Geotechnical and Geophysical Site Characterisation 5 – Lehane, Acosta-Martínez & Kelly (Eds)

Frequency Spectrum "Bell-Curve" Fitting as a Component of SCPT Interval Velocity Accuracy Assessment

E. Baziw & G. Verbeek
Baziw Consulting Engineers Ltd., Vancouver, Canada

ABSTRACT: Seismic Cone Penetration Testing (SCPT) is an important geotechnical testing technique for site characterization that provides low strain ($<10^{-5}$) *in-situ* interval compression (V_p) and shear (V_s) wave velocity estimates. A challenging problem is to obtain an accuracy assessment of the quality of these calculated interval velocities. The accuracy assessment should take into account various independent characteristics from the acquired data for each depth increment, which are then fused together into a classification or "rank" that quantifies the accuracy of the estimated interval velocities. In 2015, the authors introduced an initial reliable and unique interval velocity classification (IVC) technique which utilizes linearity estimates from Polarization Analysis (PA) in conjunction with Cross Correlation Coefficient (CCC) calculations of the full waveforms. At that time it was suggested that the quality assessment of the estimated interval velocities should be based on the equal weight of the PA linearity and the CCC between the full waveforms calculated at sequential depths. In this paper, the mathematical and implementation details of a new parameter introduced into the IVC are outlined. This new parameter quantifies the deviation of the source wave frequency spectrum from a desirable bell-shaped curve. SCPT seismic traces with high signal to noise ratios were found to have characteristically bell-shaped curves similar to the probability density of a normal distribution.

1 INTRODUCTION

1.1 *Background*

Seismic Cone Penetration Testing (SCPT) (ASTM D7400, 2013) is a common geotechnical technique for measuring *in-situ* compression and shear wave velocities (V_P and V_S respectively). The main goal in SCPT is to obtain arrival times as the source wave travels through the soil profile of interest, and from these arrival times interval velocities are then calculated. Once the velocities have been calculated it is highly desired to assign a confidence level to the estimate. Typically, investigators have utilized the Cross Correlation Coefficient (CCC) between seismic traces recorded at successive depths. The CCC gives an indication of the similarity between the traces (Baziw, 1993) and ranges in value from zero (no similarity) to unity (perfect similarity).

However, the CCC has proven to be an unreliable indicator of the accuracy of the calculated interval velocity, because measurement noise (both random and systematic) can also be correlated and generate erroneously high CCC values. In addition, the CCC values are affected by the digital frequency filters that are applied to increase the Signal to Noise Ratios (SNRs) of the acquired seismic data files.

In order to overcome the limitations of the CCC the authors developed (Baziw and Verbeek, 2016) a new and significantly more comprehensive approach to SCPT interval velocity classification (*IVC*) and assessment, similar to the methodology Ge (Ge and Mottahed, 1993 and 1994, Ge, 2003) developed for microseismic source location estimation. In this approach various seismic data characteristics obtained from the acquired microseismic data for each event are fused together based upon analytic, derived and evolving empirical relationships. The outcome is a microseismic source location estimate "rank", varying from A (very good), B (good), C (acceptable), and D (not acceptable).

In the initial approach the authors combined linearity estimates from Polarization Analysis (PA) (Baziw, 2004, Baziw and Verbeek, 2016, Kanasewich, 1981) with the CCC values to assess the quality of the calculated interval velocities. This approach assumes that full waveforms and ray path refraction are utilized within the interval velocity calculation (Aki and Richards, 2002, Baziw, 2002 and 2004, Baziw and Verbeek, 2012).

The PA linearity approaches unity when the rectilinearity of the responses on the X and Y axes is high, which will be the case for seismic traces recorded in a transverse isotropic (TI) medium (Shearer, 1999, Aki and Richards, 2002) with minimal measurement noise, clean source waves, and no signal distortions (e.g., reflections). The more any of these aspects are present the lower the linearity value will be and low linearity values will also be obtained when shear-wave splitting or shear wave birefringence occur (Gibowicz and Kijko, 1994).

The authors suggested that the quality assessment should be based on the equal weight of the linearity and the CCC between the full waveforms calculated at sequential depths:

$$IVC = (LinearityDepth_1 + LinearityDepth_2 + CCC_{between\ Depth1\ and\ Depth2})/3). \quad (1)$$

This initial weighting was based on a re-evaluation of many data sets previously processed (from over 40 different sites around the world, covering over 4000 seismic traces).

