Identification of Thin Soil Layers Utilizing the q_m*HMM-IFM* Algorithm on Cone Bearing Measurements

Erick Baziw, Ph.D.,P.Eng.[1] and Gerald Verbeek, B.Sc., M.Sc.[3]

[1] Baziw Consulting Engineers Ltd., 3943 West 32nd Avenue, Vancouver B.C., Canada V6S 1Z4; E-mail: ebaziw@bcengineers.com
[2] Baziw Consulting Engineers Ltd., 1411 Cumberland Road, Tyler, TX USA 75703; E-mail: gverbeek@bcengineers.com

ABSTRACT: The cone penetration test (CPT) is a widely used geotechnical tool which provides excellent stratigraphic detail and information for estimating a wide range of soil properties. CPT consists of pushing at a constant rate an electronic penetrometer into penetrable soils and recording the resistance to the cone tip or cone bearing (q_m). The q_m values are utilized to characterize the soil profile along with measured sleeve friction and pore pressure. Cone bearing measurements at a specific depth are blurred or averaged due to q_m values being strongly influenced by soils within 10 to 30 cone diameters from the cone tip. This blurring of the true cone tip readings results in the inability to identify thin and low bearing soils which are masked by the properties of the adjacent soils. The q_m*HMM-IFM* algorithm was developed was developed to address the q_m blurring\averaging limitation. The q_m*HMM-IFM* algorithm implements a hybrid hidden Markov model and iterative forward modelling technique so that true cone bearing are obtained from the averaged/blurred q_m values. This paper demonstrates the performance of the q_m*HMM-IFM* algorithm by analyzing very challenging test bed simulations where thin soil layers are interspersed within uniform soils. It is critical to characterize a proposed algorithm capabilities by analyzing numerous test bed simulations prior to implementing on real data sets. It has been our experience in the industry that developed geotechnical signal processing software typically carries out minimal test bed simulations to verify an algorithms performance prior to implementing on real data sets.

INTRODUCTION

The Cone Penetration Test (CPT) is extensively utilized in geotechnical engineering to identify of sub-surface soils and their associated geotechnical properties (Lunne et al., 1997; Robertson, 1990; ASTM D6067, 2017). In addition, the CPT is utilize to estimate toe bearing capacity of piles (Eslami and Fellenius, 1995 and 1997). In CPT a steel cone is pushed vertically into the ground at a typical standard rate of 2cm per second and data are recorded at constant rate during penetration (typically every 1cm to 2cm). The cone penetrometer has electronic sensors to measure penetration resistance at the tip (q_m) and friction in the shaft (friction sleeve) during penetration. A CPTU probe is equipped with a pore-water pressure sensor and is called a piezo-cones. For piezo-cones with the filter element right behind the cone tip (i.e., the u_2 position) it is standard practice to correct the recorded tip resistance for the impact of the pore pressure on the back of the cone tip. Figure 1 illustrates a schematic and the associated terminology of a cone penetrometer.

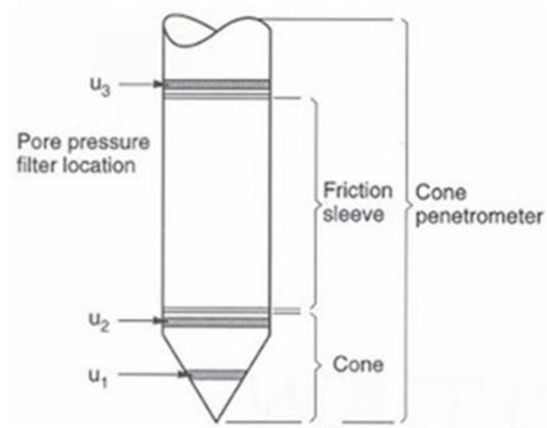

Fig. 1. Schematic and terminology for cone penetrometer (Lunne et al., 1997).

