

Implementation of Kalman Filtering Techniques for Filtering CPT Cone Bearing Measurements

Erick Baziw

Baziw Consulting Engineers Ltd., Vancouver, Canada

Gerald Verbeek

Baziw Consulting Engineers Ltd., Vancouver, Canada

ABSTRACT:

Cone penetration testing (CPT) is a popular and cost effective tool for geotechnical site characterization. CPT consists of pushing at a constant rate an electronic penetrometer into penetrable soils and recording cone bearing (q_m), sleeve friction (f_s) and dynamic pore pressure (u) with depth. The measured q_m , f_s and u values are utilized to estimate soil type and associated properties based predominantly on empirical correlations. The cone bearing measurements are highly susceptible to additive measurement noise for sandy, silty and gravelly soils. This measurement noise results in high peaks due to interbedded gravels and stones and low peaks due to softer materials or local pore pressure build-up. To date there has been little progress in filtering this additive noise aside from *ad hoc* techniques, which include discarding q_m measurements and smoothing/averaging q_m measurements. This paper outlines a q_m filtering technique which attempts to put structure into the estimation problem. In this case, the dynamics of the cone bearing measurements are modelled within a state-space mathematical formulation and a Kalman filter (KF) is then utilized to obtain optimal estimates of q_m . The mathematical details of the q_m KF algorithm are outlined in this paper along with the results from a challenging test bed simulation.

1. INTRODUCTION

The Cone Penetration Test (CPT) is used to determine the subsurface stratigraphy, estimate geotechnical parameters of the soils present (Lunne et al., 1997; Robertson, 1990; ASTM D6067, 2017; Cai et. al, 2006), and estimate toe bearing capacity of piles (Eslami and Fellenius, 1995 and 1997). In CPT a steel cone on the end of a series of rods is pushed into the ground at a constant rate and data are recorded at a constant rate during penetration. The cone penetrometer has electronic sensors to measure penetration resistance at the tip (q_m), friction in the shaft (friction sleeve) and dynamic pore-water pressure during penetration. For 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. This corrected cone tip resistance is normally referred to as q_t .

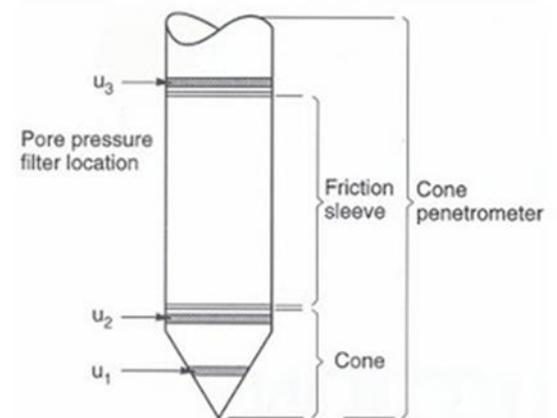


Fig. 1. Schematic and terminology for cone penetrometer (Lunne et al., 1997).

Figure 1 illustrates a schematic and the associated terminology of a cone. Geotechnical engineers use the CPT measurements to characterize and quantify soil properties and ground water conditions so that the infrastructure (e.g., bridges, roads, buildings) construction requirements can be determine.

The cone tip resistance is a smoothing or averaging process (Baziw and Verbeek, 2021) where cone tip values measured at a particular depth are affected by values above and below the depth of interest as illustrated in Fig. 2. The measured cone penetration tip resistance q_m can then be described as

$$q_m(d) = \sum_{j=1}^{60 \times \left(\frac{C_d}{\Delta}\right)} w_c(j) \times q_v(\Delta_{qt} + j) + v(d) \quad (1)$$

$$\Delta_{qt} = (d - \Delta_{wc}), \quad \Delta_{wc} = 30 \times \left(\frac{C_d}{\Delta}\right)$$

where

- d the cone depth
- C_d the cone tip diameter
- Δ the q_t sampling rate
- $q_m(d)$ the measured cone penetration tip resistance
- $q_v(d)$ the true cone penetration tip resistance
- $w_c(d)$ the $q_v(d)$ averaging function
- $v(d)$ additive noise, generally taken to be white with a Gaussian pdf

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_l (Baziw and Verbeek, 2021)):

$$w_c = \frac{w_1 w_2}{\sum w_1 w_2} \quad (2a)$$

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

$$w_2 = \sqrt{\frac{2}{1 + \left(\frac{q_{t,z'}}{q_{t,z'=0}} \right)^{m_q}}} \quad (2c)$$

where:

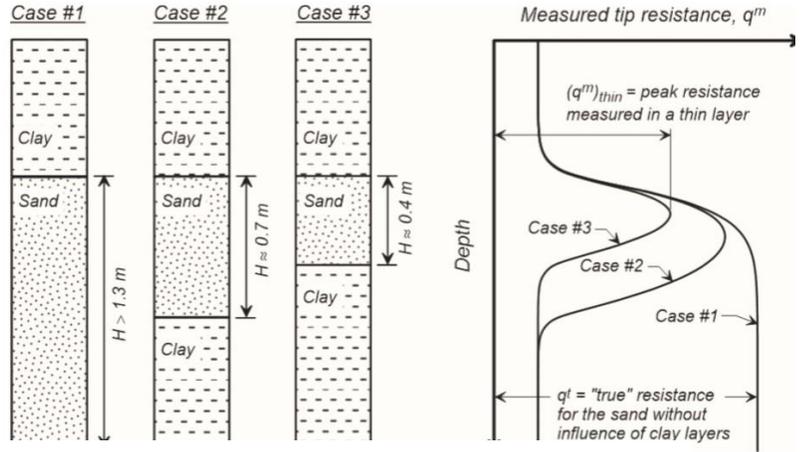


Figure 2. Schematic of thin layer effect for a sand layer embedded in a clay layer (Boulanger and DeJong, 2018).

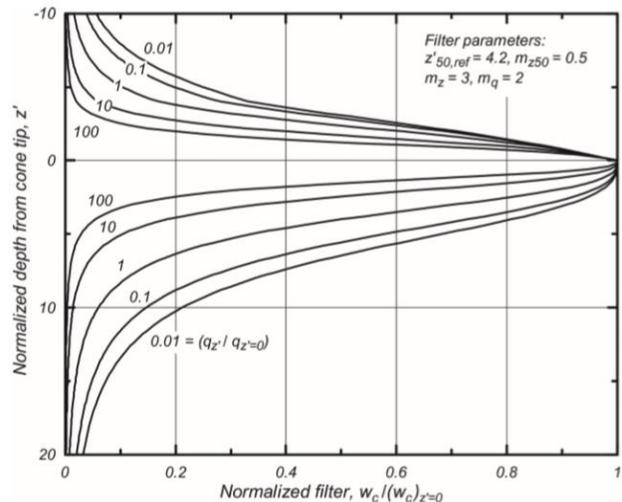


Figure 3. Schematic of thin layer effect for a sand layer embedded in a clay layer (Boulanger and DeJong, 2018).