The *IVC* value was then converted into a grade ranging from *A* to *F* as shown below, where A is highly desirable and *F* is unusable:

Classification	
[0-1]	***[A-F]***
0.9 to 1.0	***A***
0.8 to 0.9	***B***
0.7 to 0.8	***C***
0.6 to 0.7	***D***
< 0.6	***F***

It should be noted that the authors indicated that this classification would be refined in the future as it is applied on a routine basis for new data sets.

After proposing the initial approach the authors have reviewed numerous SCPT data sets and attempted to identify other independent characteristics which define the quality of the acquired SCPT seismic data, especially for horizontal shear (SH) wave analysis where dominant responses reside on the X and Y axis. As part of this process they considered among others the amplitude responses on the Z component of a triaxial sensor package (minimal Z component responses would be expected for SH wave analysis), the correlation between right and left polarized wave interval velocity estimates, and how well true Vertical Seismic Profile (VSP) arrival times match–up with arrival times derived from relative arrival times obtained from cross-correlating full waveforms at successive depths. Unfortunately, the first proved to be ad hoc, while the last two were not conducive to automated data processing.

A more promising feature for incorporation into the *IVC* was the shape of the seismic trace frequency spectrum, or better the deviation of the source wave frequency spectrum from a desirable bell-shaped curve, as SCPT seismic traces with high signal to noise ratios were found to have characteristically bell-shaped curves similar to the probability density of a normal distribution. This approach is presented in this paper.

2 FREQUENCY SPECTRUM "BELL CURVE" FITTING

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

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

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 1 illustrates example of normal pdfs for varying μ and σ^2 values. All the curves in Fig. 1 have the classical bell-shape.

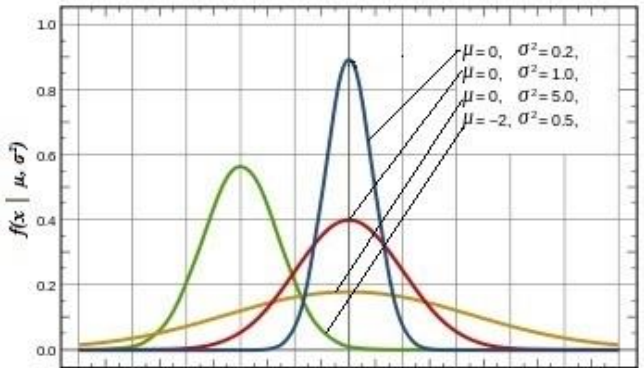

Figure 1. illustrates example of normal pdfs for varying μ and σ2 values. (after, http://www.dplot.com/probability-scale.htm)

Figure 2 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

$$w(t) = AH(t)t^n e^{-ht}\cos(2\pi ft + \emptyset) \quad (3)$$

where $H(t)$ is the Heaviside unit step function [$H(t) = 0$ for $t \leq 0$ 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 3 illustrates the frequency spectrum (solid black line) of the Berlage source wave shown in Fig. 2 with the normal pdf approximation shown as a dotted grey line, with μ = 69 Hz and σ = 32.5 . As is evident from Fig. 3, the frequency spectrum of the simulated Berlage source wave closely matches that of a bell-shaped curve.

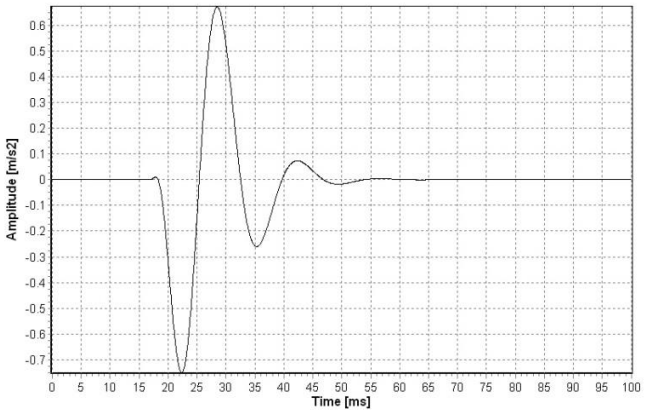

Figure 2. Berlage source wave with of $f = 70$ Hz, $n = 2$, $h = 270$ and $\phi = 40^o$ specified.