The cone tip resistance measured at a particular depth is affected by the values above and below the depth of interest which results in an averaging or blurring of the q_v values (Boulanger and DeJong, 2018; Robertson, 1990; Baziw and Verbeek, 2021A). This phenomenon is especially of concern when mapping thin soil layers which is critical for liquefaction assessment. Mathematically the measured cone tip resistance q_m is described as (Baziw and Verbeek, 2021A)

$$q_m(d) = \sum_{j=1}^{60\times\left(\frac{d_c}{\Delta}\right)} w_c(j) \times q_v\left(\Delta_{qm} + j\right) + v(d)$$
$$\Delta_{qm} = (d - \Delta_{wm}), \;\; \Delta_{wm} = 30 \times \left(\frac{d_c}{\Delta}\right) \tag{1}$$

where

- d cone depth (m)
- d_c cone tip diameter (m)
- Δ q_m sampling rate (m)
- $q_m(d)$ measured cone penetration tip resistance (MPa)
- $q_v(d)$ true cone penetration tip resistance (MPa)
- $w_c(d)$ the $q_v(d)$ averaging function (dimensionless)
- $v(d)$ additive noise, generally taken to be white with a Gaussian probability distribution function (PDF) (MPa)

In eq. (1) it assumed that w_c averages q_v over 60 cone diameters centered at the cone tip. Boulanger and DeJong (Boulanger and DeJong, 2018) outline how to calculate w_c below (after correcting the equation for w_1 (Baziw and Verbeek, 2021)):

$$w_c = \frac{w_1 w_2}{\sum w_1 w_2} \tag{2a}$$

$$w_1 = \frac{C_1}{1 + \left|\left(\frac{z'}{z'_{50}}\right)^{m_z}\right|} \tag{2b}$$

$$w_2 = \sqrt{\frac{2}{1 + \left(\frac{q_{v,z'}}{q_{v,z'=0}}\right)^{m_q}}} \tag{2c}$$

where:

- w_1 accounts for the relative influence of any soil decreasing with increasing distance from the cone tip.
- w_2 adjusts the relative influence that soils away from the cone tip will have on the penetration resistance based on whether those soils are stronger or weaker.
- z' the depth relative to the cone tip normalized by the cone diameter.
- z'_{50} the normalized depth relative to the cone tip where $w_1 = 0.5\ C_1$.
- C_1 equal to unity for points below the cone tip, and linearly reduces to a value of 0.5 for points located more than 4 cone diameters above the cone tip.
- m_z exponent that adjusts the variation of w_1 with z'.
- m_q exponent that adjusts the variation of w_2 with $\left(\frac{q_{v,z'}}{q_{v,z'=0}}\right)$.

Boulanger and DeJong (2018) provide a thorough outline and review on the setting of the parameters given in eq. (2) based upon extensive research and modelling. In general terms, soils in front of the cone tip have a greater influence on penetration resistance than the soils behind the cone tip. In the subsequently outlined test bed simulations the parameters in eq. (2) are set identical to those outlined by Boulanger and DeJong. In this case, exponents $m_q = 2$ and $m_z = 3$.

Baziw and Verbeek (2021A) developed an algorithm (the so called q_m*HMM-IFM*) to optimally obtain true q_v cone bearing estimates from blurred measurements q_m. The q_m*HMM-IFM* algorithm combines the Bayesian recursive estimation (BRE) Hidden Markov Model (HMM) filter with Iterative Forward Modelling (IFM) parameter estimation in a smoother formulation for optimal estimation. Preliminary test bed analysis of the q_m*HMM-IFM* algorithm demonstrated it to be a very promising mathematical tool for obtaining q_v estimates from measured cone bearing values. Subsequent to the initial work of the authors (Baziw and Verbeek (2021A)) upgrades and modifications of the q_m*HMM-IFM* algorithm have been made and additional challenging and extensive new test bed analysis has been carried out. This paper briefly outlines the current q_m*HMM-IFM* algorithm formulation and present the results from very challenging test bed simulations. The test bed simulations focused on extracting masked thin bed layers.

q_m*HMM-IFM* ALGORITHM FILTER FORMULATION

The q_m*HMM-IFM* algorithm implements a hybrid BRE HMM filter and IFM filter formulation. Baziw and Verbeek (2021A) outline the details of the BRE, HMM and IFM signal processing and optimal estimation tools. For completeness the details of q_m*HMM-IFM* algorithm HMM and IFM components are summarized.

