A procedure for determining the characteristic value of a geotechnical parameter

A. J. Bond
Geocentrix Ltd, Banstead, Surrey, United Kingdom

ABSTRACT: Practising engineers are shown to be poor at predicting the appropriate degree of caution needed to select the 'characteristic' value of a geotechnical parameter, as defined by Eurocode 7. The paper presents a procedure for determining this characteristic value, based on simple statistical methods provided in readily available business software. The procedure is illustrated with data from two sites, obtained using cone penetration and standard penetration tests. Limitations of the procedure are discussed.

Keywords: characteristic value; statistics; Eurocode 7; worked examples

1 INTRODUCTION

Eurocode 7 defines the characteristic value of a geotechnical parameter as 'a cautious estimate of the value affecting the occurrence of the limit state'. For limit states that depend on the strength of the ground (typical of many ultimate limit states), it is the mean strength mobilized along the failure surface that is required – Eurocode 7 suggests that this should be selected with 95% confidence.

The author has conducted a series of experiments in which practising engineers have been asked to choose the characteristic value of various geotechnical parameters that vary with depth. The results reveal a very wide range of interpretations of the data – and that this variation increases as the data becomes more 'noisy'. The gap between the uppermost and lowermost interpretations is large enough to lead to significantly different design outcomes.

To help reduce this variation in interpretation, this paper proposes a simple procedure that engineers could follow to produce a repeatable characteristic value that is consistent with its definition in Eurocode 7. In outline, it involves the following steps:

1) Using readily-available statistical tools (such as those available in Microsoft Excel), determine the best-fit line through the data taking account of its variation with depth
2) Determine the residual (or fitting error) of each data point
3) Calculate the standard deviation of the residuals, s_X
4) Determine the appropriate degree of caution needed to establish a 95% confident mean value (using Student's statistical coefficient k_n, which takes account of the number of data points available)
5) Plot the resulting characteristic line parallel to the best fit line, using the expression $X_k = X_{mean} - k_n s_X$, where k_n is the statistical coefficient from Step 4 and s_X the standard deviation from Step 3

The paper illustrates this procedure with the results from a number of sites where the geotechnical parameter varies linearly with depth and compares the outcome with more rigorous statistical analysis. A variation on the procedure is outlined for situations where the geotechnical parameter does not vary linearly with depth.

2 DEFINITION OF THE CHARACTERISTIC VALUE

Eurocode 7 defines the *characteristic value* of a geotechnical parameter as:

a cautious estimate of the value affecting the occurrence of the limit state

Since the volume of ground that controls the occurrence of a limit state is usually much larger than a test sample, Eurocode 7 further states says that the characteristic value should be selected as:

a cautious estimate of … the mean of a range of values covering a large surface or volume of the ground

In structural engineering, the resistance of the structure usually depends on the strength of an individual structural element. For example, the resistance of a concrete beam is limited by the strength of the concrete across its weakest section. The strength of the concrete across this section does not vary greatly, although it might differ across different sections. In the Eurocode system, the characteristic value in this case is selected as the 5% fractile, i.e. a value that will be exceeded in 95% of all tests.

A key aspect of geotechnical engineering, which is alluded to in thes second quote above, is that the resistance of a foundation usually depends on the strength of the continuum, not just an element of the ground. For example, the bearing resistance of a footing on clay is limited by the undrained strength of the soil along the external and internal boundaries of the failure mechanism. Any variation in strength of the clay along those boundaries is 'averaged out' over the whole mechanism. The characteristic value is a cautious estimate of that average value. In the Eurocode system, the characteristic value in this case is selected as the 50% fractile (i.e. the mean value) at the 95% confidence level.

3 ENGINEERS' ASSESSMENT OF CHARACTERISTIC VALUES

The task of assessing 'a cautious estimate' of a geotechnical parameter is not an easy one. This is most vividly demonstrated by comparing the estimates made by more than one hundred engineers who were asked to assess the characteristic value of various parameters from typical site investigation data.