- z' the depth relative to the cone tip normalized by the cone diameter
- z'_{50} the normalized depth relative to the cone tip where $w_1 = 0.5 C_1$

The cone penetration averaging function w_c for varying $q_{t,z'}/q_{t,z'=0}$ ratios is illustrated in Fig. 3.

The cone bearing measurements q_m are highly susceptible to anomalous peaks and troughs due to the relatively small diameter cone tip penetrating sandy, silty and gravelly soils. The high peaks are due to interbedded gravels and stones and low peaks are due to softer materials or local pore pressure build-up. Lunne, Robertson and Powell (1997) give a detailed outline of these undesired peaks and troughs and refer to them as anomalous and spurious q_m data. Figure 4 is a schematic of the spurious cone bearing data. Figure 5 illustrates an example of a cone bearing profile which contains significant anomalous/spurious q_m data.

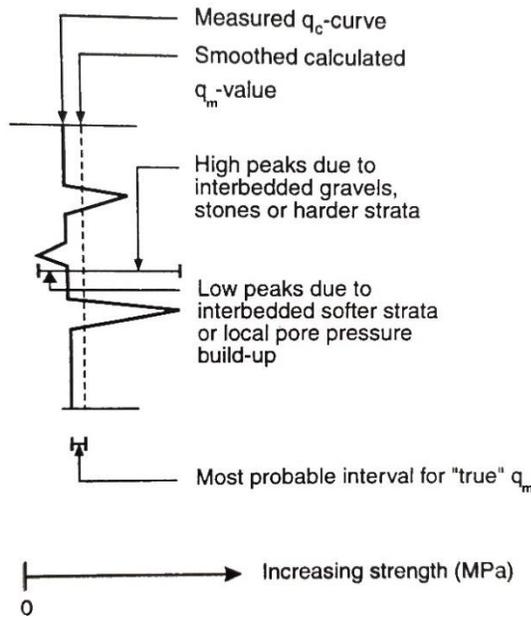


Figure 4. Schematic of anomalous/spurious cone bearing data (after Mortensen et al., 1991).

The spurious q_m data closely resembles additive Gauss-Markov correlated white measurement noise which is extensively present in engineering measurement sensors such as marine navigation dead reckoning devices (Baziw, 1994 and 1996). This type of measurement noise is also found in seismic data acquisition systems (Baziw and Verbeek, 2012; Baziw and Weir-Jones, 2002). As previously outlined, the cone bearing measurement is an averaging operation where layers above and below the cone tip affect the measured tip

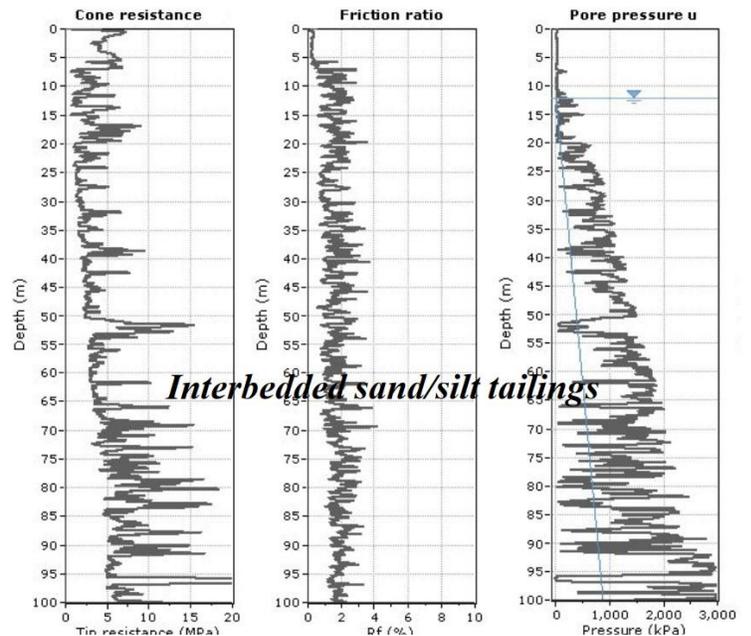


Figure 5. Example of anomalous/spurious cone bearing data (after Robertson, 2015).

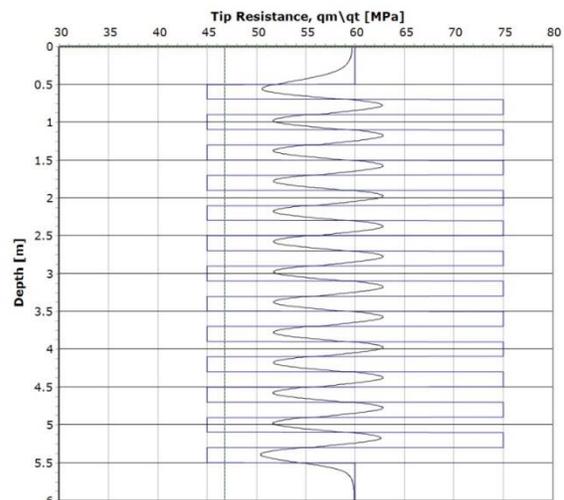


Figure 6. Example of smoothing/averaging highly variable (5MPa and 15MPa) interbedded soils. Blue series is the true cone bearing q_t and black series is smoothed q_m cone bearing.

resistance; therefore, sharp peaks and troughs should not be present and are considered measurement noise. For example, Fig. 6 illustrates a simulation of thin bed layering (0.2m) where there is alternating true q_t values of 45MPa and 75MPa (blue trace). As is shown in Fig. 6 there is a resulting oscillation averaged/smooth q_m trace (black trace) with no sharp peaks or troughs.