HMM Filter Formulation

The HMM filter (also termed a grid-based filter) has a discrete state-space representation and has a finite number of states. In the HMM filter the posterior PDF is represented by the delta function approximation as follows:

$$p(x_{k-1}|z_{1:k-1}) = \sum_{i=1}^{N_s} w^i_{k-1\backslash k-1}\delta\left(x_{k-1} - x^i_{k-1}\right) \tag{3}$$

where x^i_{k-1} and $w^i_{k-1|k-1}$, $i = 1,\ldots,N_s$, represent the fixed discrete states and associated conditional probabilities, respectively, at time index *k-1*, and N_s the number of particles utilized. In the case of the q_m*HMM-IFM* algorithm the HMM discrete states represent possible q_v values where maximum, minimum and resolution values are specified. The HMM governing equations are outlined in Table 1.

Iterative Forward Modelling

Iterative forward modeling (IFM) is a parameter estimation technique which is based upon iteratively adjusting the parameters until a user specified cost function is minimized. The desired parameter estimates are defined as those which minimize the user specified cost function. The IFM technique which is utilized within the q_v estimation algorithm is the downhill simplex method (DSM) originally developed by Nelder and Mead (Nelder and Mead, 1965). A simplex defines the most elementary geometric figure of a given dimension: a line in one dimension, the triangle in two dimensions, the tetrahedron in three, etc; therefore, in an N-dimensional space, the simplex is a geometric figure that consists of N+1 fully interconnected vertices. The DSM starts at N + 1 vertices that form the initial simplex. The initial simplex vertices are chosen so that the simplex occupies a good portion of the solution space. In addition, it is also required that a scalar cost function be specified at each vertex of the simplex. The general idea of the minimization is to keep the minimum within the simplex during the optimization, at the same time decreasing the volume of the simplex. The DSM searches for the minimum of the costs function by taking a series of steps, each time moving a point in the simplex away from where the cost function is largest. The simplex moves in space by variously reflecting, expanding, contracting, or shrinking. The simplex size is continuously changed and mostly diminished, so that finally it is small enough to contain the minimum with the desired accuracy.

Table 1. HMM Governing Equations			
STEP	DESCRIPTION	MATHEMATICAL REPRESENTATION	EQ.
1	Initialization (k=0) – initialize particle weights.	e.g., $w_k^i \sim 1/N_s$, $i = 1,\ldots,N_s$.	(4)
2	Prediction - predict the weights.	$w_{k\backslash k-1}^i = \sum_{j=1}^{N_s} w_{k-1\backslash k-1}^j p\left(x_k^i \mid x_{k-1}^j\right)$	(5)
3	Update - update the weights.	$w_{k\backslash k}^i = \frac{w_{k\backslash k-1}^i p\left(z_k \mid x_k^i\right)}{\sum_{j=1}^{N_s} w_{k\backslash k-1}^j p\left(z_k \mid x_k^j\right)}$	(6)
4	Obtain optimal minimum variance estimate of the state vector and corresponding error covariance.	$\hat{x}_k \approx \sum_{i=1}^{N_s} w_{k\mid k}^i x_k^i$ $P_{\hat{x}_k} \approx \sum_{i=1}^{N_s} w_{k\mid k}^i (x_k^i - \hat{x}_k)(x_k^i - \hat{x}_k)^T$	(7)
5	Let k = k+1 & iterate to step 2.		
In the above equations it is required that the likelihood pdf $p(z_k \mid x_k^i)$ and the transitional probabilities $p(x_k^i \mid x_{k-1}^j)$ be known and specified.			

