Force the area under *S(f)* to approach unity by uniformly modifying the amplitudes within *S(f)*. This step is outlined below by eqs. 5(a) and 5(b).

$$Area_{S(f)} = \Delta f \sum_{i+1}^{n} S(f)_i \tag{5a}$$

$$\sum_{i+1}^{n} S(f)_i = S(f)_i / Area_{s(f)} \tag{5b}$$

   In eq. 5(a), Δf denotes the frequency increment resolution.
4. Determine μ (dominant frequency), $p(\mu)$ (maximum spectral amplitude), and $\sigma = 1/(p(\mu) \sqrt{(2\pi)})$ utilizing an iterative forward modelling (IFM) technique such as the Simplex method (Baziw, 2002, 2011). In this IFM case the cost function to minimize is the RMS difference between the normalized area under *S(f)* and the derived area (using eq. (3)) from a normal pdf which utilizes the currently estimated μ and σ values.
5. Calculate *p(f)* via equation (1) utilizing the IFM estimates μ and σ from Step 4.
6. Calculate $\epsilon_1 = \sum_{i+1}^{n} abs\big(S(f)_i - pdf(f)_i)\big)$
7. Calculate $\epsilon_2 = \sum_{i+1}^{n} abs(S(f)_i)$
8. Calculate parameter R which is defined as $R = \varepsilon_1/\varepsilon_2$
9. Signal Shape Parameter (*SSP*) is then calculated as *SSP = 1-R*.

Parameter 4: PSD

The "Peak Symmetry Differential" (PSD) parameter facilitates the identification of traces whose peak source wave responses have been significantly skewed due to measurement noise or source wave reflection interference. Fig. 8 illustrates this phenomenon. In Fig. 8(A) we have an ideal source wave recording where no interference is present. In this case the time difference between the two zero crossings bounding the peak response (A_1) are identical ($\Delta t_1 = \Delta t_2$). In Fig. 8(B) we have a source wave recording with interference, resulting in skewing or time shifting of the peak source wave response. The "peak symmetry" error assessment is also carried out on the adjacent peaks and/or troughs if the amplitude exceeds 70 % of that for the peak response

The PSD parameter is determined as follows:

1. Apply a frequency filter to the entire seismic trace to eliminate irrelevant zero crossings
2. Identify the largest peak/trough of the seismic trace and determine the time differential from the moment the peak occurs to the zero crossings on either side (Δt_1 and Δt_2 respectively)
3. Calculate $\Delta t = |\Delta t_1 - \Delta t_2|$.
4. If the amplitude of the adjacent trough/peak on either side exceeds 70 % of that for the largest peak/trough response calculate the Δt value for that trough/peak.
5. Determine the maximum Δt value

6. The PSD is then: for $\Delta t \le 0.02$ PSD = 1
for $0.02 \le \Delta t \le 0.8$ PSD = $1.026 - \Delta t/0.78$
for $\Delta t \ge 0.8$ PSD = 0