To date there has been minimal progress in removing the anomalous q_m data aside from *ad hoc* techniques which include discarding q_m measurements and smoothing/averaging (arithmetic and/or geometric) q_m measurements over a specific depth interval. The software program Settle3D developed by Rocscience (2016) describes a q_m filtering technique where q_m data spikes are removed. In this technique the cone bearing profile is divided into n sections, where $n = \text{depth}/(\text{window size})$ (default window size is set to 0.25m). The q_m mean (q_{m_mean}) and standard deviation (σ_i) are then calculated for each section. A bandwidth is W_{bi} is calculated for each section as outlined below and if the measured q_m value exceeds W_{bi} then the q_m value is removed.

$$\sigma_{ai} = \sqrt{(\sigma_{i-1}^2 + \sigma_i^2)} \quad (3a)$$

$$\sigma_{bi} = \sqrt{(\sigma_{i+1}^2 + \sigma_i^2)} \quad (3b)$$

$$W_{bi} = q_{m_mean} + BS \times \sigma_{ai} \text{ if } \sigma_{ai} < \sigma_{bi} \quad (3c)$$

$$W_{bi} = q_{m_mean} + BS \times \sigma_{bi} \text{ if } \sigma_{ai} > \sigma_{bi} \quad (3d)$$

BS is a filtering constant that has a default value of 1

In this paper the q_m measurement is mathematically modeled in a state-space formulation. The state-space formulation allows for the implementation of Bayesian recursive estimation techniques such as Kalman Filtering (KF). BRE techniques are very robust where multiple input measurements can be inputted into the estimation algorithm. The optimal q_m estimation algorithm where spurious data has been removed or minimized is referred to as the q_m KF algorithm.

2. MATHEMATICAL BACKGROUND

Bayesian Recursive Estimation

Bayesian Recursive Estimation (BRE) is a filtering technique based on state-space, time-domain formulations of physical problems (Arulampalam et al., 2002; Baziw, 2007). Application of this filter type requires that the dynamics of the system and measurement model, which relates the noisy measurements to the system state equations, be describable in a mathematical representation and probabilistic form that uniquely define the system behaviour. The potentially nonlinear discrete stochastic equation describing the system dynamics is defined as follows:

$$\mathbf{x}_k = \mathbf{f}_{k-1}(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}) \leftrightarrow p(\mathbf{x}_k | \mathbf{x}_{k-1}) \quad (4)$$

In eq. (4), the vector \mathbf{f}_k is a function of the state vector \mathbf{x}_k and the process or system noise \mathbf{u}_k . It is assumed that eq. (4) describes a Markov process of order one. The sampled potentially nonlinear measurement equation is given as

$$z_k = h_k(\mathbf{x}_k, \mathbf{v}_k) \leftrightarrow p(z_k | \mathbf{x}_k) \quad (5)$$

In eq. (5), \mathbf{h}_k depends upon the index k , the state \mathbf{x}_k , and the measurement noise \mathbf{v}_k at each sampling time. The probabilistic state-space formulation described by eq. (5) and the requirement for updating the state vector estimate based upon the newly available measurements described by eq. (5) are ideally suited for the Bayesian approach to derive the optimal estimation. In this approach it is attempted to construct the posterior estimate of the state given all available measurements. In general terms, it is desired to obtain estimates of the discretized system equation states \mathbf{x}_k based on all available measurements up to time k (denoted as $\mathbf{z}_{1:k}$) by constructing the posterior $p(\mathbf{x}_k | \mathbf{z}_{1:k})$. The posterior *Probability Density Function* (PDF) then allows the calculation of the conditional mean estimate of the state ($E[\mathbf{x}_k | \mathbf{z}_{1:k}]$).

BRE is a two step process consisting of prediction and update. In the prediction step the system equation defined by eq. (6) is used to obtain the prior PDF of the state at time k using the Chapman-Kolmogorov equation, which is given as

$$p(\mathbf{x}_k | \mathbf{z}_{1:k-1}) = \int p(\mathbf{x}_k | \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}) d\mathbf{x}_{k-1} \quad (6)$$

The update step then computes the posterior PDF from the predicted PDF and the newly available measurement as follows:

$$p(\mathbf{x}_k | \mathbf{z}_{1:k}) = \frac{p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{z}_{1:k-1})}{p(\mathbf{z}_k | \mathbf{z}_{1:k-1})} \quad (7)$$

The recurrence eqs. (6) and (7) form the basis for the optimal Bayesian solution. The BRE of the posterior density can generate an exact solution when the state-space equations fit into a *Kalman Filter* formulation or a *Hidden Markov Model*.

Kalman Filter

The Kalman Filter (KF) equations (Table 1) can be implemented as an optimal solution to the BRE when the following conditions are met:

- \mathbf{u}_k and \mathbf{v}_k are zero mean independent Gaussian white noise processes
- \mathbf{f}_k is a linear function of the state vector
- process noise, \mathbf{h}_k is a linear function of the state vector and measurement noise
- the initial estimate of \mathbf{x}_0 has a Gaussian distribution.

TABLE 1
KF Governing Equations

DESCRIPTION	Mathematical Representation	Eq
System equation	$\mathbf{x}_k = \mathbf{F}_{k-1} \mathbf{x}_{k-1} + \mathbf{G}_{k-1} \mathbf{u}_{k-1}$	8
Measurement equation	$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k$	9
State estimate extrapolation	$\hat{\mathbf{x}}_{k k-1} = \mathbf{F}_{k-1} \hat{\mathbf{x}}_{k-1 k-1}$	10
Error covariance extrapolation	$\mathbf{P}_{k k-1} = \mathbf{F}_{k-1} \mathbf{P}_{k-1 k-1} \mathbf{F}_{k-1}^T + \mathbf{G}_{k-1} \mathbf{Q}_{k-1 k-1} \mathbf{G}_{k-1}^T$	11
Measurement extrapolation	$\hat{\mathbf{z}}_k = \mathbf{H}_{k-1} \hat{\mathbf{x}}_{k k-1}$	12
Innovation	$\Delta_k = \mathbf{z}_k - \hat{\mathbf{z}}_k$	13
Variance of innovation	$\mathbf{S}_k = \mathbf{H}_k \mathbf{P}_{k k-1} \mathbf{H}_k^T + \mathbf{R}_k$	14
Kalman gain matrix	$\mathbf{K}_k = \mathbf{P}_{k k-1} \mathbf{H}_k (\mathbf{S}_k)^{-1}$	15
State estimate update	$\hat{\mathbf{x}}_{k k} = \hat{\mathbf{x}}_{k k-1} + \mathbf{K}_k \Delta_k$	16
Error covariance update	$\mathbf{P}_{k k} = [\mathbf{I} - \mathbf{K}_k \mathbf{H}_k] \mathbf{P}_{k k-1}$	17

In (8) and (9) \mathbf{v}_k and \mathbf{u}_k are *i.i.d* Gaussian zero mean white noise processes with variances of \mathbf{Q}_k and \mathbf{R}_k , respectively (i.e., $\mathbf{v}_k \sim N(0, \mathbf{R}_k)$ and $\mathbf{u}_k \sim N(0, \mathbf{Q}_k)$).