q_mHMM-IFM **Algorithm**

The HMM portion of the *q_mHMM-IFM* algorithm implements a BRE smoother. BRE smoothing uses all measurements available to estimate the state of a system at a certain time or depth in the q_v estimation case (Arulampalam et al., 2002; Baziw, 2007; Gelb, 1974). This requires both a forward and backward filter formulation. The forward HMM filter ($\hat{q}_k^F$) processes measurement data (q_m) above the cone tip ($j = 1\ to\ 30 \times \left(\frac{d_c}{\Delta}\right)$ in (1)). Next the backward HMM filter ($\hat{q}_k^B$) is implemented, where the filter recurses through the data below the cone tip ($j = 30 \times \left(\frac{d_c}{\Delta}\right)\ to\ 60 \times \left(\frac{d_c}{\Delta}\right)$ in (1)) starting at the final q_m value. The optimal estimate for q_v is then defined as

$$\hat{q}_k^v = (\hat{q}_k^F + \hat{q}_k^B)/2 \tag{8}$$

where the index k represents each q_m measurement.

In both the forward and backward HMM filter formulation a bank of discrete q_v values ($i = 1$ *to* N) varying from low (q_{tL}) to high (q_{tH}) and a corresponding q_t resolution q_{tR} are specified. The required number of fixed grid HMM states is given as $N_S = (q_{tH} - q_{tL})/q_{tR}$. In Table 1 the notation of the states x^i is mapped to q^i to reflect the bank of q_t values. The current *q_mHMM-IFM* formulation automatically sets the minimum and maximum limits of the q_v values based upon the minimum and maximum cone bearing values q_m measured.

$$q_{tL} = q_{min} - 0.5q_{min}, \; where \; q_{min} \geq 0 \tag{9a}$$

$$q_{tH} = q_{max} + 0.5q_{max} \tag{9b}$$

In eqs. 9(a) and 9(b) q_{min} and q_{max} denote the minimum and maximum q_m values measured, respectively.

In the q_m*HMM-IFM* HMM forward and backward filter formulation the transitional probabilities (i.e., $p\left(x_k^i|x_{k-1}^j\right)$ or $p\left(q_k^i|q_{k-1}^j\right)$) for each HMM state (i.e., discrete cone tip, q^i) is set equal due to the fact that there is equal probability of moving from a current cone tip value to any other value between the range q_{tL} to q_{tH}. The likelihood PDF $p(z_k|q_k^i)$ in the HMM filter outlined in Table 1 is calculated based upon an assumed Gaussian measurement error as follows:

$$p(z_k|q_k^i) = \frac{1}{\sqrt{2\pi\sigma}} e^{\left[-\frac{\left(q_m(d)-z_k^i\right)}{2\sigma^2}\right]} \tag{10}$$

where $\sigma^2$is the variance of the measurement noise. Baziw and Verbeek (2021A) outline the details of the q_m*HMM-IFM* algorithm HMM forward and backward filter formulation.

IFM is incorporated into the q_m*HMM-IFM* algorithm so that initially estimated IFM q_v estimates are feed into the HMM smoothing filter as initial values. This results in significantly more accurate results. Instead of attempting to estimate all the unknown q_v values with the HMM smoother (below the cone depth for the forward HMM analysis, and above the cone for the backward HMM analysis) IFM is utilized where only a fraction of the q_v values are required to be estimated. In this process constant layer q_v values and their corresponding depth extents are estimated for a maximum number of layers within the next w_c window. Baziw and Verbeek (2021) elaborate on the IFM portion of the q_m*HMM-IFM* algorithm.

q_tHMM-IFM **THIN LAYER TEST BED SIMULATIONS**

The performance of the q_m*HMM-IFM* algorithm was evaluated by carrying out challenging test bed simulations of variable thin bed layering. In addition, the thin bed layer challenges outlined by Boulanger and DeJong (2018) is revisited. In the test bed simulations outlined below the measured blurred\averaged cone bearing values q_m are generated by applying eq. (1) to the true cone bearing values q_v.

Test Bed Simulation 1

In this test bed simulation strong thin layers are interspersed within a weak uniform cone soil profile bearing. Figure 3 illustrates a true cone bearing profile (light grey trace) of uniform soil with a cone bearing of 5 MPa except for thin intervals that are 0.1 m, 0.08 m, and 0.15 m thick with corresponding cone bearing values of 100 MPa, 80 MPa, and 60 MPa, respectively. The resulting q_m values were then calculated (black line in Fig. 3). Using the q_m*HMM-IFM* algorithm the q_v values were then estimated based on the q_m values (black dotted line in Fig. 3). It shall be obvious that the algorithm performed well as the derived q_v values closely matched the originally specified q_v values. This test bed simulation analysis is important in that it highlights strong thin layers maybe perceived to be much weaker when they are interspersed within a weak uniform soil. This may erroneously trigger liquefaction concerns.