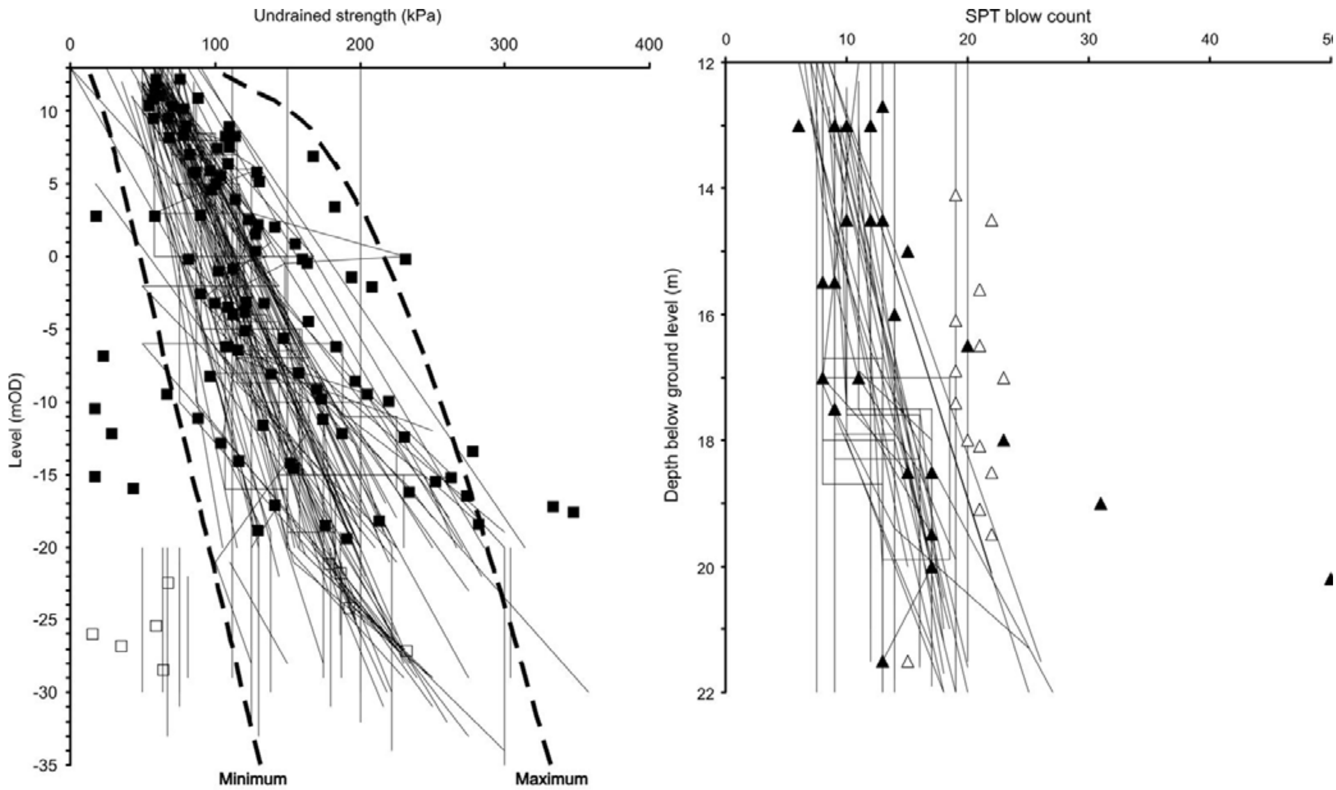

Figure 1, engineers' interpretation of the characteristic value of (left) the undrained strength of London and Lambeth clays from results of triaxial compression tests; (right) SPT blow-count in Thames Gravels

Figure 1 (from Bond and Harris, 2008) shows, on the left, the results of triaxial compression tests on London and Lambeth clays and, on the right, blow-counts measured in standard penetration tests in Thames Gravels. The symbols on these graphs represent individual data points and the superimposed lines are engineers' assessment of the characteristic value based on this data (alone). The scatter in the data is not at all unusual in these materials; the spread of the lines, however, is worrying, since it indicates little agreement between different engineers regarding the most appropriate value to select as 'characteristic'. Bond and Harris concluded that "engineers are not particularly good at selecting a cautious estimate of the characteristic value, particularly when the available data is scattered. Statistical treatment of large data sets … may help to guide engineers in this task".

The remainder of this paper presents a simple procedure that can help in this task.

4 PROCEDURE FOR DETERMINING THE CHARACTERISTIC VALUE

4.1 *Step 1 – determine the best-fit line through the data*

The first step in the proposed procedure is to establish the best-fit line through the data, taking account any trend for the data values to increase or decrease with depth below ground surface.

For example, consider the results of four cone penetration tests (CPTs) conducted in dense sand, as shown in **Figure 2** (left). The data is taken from ETC 10 Design Example 2.1 (ETC 10, 2009) and will be used in this paper to illustrate the procedure for establishing 'the' characteristic value. As can be seen from **Figure 2**, there is a marked tendency for the measured cone resistance to increase with depth, as is commonly the case in dense sand. The water table at this site is located at 6 m below ground level.