In Table 1 \mathbf{x}_k denotes the state to be estimated, \mathbf{F}_{k-1} denotes the state transition matrix which describes the system dynamics, \mathbf{u}_{k-1} the process or system noise (model uncertainty), \mathbf{G}_{k-1} describes the relationship between \mathbf{x}_k and \mathbf{u}_{k-1} , and \mathbf{H}_k the relationship between the state and the available measurement (measured cone resistance q_m). The KF can be applied to problems with linear time-varying systems and with non-stationary system and measurement statistics. Problems with nonlinearities are handled by linearizing the system and measurement equations. The Kalman Filter is readily applied to estimation, smoothing and prediction.

Figure 7 illustrates the essential relation between the system, the measurements and the Kalman Filter. Figure 7 indicates the scope of information the KF takes into account. As can be seen, the statistics of the measurement and state errors are essential components of the filter. The *a priori* information provides for optimal use of any number, combination and sequence of external measurements.

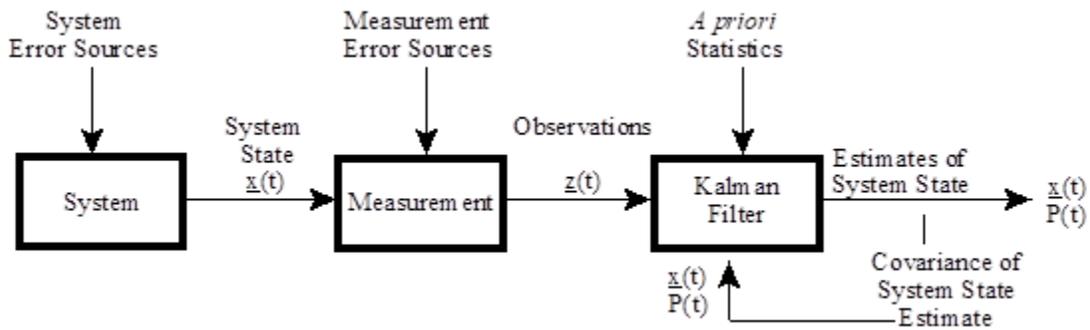


Figure 7. Block diagram of system, measurement, and Kalman Filter (Baziw and Weir-Jones, 2002).

The computational sequence for the discrete KF is outlined as follows:

- I. at time index $k = 0$, specify initial conditions x_0 and P_0 , and compute F_0 and $G_0 Q_0 G_0^T$
- II. at time index $k=1$, compute $\hat{x}_{1|0}$, $P_{1|0}$, H_1 , R_1 , and the gain matrix K_1
- III. using the measurement z_1 at time index $k=1$, the best estimate of the state at $k=1$ is given by

$$\hat{x}_{1|1} = \hat{x}_{1|0} + K_1 \Delta_1$$

$$\Delta_1 = z_1 - \hat{z}_1$$

$$\hat{z}_1 = H_1 \hat{x}_{1|0}$$

- IV. Update the error covariance matrix $P_{1|1}$

At time index $k=2$, a new measurement z_2 is obtained and the computational cycle is repeated.

Gauss-Markov Model

In the subsequently outlined $q_m KF$ algorithm the state-space mathematical model for the cone bearing measurements q_m is described kinematically via the position, velocity and acceleration of

the averaged/smoothed cone bearing q_m measurement. This is similar to marine navigation (Baziw, 1994 and 1996) where multiple measurements are inputted into a kinematics KF and optimal estimates of the vessel's real time position are made based upon the dynamics of the vessel and the available measurements. In the q_mKF algorithm the measured cone bearing acceleration is modelled as a Gauss-Markov process. The Gauss-Markov (Gelb, 1974; Lear, 1985) process has a relatively simple mathematical description.

As in the case of all stationary Gaussian processes, specification of the process autocorrelation completely defines the process. The variance, σ^2 , and time constant, T_c (ie., $\beta = 1/T_c$), define the first-order Gauss-Markov process. The discrete mathematical equation for a Gauss-Markov process is given as

$$\begin{aligned} n_{k+1} &= a_w n_k + b_w w_k \\ a_w &= e^{-\beta\Delta} \text{ and } b_w = \sigma \sqrt{1 - e^{-2\beta\Delta}} \end{aligned} \quad (18)$$

In eq. (18), Δ is the sampling rate and w_k is a zero-mean, timewise-uncorrelated, unit-variance sequence with a Gaussian probability distribution function. n_k is therefore a zero-mean, exponentially-correlated random variable whose standard deviation is σ . The constant a_w can have a range of values from -1 to +1. For a stable variable, a_w is restricted to values between 0 and +1. For $a_w \rightarrow 0$, n_k changes rapidly and tends to be uncorrelated from sample to sample. For $a_w \rightarrow 1$, the behavior of n_k becomes more sluggish and it tends to change little from sample to sample. The time constant term (T_c where $a_w = e^{-\Delta/T_c}$) of the q_m acceleration model is a very robust parameter within the q_mKF algorithm. Preferably the q_m acceleration model results in a smooth trajectory of the cone bearing while at the same time allowing for sufficient maneuverability. A value of $a_w = 0.999$ was found to work well for modeling a variety of q_m profiles. The estimate of the variance of the acceleration is provided by a polynomial best fit as is subsequently described.

3 q_m KF ALGORITHM OUTLINE

3.1 State-Space Formulation and Measurement Model

System Model

To specify the systems equations in the standard KF state-space form, the following states need to be defined

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \equiv \begin{bmatrix} q_m \text{ cone bearing position} \\ q_m \text{ cone bearing velocity} \\ q_m \text{ cone bearing acceleration} \end{bmatrix} \quad (19)$$

The discrete system equation (eq. (8)) is given as

$$\begin{bmatrix} x_{1k+1} \\ x_{2k+1} \\ x_{3k+1} \end{bmatrix} = \begin{bmatrix} 1 & \Delta & \Delta^2/2 \\ 0 & 1 & \Delta \\ 0 & 0 & a_w \end{bmatrix} \begin{bmatrix} x_{1k} \\ x_{2k} \\ x_{3k} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ b_w \end{bmatrix} w_k \quad (20)$$

In eq. (19) where Δ is the sampling rate and a_w , b_w and w_k are defined in eq. (18).

The discrete covariance structure (eq. 11)

$$G_{k-1} Q_{k-1|k-1} G_{k-1}^T = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & b_w^2 \end{bmatrix} \quad (21)$$

Measurement Model

Currently two synthesized measurements are incorporated into the q_m KF algorithm: 1) The best fit seventh degree polynomial to the q_m profile and 2) Output after applying an fourth order low pass frequency filter to the q_m profile. At a later date it is envisioned that additional measurements could be incorporated into the q_m KF algorithm such as the output from the vane shear strength. In this case the estimated vane shear test undrained strength estimate is used to calculate q_m measurements from established empirical correlations.