Test Bed Simulation 2

In this test bed simulation challenging cone bearing profile is processed where there strong and weak thin soil layers interspersed within a uniform cone soil profile. Figure 4A illustrates a true cone bearing profile (light grey trace) of uniform soil with a cone bearing of 80 MPa except for thin intervals that are 0.1 m, 0.08 m, 0.2 m, 0.15 m, and 0.4 m thick with corresponding cone bearing values of 100 MPa, 30 MPa, 100 MPa, 40 MPa and 120 MPa, respectively. The resulting q_m values were then calculated (black line in Fig. 4A). Figure 4B illustrates the percentage difference between q_v and q_m (black line). Using the q_m*HMM-IFM* algorithm the q_v values were then estimated based on the q_m values (black dotted line in Fig. 4A). It shall be obvious that the algorithm performed well as the derived q_v values closely matched the originally specified q_v values. Figure 4B illustrates the percentage difference between the estimated q_v values and the true q_v values (black dotted line).

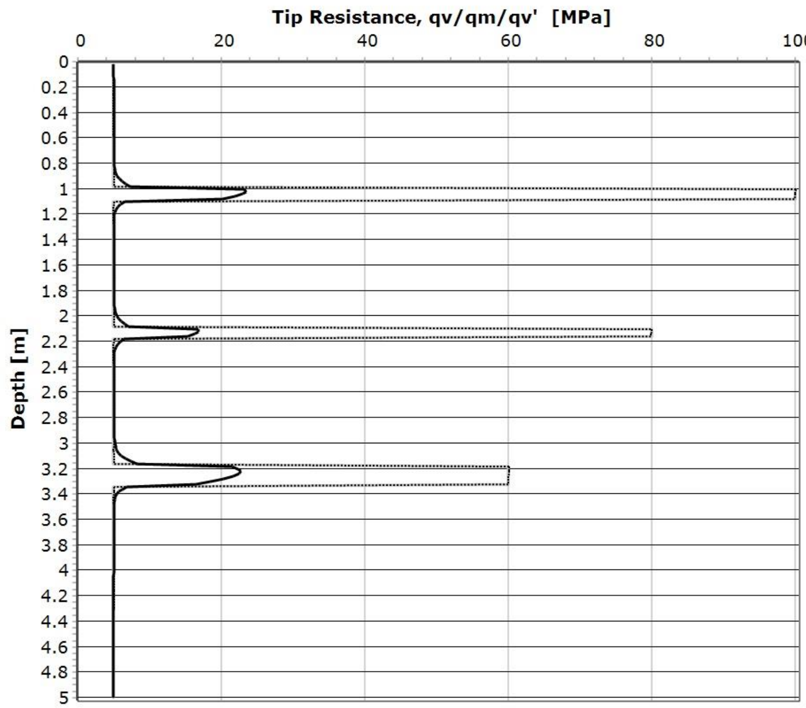

Figure 3. TEST BED 1 Specified q_v values (grey line), derived q_m values (black line) and estimated q_v values based on q_m values (black dotted line).

Test Bed Simulation 3

In this test bed analysis the thin bed layer challenges outlined by Boulanger and DeJong (2018) is revisited. Boulanger and DeJong (2018) simulated the potential effects of what they characterize as "noise" with inverting measurements from very thin interlayers using their "inverse" filtering technique defined by eqs. (9) to (14) of their paper. Boulanger and DeJong illustrate this "noise" by processing the variable thin layers illustrated in Fig. 5A. Figure 5A illustrates a profile of uniform soil with a cone bearing of 10 except for thin intervals that are 1.1, 1.7, 2.2, and 3.4 cone diameters thick and cone bearing values of 12. The simulated data was generated with a 2 cm sampling interval. This corresponds to thin depth intervals having 2, 3, 4, and 6 data points, respectively, with cone bearing values of 12. This simulated cone bearing profile was then feed into Boulanger's and DeJong's "inverse" filtering algorithm to give the very unstable estimates illustrated in Fig. 5A. These unstable results led Boulanger and DeJong to develop and incorporate *ad hoc* low-pass spatial filter and smoother into their "inverse" filtering algorithm.