A trend line through this data can be obtained using simple linear regression, using (for example) Microsoft Excel's 'Linear Trendline' feature. For one of the cone tests (CPT3), this produces the best-fit line shown in **Figure 2** (right) given – with a coefficient of determination $R^2 = 0.7521$ – by the equation:

$$y = 0.487x - 4.66 \qquad (1)$$

where y = depth below ground surface (z, in m) and x = measured cone resistance (q_c, in MPa). On rearranging the coefficients, we get the more useful expression:

$$q_c = 9.57 + 2.05z \qquad (2)$$

which is illustrated by the dashed line on **Figure 2** (right). Although this trend-line is a reasonable fit to the data down to about 7 m, it overestimates the cone resistance below that depth.

The 'coefficient of determination' R^2 is the square of Pearson's correlation coefficient R and is an imperfect measure of the trendline's 'goodness of fit'. See Wikipedia for a simple and easily-accessed explanation.

It is worth noting that Excel's 'Linear Trendline' feature (which one of its Chart Tools) does not always give reliable answers. In processing the results for CP4 for this paper, the predicted trendline was seriously in error and differed significantly from that produced by Excel's alternative Regression tool (in its Data Analysis pack) and other statistical software. This error became apparent when comparing the predicted trendline with the raw data. This reiterates the importance of *looking* at the data, not just processing the numbers!

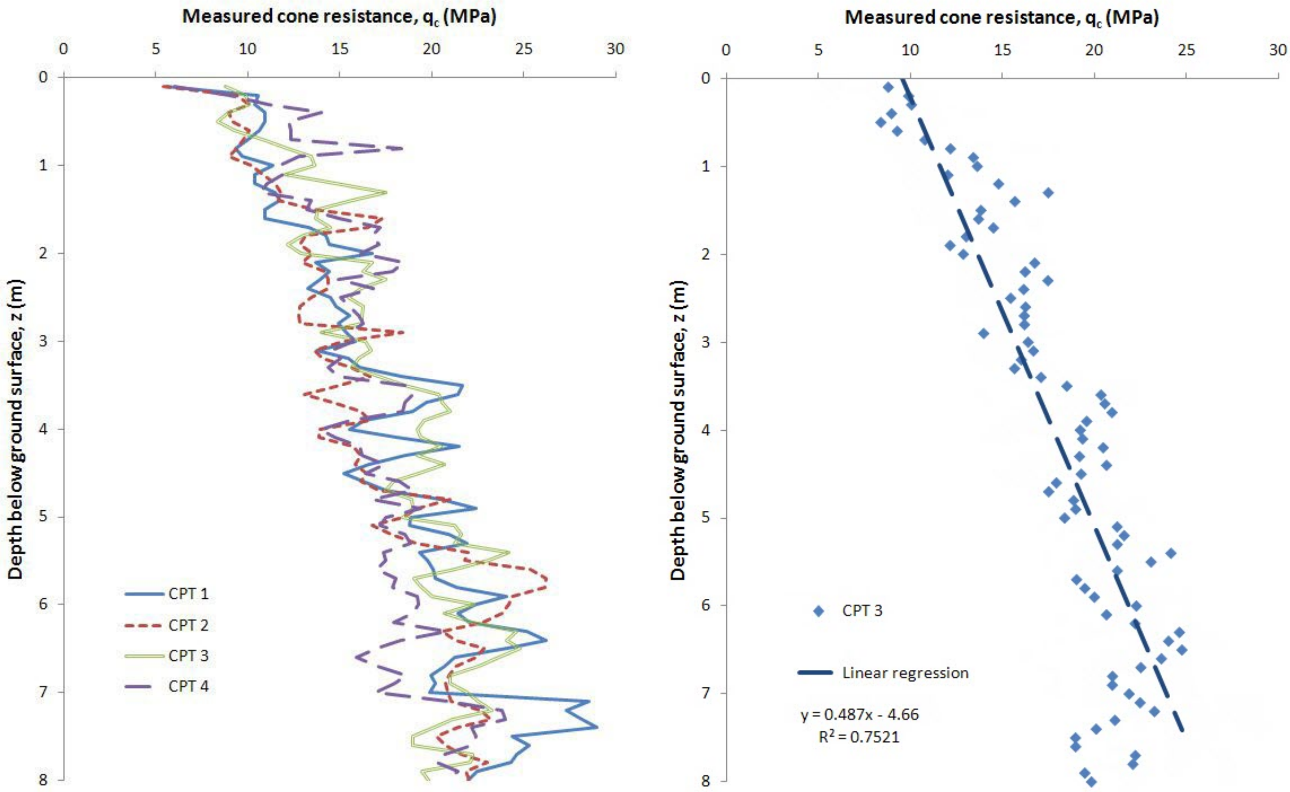

Figure 2, (left) measured cone resistance vs depth from four CPTs; (right) best-fit linear regression through data from CPT3

4.2 *Step 2 – determine the residual (or fitting error) of each data point*

The second step in the procedure is to determine the difference between each data value and that predicted by the best-fit line – in other words, the horizontal separation of each data point from the trend line shown in **Figure 2** (right). These 'residuals' are plotted (to an enlarged scale) versus depth in **Figure 3** (left) and as a histogram in **Figure 3** (right).