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \equiv \begin{bmatrix} \text{seventh order polynial best fit} \\ \text{output after applying low pass filter} \end{bmatrix} \quad (22)$$

The linear measurement matrix, H, is given by the following equation

$$H_k = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \quad (23)$$

7th Degree Polynomial Measurement Model

A best fit 7th degree polynomial is made to the q_m measurements every 1.4 m depth increment so that the ‘‘spurious’’ noise is minimized. This polynomial is then fed into the q_m KF algorithm as a measurement. The order of the polynomial and depth increment were selected due to the averaging/blurring of the q_m measurements where it would be highly unlikely that there would be

greater than 6 turnings¹ in a 1.4m depth increment. This assumption was tested with extensive test bed simulation. The estimation of the series of polynomials over the complete q_m profile is done utilizing a KF.

In terms of kinematics, the position (z_k), velocity (\dot{z}_k), and acceleration (\ddot{z}_k) of a 7th degree polynomial is given as

$$z_k = a_0 + a_1d + a_2d^2 + a_3d^3 + a_4d^4 + a_5d^5 + a_6d^6 + a_7d^7 \quad (24a)$$

$$\dot{z}_k = a_1 + 2a_2d + 3a_3d^2 + 4a_4d^3 + 5a_5d^4 + 6a_6d^5 + 7a_7d^6 \quad (24b)$$

$$\ddot{z}_k = 2a_2 + 6a_3d + 12a_4d^2 + 20a_5d^3 + 30a_6d^4 + 42a_7d^5 \quad (24c)$$

where $d = k\Delta$, Δ is the sampling rate

The following states need to be defined for estimating the 7th degree polynomial coefficients with a KF

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \end{bmatrix} \equiv \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \end{bmatrix} \quad (25)$$

The discrete system equation (eq. (8)) is given as

$$\begin{bmatrix} x_{1k+1} \\ x_{2k+1} \\ x_{3k+1} \\ x_{4k+1} \\ x_{5k+1} \\ x_{6k+1} \\ x_{7k+1} \\ x_{8k+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{1k} \\ x_{2k} \\ x_{3k} \\ x_{4k} \\ x_{5k} \\ x_{6k} \\ x_{7k} \\ x_{8k} \end{bmatrix} \quad (26)$$

The linear measurement matrix, H , is given by the following equation

$$H_k = [1 \quad d \quad d^2 \quad d^3 \quad d^4 \quad d^5 \quad d^6 \quad d^7] \quad (27)$$

As previously outlined, the best fit 7th degree polynomial is made to the q_m measurements every 1.4m depth increment (i.e., $N = (1.4/\Delta)$, where N is the total number of points to be processed at each depth increment). The value of the variance of the polynomial acceleration at each depth increment is calculated using eq. (24c). The maximum value is then used within eq. (18). For continuity in the polynomial estimated trace, the last estimated value of the polynomial “position” (z_N) is set to coefficient a_0 and the last estimated value of the polynomial “velocity”

¹ The maximum number of turnings of a polynomial function is always one less than the degree

(z'_N), is set to coefficient a_1 in the subsequent depth increment 7th degree polynomial KF estimation.

4th Order Low Pass Filter

A 4th order 250Hz Butterworth (low pass frequency filter is applied to the q_m measurements measurement so that the “spurious” noise is minimized. This 250Hz low passed frequency filtered trace is then fed into the q_m KF algorithm as a measurement. The Butterworth filter (Kanasewich, 1981) is a common form of a low-pass filter, and it can be defined by

$$|G(\omega)|^2 = 1/\{1 + (\omega/\omega_0)^{2N}\} \quad (28)$$

where ω_0 is the “cutoff” frequency (250Hz in this case) and N determines the sharpness of the cutoff.

The advantages associated with the Butterworth filters are as follows:

- Their transfer functions are smooth and maximally flat both inside and outside the passband.
- The squared filter (i.e., the input is filtered twice so that the amplitude responses is $|G(\omega)|^2$) produced zero phase shift and its power is down 3dB (factor of 1/2) at the cutoff frequency².

The 250Hz cutoff frequency was selected due to the averaging/blurring of the q_m measurements where it would be highly unlikely that there would be frequencies greater than 250Hz. This assumption was tested with extensive test bed simulation.

q_m KF Test Bed Examples:

The performance of the q_m KF algorithm was evaluated by carrying out challenging test bed simulations. This section outlines two of these challenging test bed simulations.

Table 2 and Figure 8 outline and illustrate, respectively, the simulated cone bearing profile where the background q_t values linearly increase from 10 MPa to 12 MPa to a depth of 20 m. A 1cm sampling interval was applied in the simulation.

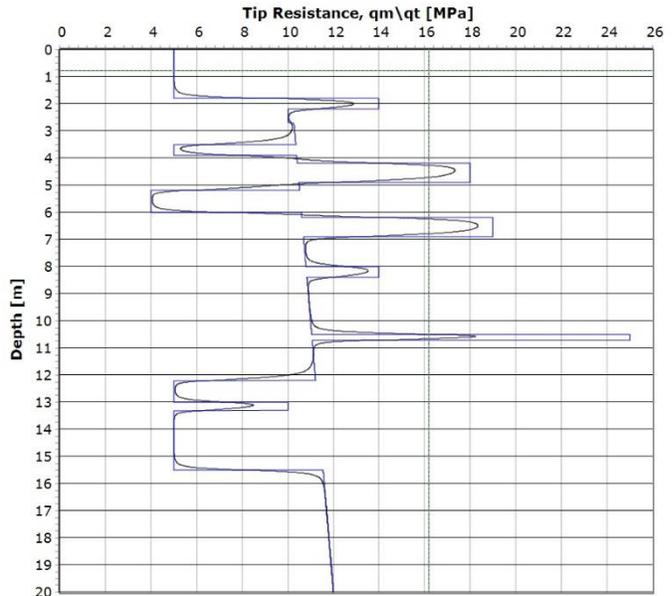


Figure 8. Simulated true cone bearing measurements q_t (red trace) and corresponding averaged/blurred q_m (black trace) measurements.

²The cutoff frequency determines the half-power point of the filter.

Figure 9 illustrates the simulated q_m data of Figure 8 (black) with additive noise to represent anomalous/spurious q_m data (grey trace). The Gauss-Markov simulated noise has a variance of 9 and a time constant of 0.3. In addition, the Gauss-Markov noise had a high pass frequency of 250 Hz applied so that realistic spurious data was generated. The red trace is the q_mKF algorithm estimated q_m profile after processing the spurious q_m data (grey trace). Figure 10 shows the traces illustrated in Fig. 9 but in this case an industry standard four point smoother was applied to the spurious q_m trace (grey trace).