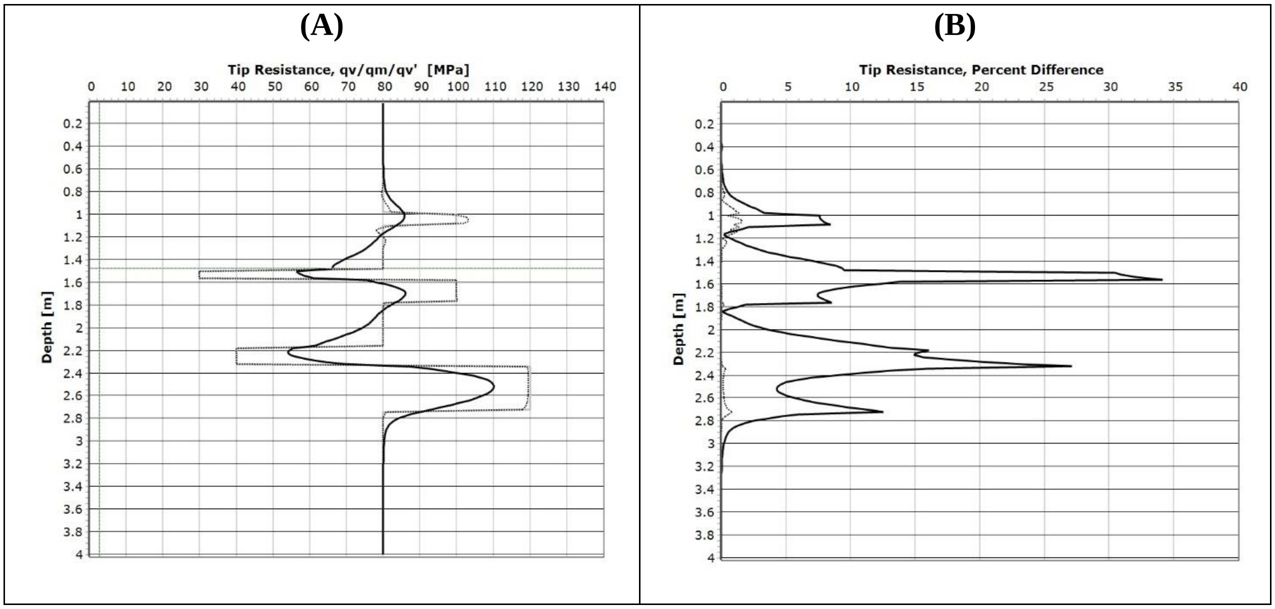

Figure 4. TEST BED 2 (A) Specified q_v values (grey line), derived q_m values (black line) and estimated q_v values based on q_m values (black dotted line). Percent differences between specified and estimated q_v values (black line) and q_m values and estimated q_v values (black dotted line).

Equation (1) outlines an averaging mathematically operation; therefore, the simulated q_m cone bearing values illustrated in Fig. 5A would not be seen in practice and is inappropriate as a test bed simulation. The q_m values have very "sharp" transitions at the 10 to 12 interfaces. This contradicts the averaging operation defined by eq. (1). Processing the q_m values of Fig. 5A results in highly variable q_v estimates. Figure 5B illustrates the near duplication of Boulanger's and DeJong's output results illustrated in Fig. 5A. Figure 5C illustrates the measured q_m values (black trace) when the measured "sharp" q_m values (i.e., mapped to true q_v values) of Fig. 5A (light grey trace) are feed into eq. (1). Clearly the q_m values illustrated in Fig. 5C values are significantly smoothed and reduced in amplitude due to the background q_v value of 10 as expected. Figure 6 illustrates the output of the q_m*HMM-IFM* algorithm when processing the q_m values of Fig. 5C. As is shown in Fig. 6, as thin layer thickness increase the q_m*HMM-IFM* algorithm does a better job of estimating the layer thickness and cone bearing value of the thin bed layer. These are expected results.