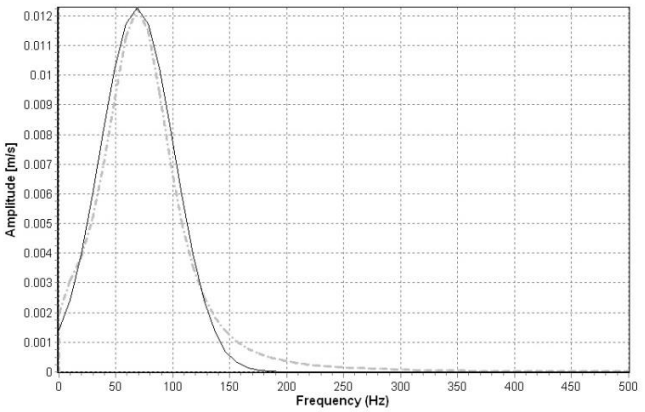

Figure 3. Frequency spectrum (solid black line) of Berlage source wave illustrated in Fig. 2 with the normal pdf approximation shown as a dotted grey line.

2.1 *Bell-Curve fitting algorithm, SNR estimation and incorporation into IVC*

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

1. Apply a digital zero-phase shift frequency filter (e.g., 200 Hz low pass frequency filter) to triaxial sensor data set so that high frequency measurement noise is removed.

2. For SH source wave analysis, calculate frequency spectra 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. 4(a) and 4(b).

$$Area_{S(f)} = \Delta f \sum_{i+1}^{n} S(f)_i \quad (4a)$$

$$\sum_{i+1}^{n} S(f)_i = S(f)_i / Area_{s(f)} \quad (4b)$$

In eq. 4(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. (2)) 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. Spectrum Rank (*SR*) component of the *IVC* is then calculated $SR = 1\text{-}R$.

10. The *IVC* is then calculated as the average of the linearity value determined by applying PA and the SR value.

$$IVC = ((Linearity_{Depth1} + SR_1)/2 + (Linearity_{Depth2} + SR_2)/2 + CCC_{between\ Depth1\ and\ Depth2})/3 \quad (5)$$

3 REAL DATA EXAMPLES

3.1 *High SNR SCPT data set*

Figure 4 illustrates a real data example of a high SNR SH source wave SCPT trace acquired in Alberta, Canada. The triaxial data was acquired at a depth of 5m by fast response accelerometers (5μs response time) with a bandwidth of 10 Hz to 10 KHz manufactured by PCB Piezotronics. The triaxial seismic data displayed in Fig. 4 had an option referred as *normalize locally* implemented. Normalize locally refers to normalizing the displayed amplitudes of the X(t), Y(t), and Z(t) axes with respect to the absolute maximum value. This provides for a clear visual of the dominant X(t), Y(t) and Z(t) dominant responses.

The frequency spectra are calculated from the filtered traces of Fig. 4 (200 Hz low pass filter applied) as outlined in Step 2. The dominant frequency components were determined to reside within the X axis recordings. Figure 5 illustrates the X axis frequency spectrum (solid black line) of triaxial source wave illustrated in Fig. 4 where the frequency spectrum amplitudes were uniformly modified so that the area under the curve was unity (Step 3).

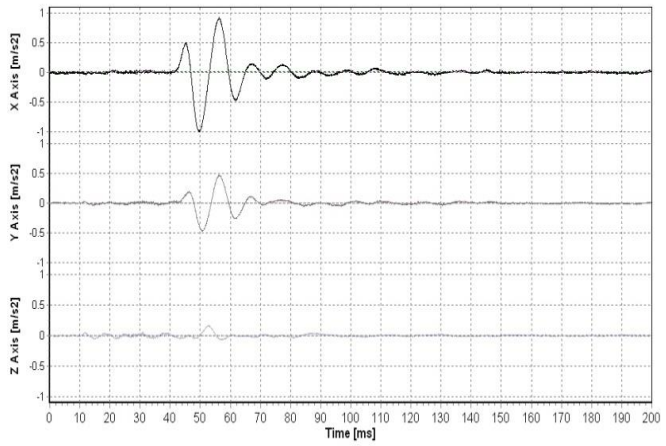

Figure 4. Unfiltered high SNR SCPT triaxial accelereometer data set.