Figure 11 illustrates the simulated q_m data of Figure 8 (black) with additive noise to represent anomalous/spurious q_m data (grey trace). The Gauss-Markov simulated noise has a variance of 50 and a time constant of 0.7. The red trace is the q_mKF algorithm estimated q_m profile after processing the spurious q_m data (grey trace). Figure 12 shows the traces illustrated in Fig. 11 but in this case an industry standard four point smoother was applied to the spurious q_m trace (grey trace).

Table 2. Simulated q_t values

Depth 1	Depth 2	q_t [MPa]
1.0	1.8	5
1.8	2.2	14
2.2	2.7	10
3.5	3.9	5
4.2	4.9	18
5.2	6.0	4
6.2	6.9	19
8.0	8.4	14
10.5	10.7	25
12.2	15.5	5
13.0	13.3	10

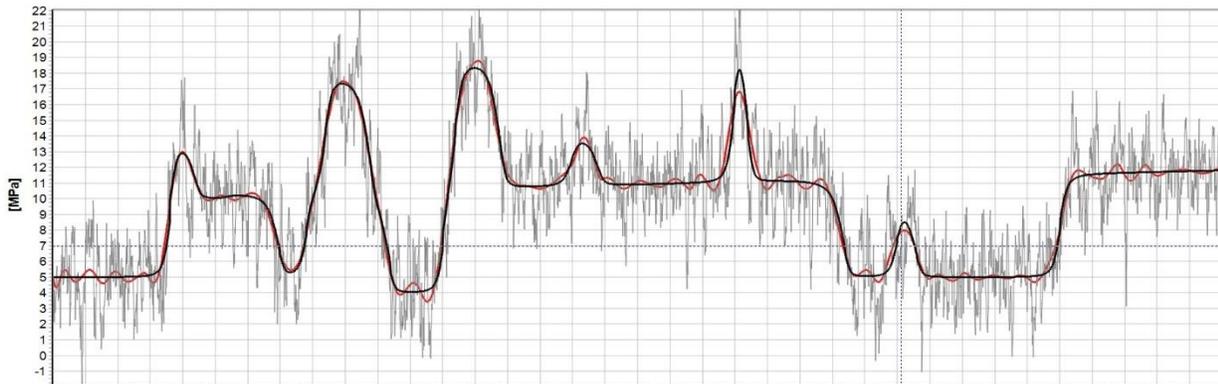


Figure 9. Simulated cone bearing averaged/blurred q_m (black trace) of Fig. 8, spurious q_m trace (grey trace $\sigma^2 = 9$ and $T_c = 0.3$) feed into the q_mKF algorithm, and the q_mKF algorithm

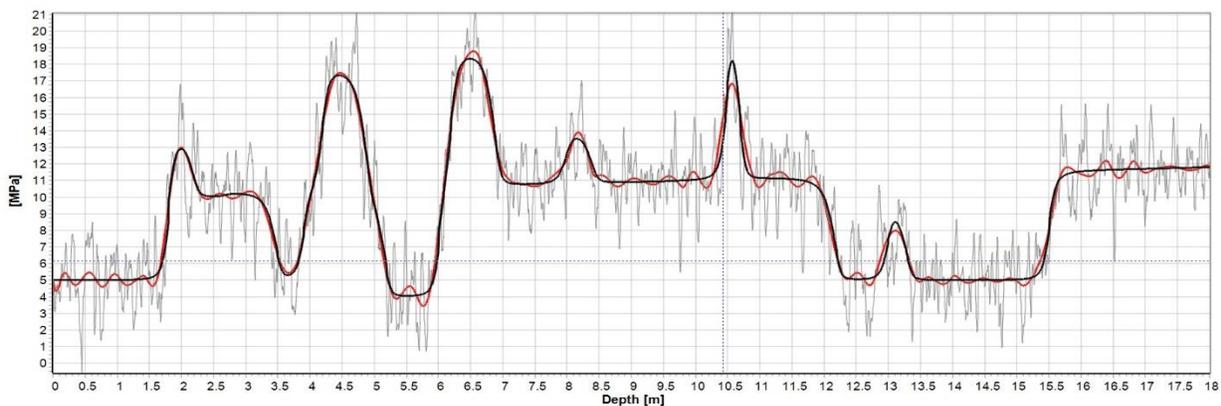


Figure 10. The traces illustrated in Fig. 9 but in this case industry standard four point smoothing/averaging was applied to the spurious q_m trace (grey trace).

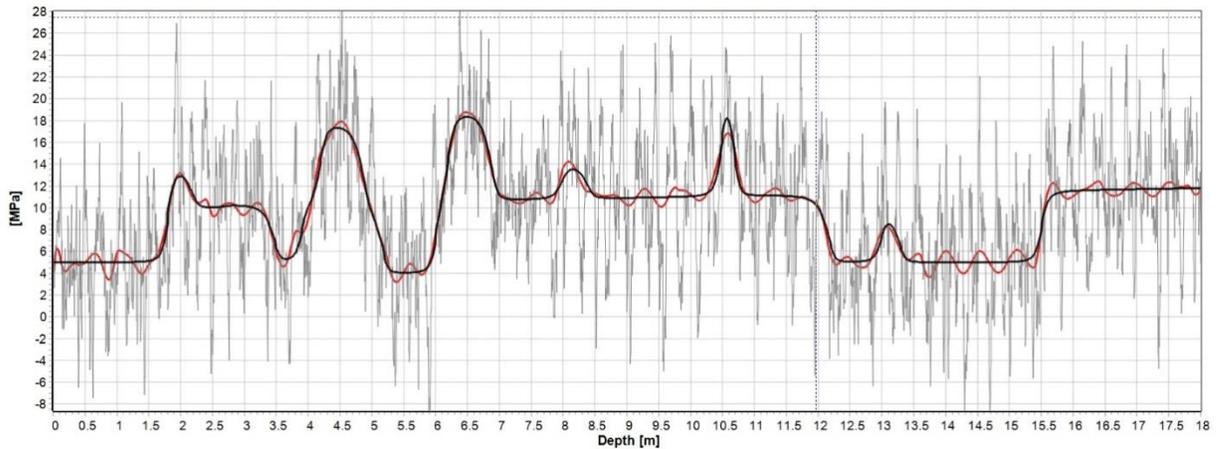


Figure 11. Simulated cone bearing averaged/blurred q_m (black trace) of Fig. 8, spurious q_m trace (grey trace $\sigma^2 = 50$ and $T_c = 0.7$) feed into the q_mKF algorithm, and the q_mKF