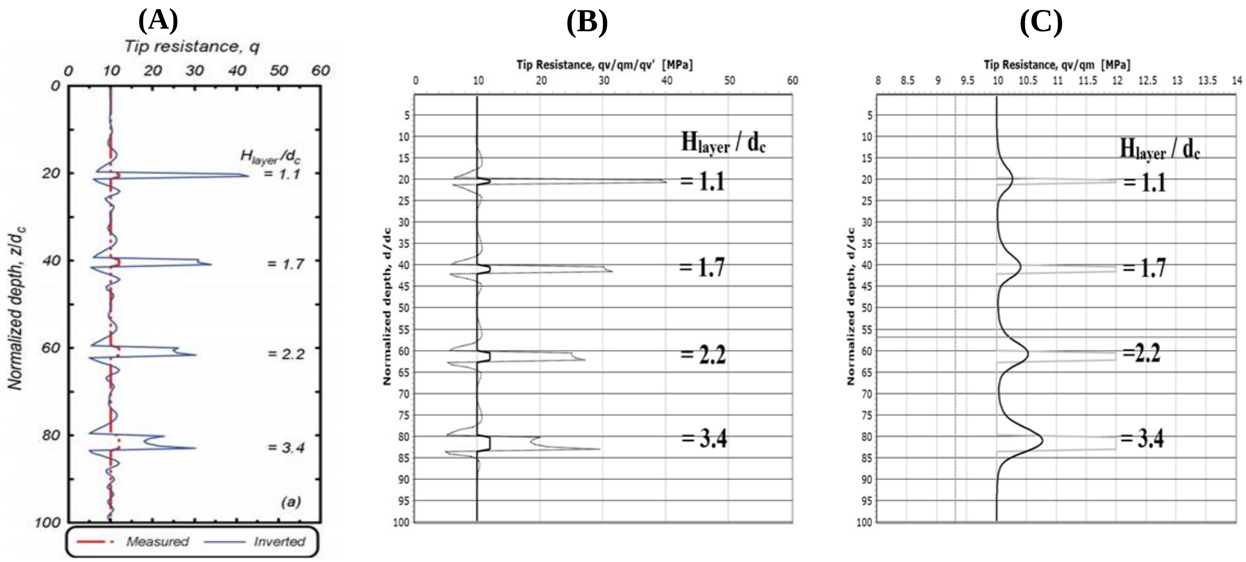

Figure 5. (A) Significant instability in the estimates of q_t when using the Boulanger and DeJong inversion estimation algorithm. (Boulanger and DeJong, 2018). (B) Duplication of the results obtained by Boulanger and DeJong (1981). (C) True q_v values (light grey trace) superimposed upon measured cone bearing values q_m (black trace) (Baziw and Verbeek, 2021B).

CONCLUSION

Cone penetrometer testing (CPT) is an effective, fast and relatively inexpensive system for determining the in-situ subsurface stratigraphy and to estimate geotechnical parameters of the soils present. In CPT, a cone on the end of a series of rods is pushed into the ground at a constant rate and resistance to the cone tip is measured (q_m). The q_m values are utilized to characterize the soil profile. Unfortunately, the measured cone tip resistances are blurred and/or averaged due to the layers above and below the cone tip affecting the measured tip resistance. The blurring of q_m measurements can result in the distortion of the soil profile characterization especially if thin soil layers are present. BCE developed the so called q_m*HMM-IFM* algorithm so that true cone bearing measurements could