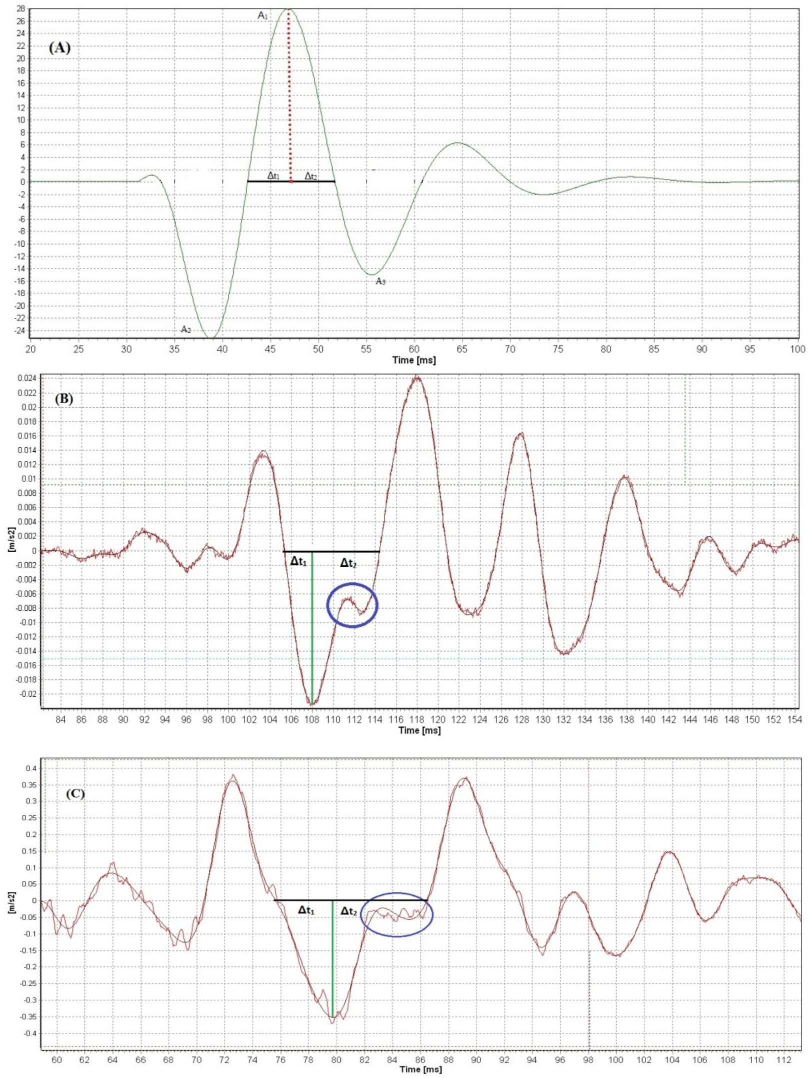

Figure 8: Source wave peak distortions due to measurement noise or source wave reflection interference. (A) Ideal source wave recording where no interference is present. (B) and (C) Source waves with interference resulting in skewing or time shifting of the peak

Parameter 5: SNR

The initial four parameters (LIN, CCC, SSP and PSD) are derived after filtering (albeit to a minimal extent) the seismic traces under analysis. The last parameter uses as input the as-recorded seismic trace and compares it with the filtered trace to assess the extent of background noise. To get a complete assessment it is essential that the recorded trace reflects accurately the signals that exist at the sensor location, in other words that the sensors display with minimal distortion the background noise and seismic source waves (see also BCE Technical Note 10). In the remainder of this Technical Note it is therefore assumed that sensors are used where the sensor output accurately represent the sensor input.

Figure 9 illustrates various seismic traces with varying signal-to-noise ratios. In each figure the filtered black traces (200 Hz low pass filter) are superimposed upon the corresponding virtually unfiltered seismic red traces (a 700 Hz low pass filter was applied to remove electrical noise). In Fig. 9(A) the unfiltered dominant source wave recording closely matches that of the filtered trace, which implies a high signal-to-noise ration. In Figs. 9(B) to (D) the unfiltered traces correlate a lot less with the filtered traces.

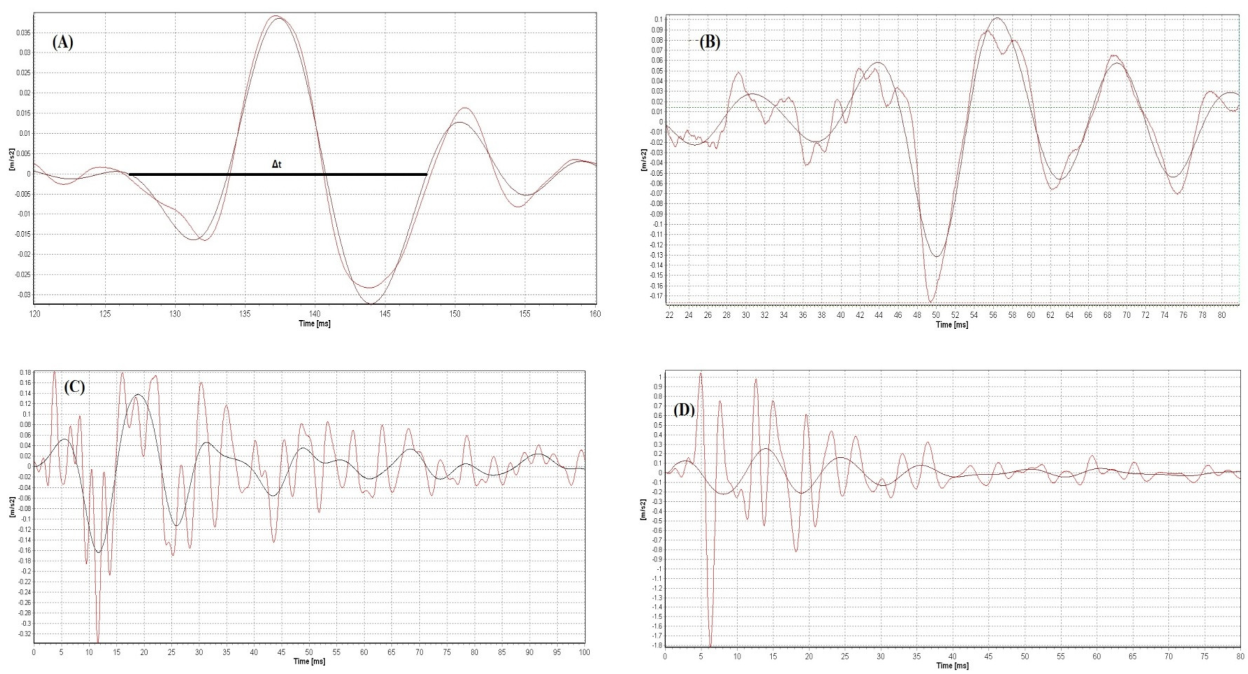

Figure 9. Examples of DST unfiltered (red traces) and corresponding filtered (black traces – low pass of 200 Hz applied) seismic time series.