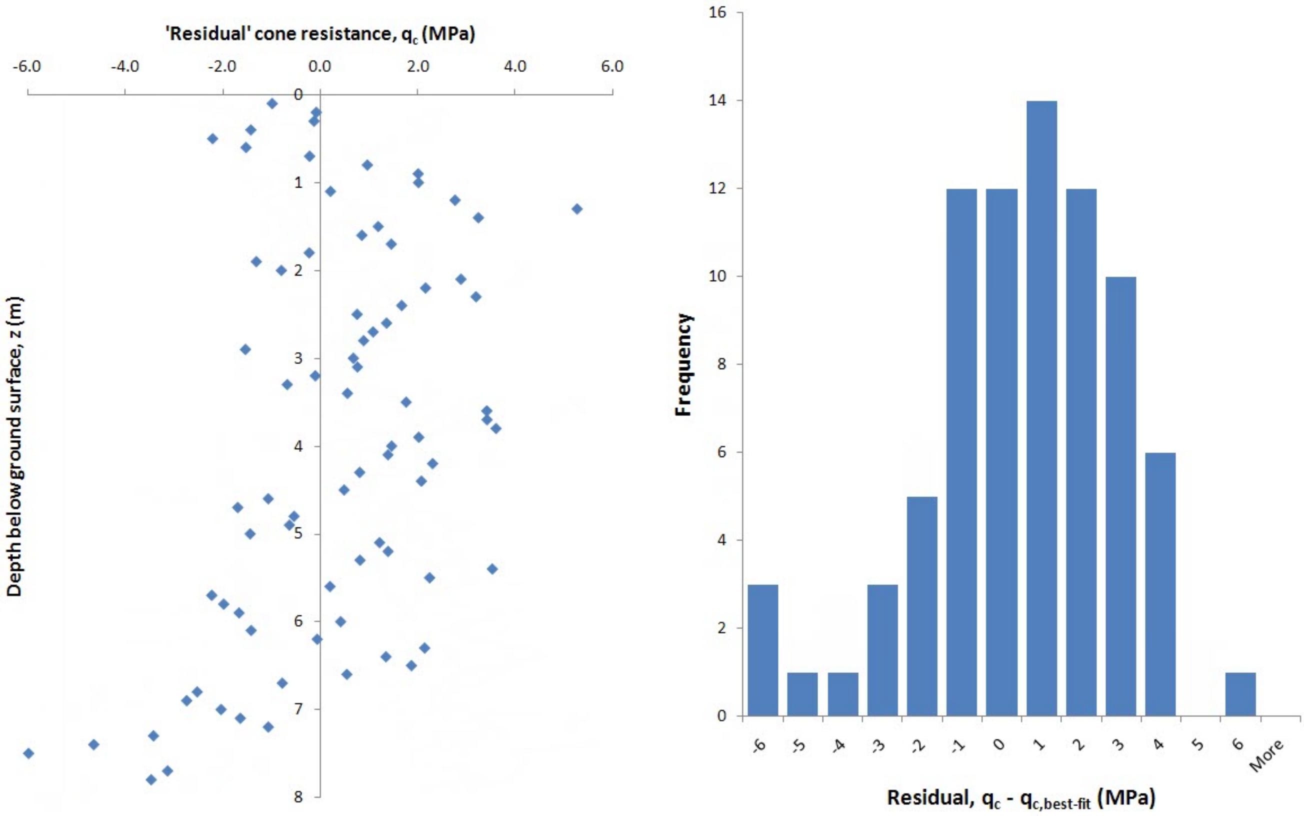

Figure 3, (left) residuals calculated for CPT3; (right) histogram of the same residuals

It is important to plot the residuals at this stage in the procedure, in order to detect any skew in the trend line. In this example, there is a marked tendency for the residuals to become negative below 7 m, suggesting that the trend line overestimates cone resistance below this depth (**Figure 3**, left). The histogram (**Figure 3**, right) makes this even more apparent, indicating that the data set as a whole, while broadly following a normal distribution (i.e. a bell-shaped curve) about a zero value, clearly is not homogenous.