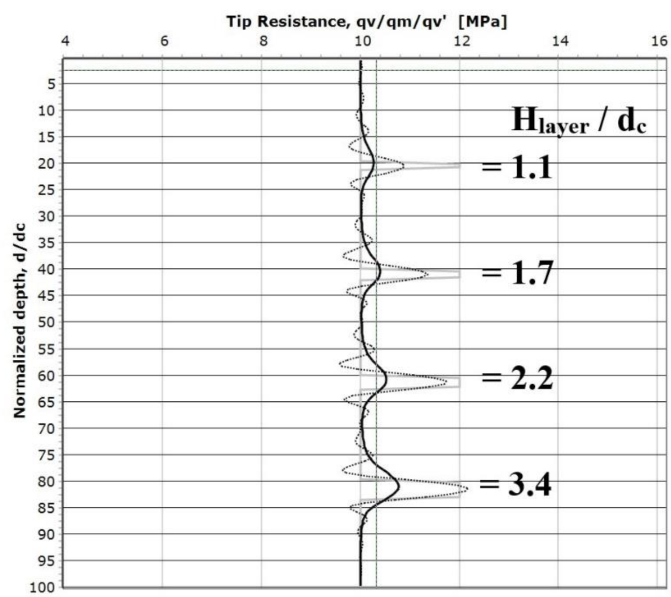

Figure 6. Specified q_v values (grey line) of Fig. 5(C), derived q_m values (black line) of Fig. 5(C) and estimated q_v values based on q_m values (black dotted line).

optimally be extracted from measured values. This paper outlined the current formulation of the q_m*HMM-IFM* algorithm where an iterative forward modelling technique is incorporated into a hidden Markov model filter. This paper demonstrated the performance of the q_m*HMM-IFM* algorithm where test bed simulation were carried out which consisted of variable this soil layers interspersed within uniform soil profiles. In addition, the thin bed layer challenges outlined by Boulanger and DeJong (2018) was revisited. The test bed simulations have demonstrated that the q_m*HMM-IFM* algorithm can derive accurate q_v values from a q_m profile when variable thin soil layers are interspersed within uniform soil profiles. The authors will carry out further test bed simulations and subsequently apply the q_m*HMM-IFM* algorithm on real data sets.

REFERENCES

Arulampalam, M.S., Maskell, S., and Clapp, T. (2002). “A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking”, IEEE Transactions on Signal Processing, vol. 50, no. 2, pp. 174-188, Feb. 2002.

ASTM D6067 / D6067M – 17 (2017). “Standard Practice for Using the Electronic Piezocone Penetrometer Tests for Environmental Site Characterization and Estimation of Hydraulic Conductivity”, ASTM Vol. 4.09 Soil and Rock (II): D5877-latest.

Baziw, E. (2007), Application of Bayesian Recursive Estimation for Seismic Signal Processing, Ph.D. Thesis, Dept. of Earth and Ocean Sciences, University of British Columbia.

Baziw, E. and Verbeek, G. (2021). " Cone Bearing Estimation Utilizing a Hybrid HMM and IFM Smoother Filter Formulation", accepted for publication within the International Journal of Geosciences (IJG) Special Issue on Geoscientific Instrumentation, Methods and Data Systems.

Boulanger, R.W. and DeJong, T.J. (2018). “Inverse filtering procedure to correct cone penetration data for thin-layer and transition effects.” Proc., Cone Penetration Testing 2018, Hicks, Pisano, and Peuchen, eds., Delft University of Technology, The Netherlands, 25-44.

Eslami, A., and Fellenius, B.H. (1995). “Toe bearing capacity of piles from cone penetration test (CPT) data”, In Proceedings of the International Symposium on Cone Penetrometer Testing, CPT ’95, Linkoping, Sweden, 4–5 October 1995. Swedish Geotechnical Society, Gothenburg, Sweden.

Eslami, A. and Fellenius, B.H. (1997). ‘‘Pile capacity by direct CPT and CPTu methods applied to 102 case histories’’, Canadian Geotechnical Journal(CGJ),6, 886–904.

Gelb, A. (1074). Applied Optimal Estimation (4th Edition). Cambridge, Mass: MIT Press.

Nelder, J.A., and Mead, R. (1965). “A simplex method for function optimization. Computing Journal”, 7: 308–313.

Robertson, P.K. (1990). "Soil classification using the cone penetration test". Canadian Geotechnical Journal 27 (1), 151-158.