To quantify this aspect the "Signal Noise Ratio" (SNR) parameter is used, which revolves around a comparison of the normalized peak response in the original seismic trace and that in the filtered trace:

The SNR parameter is determined as follows:

1. Identify the largest peak/trough of the filtered trace and derive an analysis time window by moving forward and backward in time (as shown by the solid black line in Figure 9(A)).
2. Normalize both the unfiltered and filtered responses within this time window.
3. Calculate the difference trace between the normalized and time windowed unfiltered and filtered traces. Figure 10 illustrates the difference time series for traces illustrated in Fig. 1 and implementing previously outlined Steps 1 and 2.
4. Calculate the standard deviation σ of the difference trace.
5. The SNR is then:

for $\sigma \leq 0.03$	SNR = 1
for $0.03 \leq \sigma \leq 0.7$	SNR = $1.045 - \sigma/0.67$
for $\sigma \geq 0.7$	SNR = 0

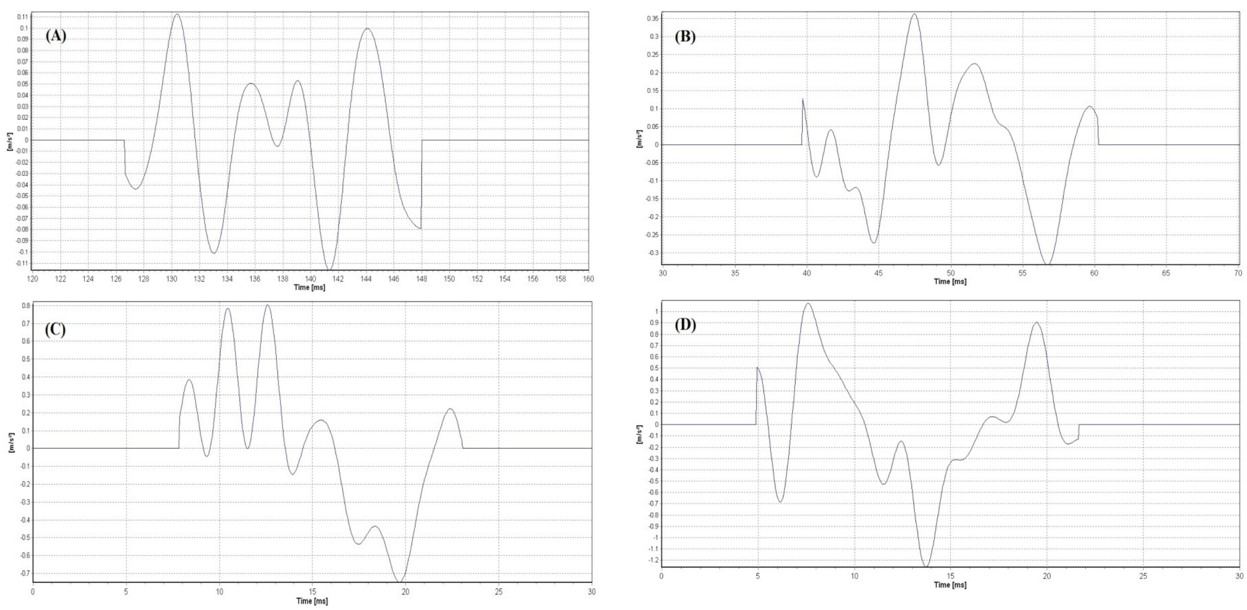

Figure 10. Calculated difference time series for traces illustrated in Fig. 9 and the SNR values (A) SNR = 0.95, (B) SNR = 0.79, (C) SNR = 0.52 and (D) SNR = 0.24.

Erick Baziw
Gerald Verbeek

BCE's mission is to provide our clients around the world with state-of-the-art seismic data acquisition and analysis systems, which allow for better and faster diagnostics of the sub-surface. Please visit our website (www.bcengineers.com) or contact our offices for additional information:

e-mail: info@bcengineers.com

phone: Canada: (604) 733 4995 – USA: (903) 216 5372