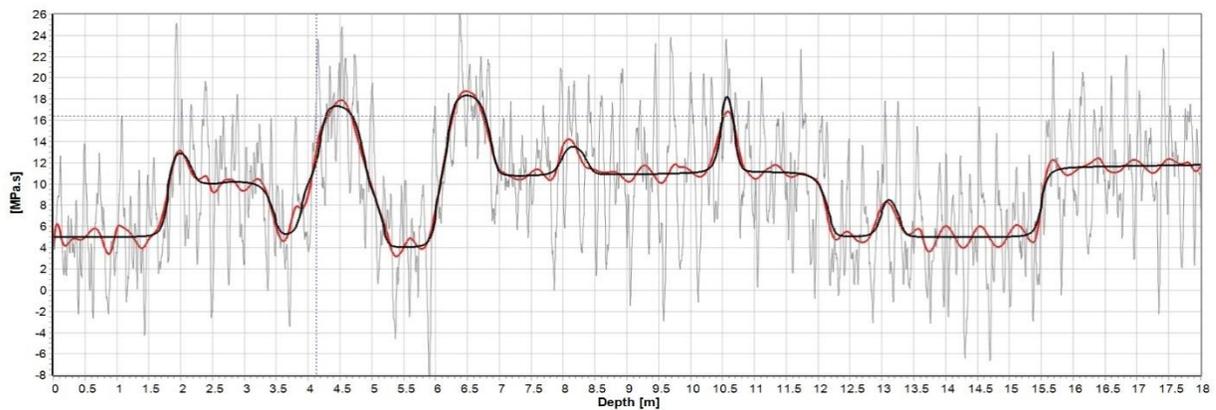


Figure 12. The traces illustrated in Fig. 11 but in this case industry standard four point smoothing\averaging was applied to the spurious q_m trace (grey trace).

The second test bed simulation addresses the issue first described in Figure 6: a simulation of thin bed layering (0.2m) where there is alternating true q_t values of 45 MPa and 75 MPa (blue trace). As was shown in Fig. 6 there is a resulting oscillation averaged/smooth q_m trace (black trace) with no sharp peaks or troughs. Figure 13 illustrates the simulated q_m data of Fig. 6 (black) with additive noise ($\sigma^2 = 70$ and $T_c = 0.2$) to represent anomalous/spurious q_m data (grey trace). The red trace is the q_mKF algorithm estimated q_m profile after processing the spurious q_m data (grey trace). Figure 14 shows the traces illustrated in Fig. 13 but in this case an industry standard four point smoother was applied to the spurious q_m trace (grey trace). Figure 15 illustrates the simulated q_m data of Fig. 6 (black) with additive noise ($\sigma^2 = 300$ and $T_c = 0.3$) to represent anomalous\spurious q_m data (grey trace). The red trace is the q_mKF algorithm estimated q_m profile after processing the spurious q_m data (grey trace). Figure 16 shows the traces illustrated in Fig. 15 but in this case an industry standard four point smoother was applied to the spurious q_m trace (grey trace).

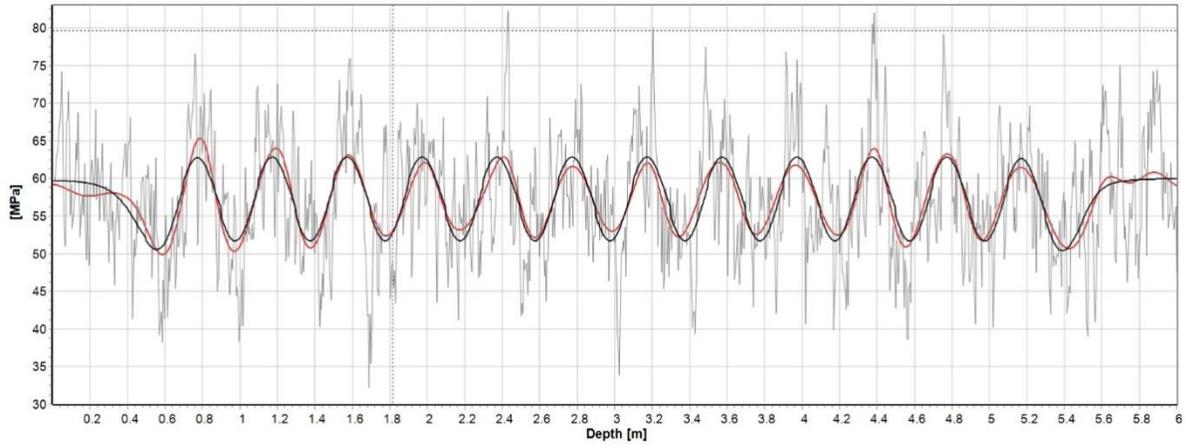


Figure 13. Simulated cone bearing averaged/blurred q_m (black trace) of Fig. 6, spurious q_m trace ($\sigma^2 = 70$ and $T_c = 0.2$) feed into the q_mKF algorithm, and the q_mKF

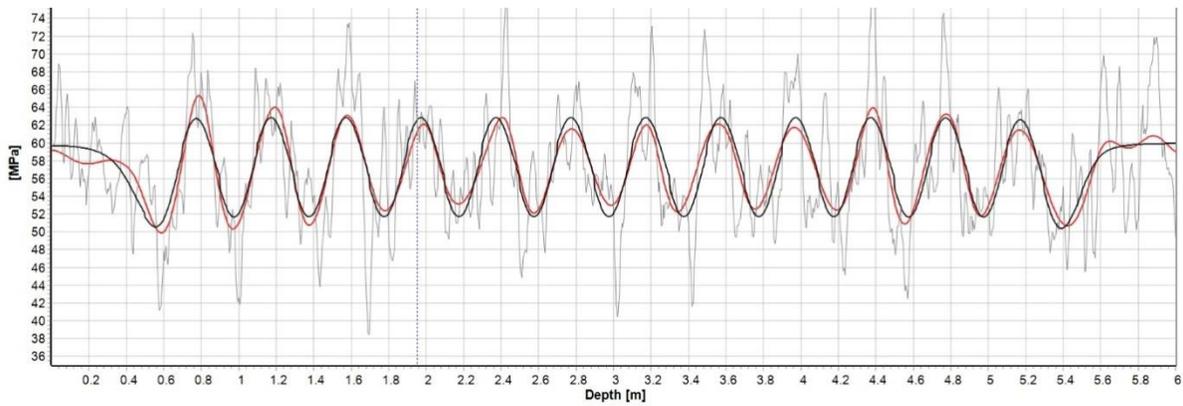


Figure 14. The traces illustrated in Fig. 13 but in this case industry standard four point smoothing/averaging was applied to the spurious q_m trace (grey trace).