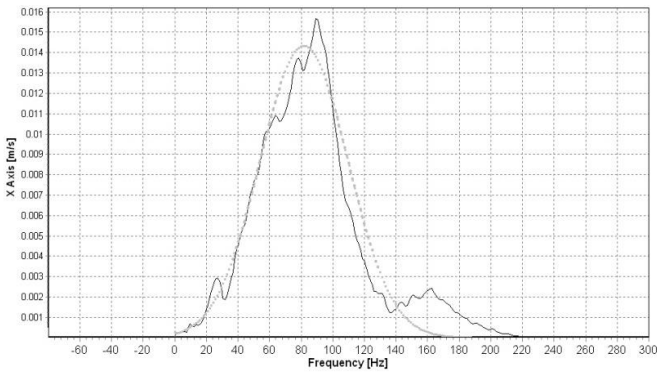

Figure 5. X axis frequency spectrum (solid black line) of triaxial source wave illustrated in Fig. 4 with the normal pdf approximation shown as a dotted grey line.

The spectrum in Fig. 5 is then fed into the IFM portion of the bell-curve fitting algorithm where optimal values of μ = 82 Hz and σ = 27.8 were obtained. Figure 5 illustrates the estimated normalized pdf superimposed upon the unity area frequency spectrum. Values of ε_1 = 0.163, ε_2 = 0.816, R = 0.20 and SR = 0.8 were obtained. This high SR value indicates a high quality seismic trace, which is confirmed by the linearity value for this trace of 0.976 (using a standard time window, see 3.2).

3.2 *Low SNR SCPT data set*

Figure 6 illustrates another real data example where we have a low SNR SH source wave SCPT trace acquired in New Zealand. The triaxial data was acquired at a depth of 8m by fast response accelerometers (5μs response time) with a bandwidth of 10 Hz to 10 KHz manufactured by PCB Piezotronics. The triaxial seismic data displayed in Fig. 6 had option *normalize locally* applied. The frequency spectrums are calculated from the filtered traces of Fig. 6 (200 Hz low pass filter applied) as outlined in Step 2. The dominant frequency components were determined to reside within the Y axis recordings. Figure 7 illustrates the 7 axis frequency spectrum (solid black line) of triaxial source wave illustrated in Fig. 6 where the frequency spectrum amplitudes were uniformly modified so that the area under the curve was unity (Step 3).

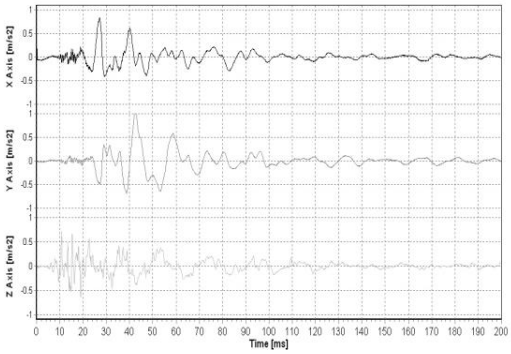

Figure 6. Unfiltered low SNR SCPT triaxial accelerometer data set

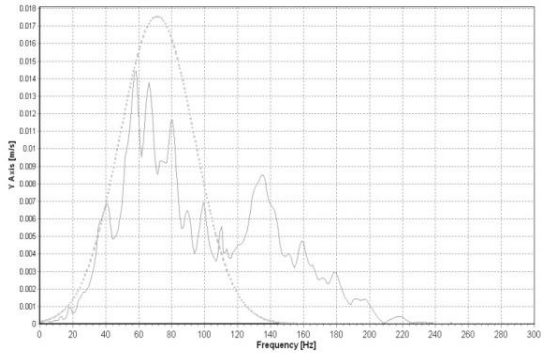

Figure 7. Y axis frequency spectrum (solid black line) of triaxial source wave illustrated in Fig. 6 with the normal pdf approximation shown as a dotted grey line.

The spectrum in Fig. 7 is then fed into the IFM portion of the bell-curve fitting algorithm where optimal values of μ = 71 Hz and σ = 22.7 were obtained. Figure 7 illustrates the estimated normalized pdf superimposed upon the unity area frequency spectrum. Values of ε_1 = 0.6, ε_2 = 0.907, R = 0.66 and SR = 0.34 were obtained. The value of SR = 0.34 identifies a low quality seismic trace. As could be expected the linearity value obtained for this trace is also lower (0.691).

It should be noted, however, that the linearity value can be significantly dependent upon the time window that is applied. The value listed above is based on a standard window covering 28 ms on either side of the first peak (the larger time window allows for greater source wave and trace characterization). If we were to reduce that to 14 ms the value increases to 0.779, and with an even narrower time the value can be as high as 0.82.