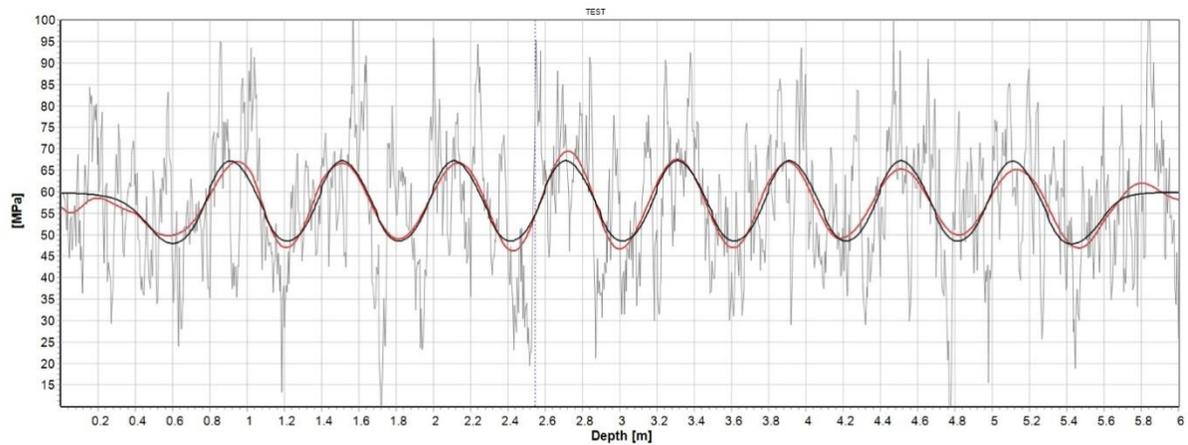


Figure 15. Simulated cone bearing averaged/blurred q_m (black trace) of Fig. 6, spurious q_m trace (grey trace $\sigma^2 = 300$ and $T_c = 0.3$) feed into the q_mKF algorithm, and the q_mKF algorithm output (red trace).

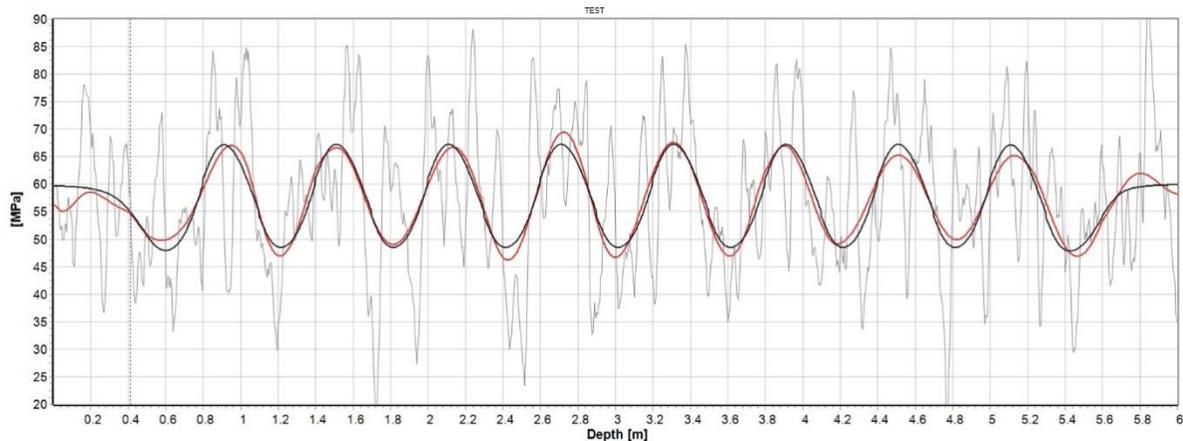


Figure 16. The traces illustrated in Fig. 15 but in this case industry standard four point smoothing\averaging was applied to the spurious q_m trace (grey trace).

CONCLUSIONS

In Cone penetration testing (CPT) the cone bearing measurements q_m are highly susceptible to anomalous peaks and troughs due to the relatively small diameter cone tip penetrating sandy, silty and gravelly soils. The high peaks are due to interbedded gravels and stones and low peaks are due to softer materials or local pore pressure build-up. The cone bearing measurement is an averaging operation where layers above and below the cone tip affect the measured tip resistance; therefore, sharp peaks and troughs should not be present and are considered as measurement noise. To date there has been minimal progress in removing the anomalous q_m data aside from *ad hoc* techniques which include discarding q_m measurements and smoothing/averaging q_m measurements over a specific depth interval. This paper has outlined a Kalman Filter (KF) algorithm (q_mKF) which optimally obtains estimates of q_m with the spurious data removed or minimized. The implementation and performance of the q_mKF algorithm was demonstrated by considering challenging test bed examples. The authors will carry out further test bed simulations and subsequently apply the q_mKF 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. 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.
- Baziw, E. and Verbeek, G. (2012), "Passive (Micro-) Seismic Event Detection by Identifying Embedded "event" Anomalies within Statistically Describable Background Noise", Pure appl. geophys., vol. 169, Issue 12, 2107-2126.

- 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 Weir Jones, I (2002), "Application of Kalman Filtering Techniques for Microseismic Event Detection", *Pure appl. geophys.*, vol. 159, pp. 449-471.
- Baziw, E. (1994), "Dead Reckoning/Target Tracking in an ECDIS Environment," Proceedings of the 1994 National Technical Meeting of The Institute of Navigation, San Diego, CA, January 1994, 747-756.
- Baziw, E. (1996), "Field Trials of ECPINS Vessel Positioning Algorithm", In Proceedings of the IEEE Position Location and Navigation Symposium (PLANS), Atlanta, Georgia, 22-26 April. IEEE Publication, 121-129.
- 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 CPT methods applied to 102 case histories", *Canadian Geotechnical Journal (CGJ)*, 6, 886-904.
- Robertson, P.K. (2015), "Introduction to Cone Penetration Testing", ISSMGE webinar <https://www.issmge.org/education/recorded-webinars/introduction-to-cone-penetration-testing>
- Gelb, A. (1974). *Applied Optimal Estimation* (4th Edition). Cambridge, Mass: MIT Press.
- Kanasewich, E.R. (1981), *Time sequence analysis in geophysics*, 3rd edition, The University of Alberta Press, Edmonton, AB. 252-277.
- Lear, W.M. (1985), *Kalman Filtering Techniques*, National Aeronautics and Space Administration Publication, Mission Planning and Analysis Division, No. JSC-20688, 172-176.
- Lunne, T., Robertson, P.K., and Powell, J.J.M. (1997). *Cone penetrating testing: in geotechnical practice*. Taylor & Francis, 1997.
- Mortensen J.K., Hansen G., Sorensen B. (1991). Correlation of CPT and field vane test for clay fills. Danish Geotechnical Society, Bulletin No. 7.
- Rocscience User's Manual (2016), "CPT Data Interpretation Theory Manual" <https://www.rocscience.com/software/settle3>