Whenever there is a low SR value and a low linearity value batch or automated processing should be avoided and arrival times should be obtained visually from VSPs and identifying first breaks or dominant peaks or trough.

4 CONCLUSIONS

SCPT is an important geotechnical testing technique for site characterization. SCPT provides low strain ($<10^{-5}$) in-situ interval compression (V_p) and shear (V_s) wave velocity estimates. A challenging problem in SCPT is to obtain an accuracy assessment of the quality of the calculated interval velocities. The accuracy assessment should take into account various independent seismic time series characteristics of the acquired data, which are then fused together into an interval velocity classification (*IVC*) or "rank" which quantifies the accuracy of the estimated interval velocities. This paper builds upon previous *IVC* development work where the authors have fused linearity estimates from polarization analysis in conjunction with cross correlation coefficient calculations of the full waveforms. In this paper, the mathematical and implementation details of a new parameter introduced into the *IVC* algorithm were outlined. This new parameter, denoted as *SR*, quantifies the deviation of the source wave frequency spectrum from a desirable bell-shaped curve. As outlined in this paper, SCPT seismic traces with high Signal to Noise Ratios (SNRs) were found to have characteristically bell-shaped curves similar to the probability density of a normal distribution. It was shown when processing real SCPT data sets that the bell-curve *SR* value is highly indicative of the SNR of the acquired trace. For seismic traces with low SNR batch or automated processing should be avoided and arrival times should be obtained visually from vertical seismic profiles identifying first breaks or dominant peaks or trough.

5 REFERENCES

Aki, K., and Richards, P.G. 2002. Quantitative Seismology (2nd Edition). Sausalito, CA: University Science Books.

ASTM. D7400-08. 2013. Standard Test Methods for Downhole Seismic Testing. ASTM Vol. 4.09 Soil and Rock (II): D5877-latest.

Baziw, E. 1993. Digital filtering techniques for interpreting seismic cone data, Journal of Geotechnical Engineering. ASCE, Vol. 119, No. 6, 98-1018.

Baziw, E. 2002. Derivation of Seismic Cone Interval Velocities Utilizing Forward Model. Can. Geotech. J., vol. 39,1-12.

Baziw, E. 2004. Two and three dimensional imaging utilizing the seismic cone penetrometer. In Proceedings of the 2nd International Conference on Geotechnical Site Characterization (ISC-2), Porto, Portugal, 19-22 Sept. Millpress Science Publishers, 1611-1618.

Baziw, E. and Ulrych, T.J. 2006. Principle Phase Decomposition - A New Concept in Blind Seismic Deconvolution, IEEE Transactions on Geosci. Remote Sensing (TGRS), vol. 44, no. 8, 2271-2281.

Baziw, E. 2011. Incorporation of Iterative Forward Modeling into the Principle Phase Decomposition Algorithm for Accurate Source Wave and Reflection Series Estimation. IEEE Transactions on Geosci. Remote Sensing (TGRS), vol. 49, no. 2, 650-660.

Baziw, E., and Verbeek, G. 2012. Deriving Interval Velocities from Downhole Seismic Data. Geotechnical and Geophysical Site Characterization 4 – Mayne (eds), CRC Press, 1019–1024.

Baziw, E. and Verbeek, G. 2014. Signal Processing Challenges when Processing DST and CST Seismic Data containing TIRs. ASTM International - Geotechnical Testing Journal (GTJ), vol. 37, no. 3, 1-21.

Baziw, E. and Verbeek, G. 2016. Classification Technique for Assessing Estimated Interval Velocity Estimates in DST, submitted to the ASTM International - Geotechnical Testing Journal.

GE, M. 2003. Analysis of source location algorithms, Parts 1 and II. J. of Acoustic Emission, vol. 21, 14-18 and 29-51.

Ge M., and Mottahed P. 1993. An automatic data analysis and source location system. Proceedings 3rd International Symposium on Rockbursts and Seismicity in Mines A A Balkema Rotterdam, 343-348.

Gelb, A. 1974. Applied Optimal Estimation, 4th ed., MIT Press: Cambridge, Mass, USA.

Gibowicz, S.J., and Kijko, A.1994. An Introduction to Mining Seismology, Academic Press, San Diego, CA, USA.

Kanasewich, E.R., 1981, Time sequence analysis in geophysics. 3rd ed., The University of Alberta Press, Edmonton, AB. Chapter 19.

Shearer, P.M. 1999. Introduction to Seismology. 1st edition, Cambridge: Cambridge University Press.