Although not shown here, similar departures from a strictly linear trendline are obtained for CPTs 2 and 4. These results necessitate re-evaluation of the chosen trendline and – although there are several techniques that could be employed to obtain a better trend – for simplicity here I am going to ignore all data below the water table at 6 m. Hence the prediction made from now on will apply solely to the dry sand.

Repeating the procedure followed thus far, but on the reduced data set for CPT3, gives the revised residuals and corresponding histogram shown in **Figure 4** and a trendline expressed by:

$$q_c = 9.05 + 2.47z \quad (3)$$

where z = depth below ground surface (in m) and q_c is cone resistance (in MPa).

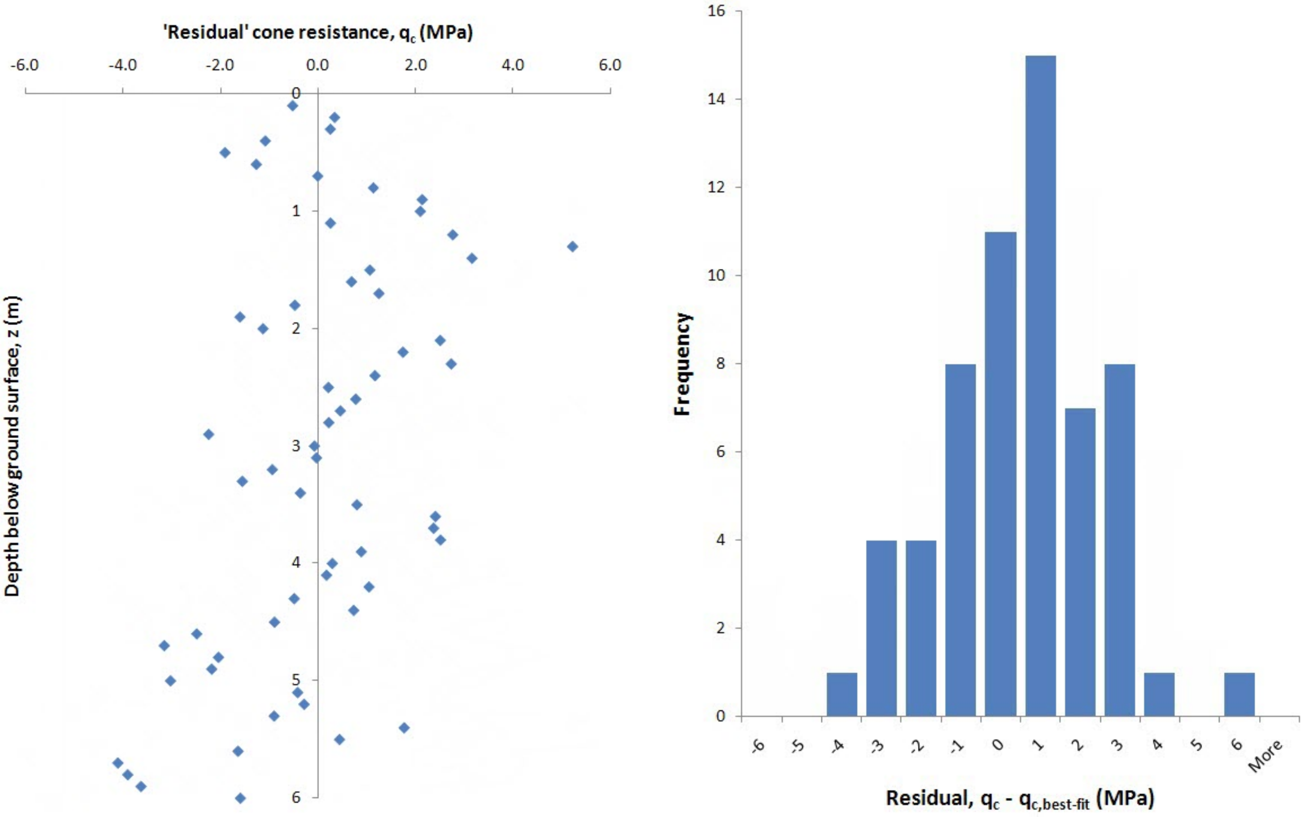

Figure 4, (left) revised residuals calculated for CPT3; (right) histogram of the same revised residuals

The residuals appear more evenly scattered about the zero line and the histogram – although not perfectly following the expected bell-shaped curve – nevertheless gives a much improved fit (cf **Figure 3**, right).

4.3 *Step 3 – calculate the standard deviation of the residuals*

The next step in the procedure is to calculate the standard deviation s_X of the residuals, assuming (in this case) a normal distribution about zero. This can once again be achieved readily using Microsoft Excel's STDEV() function applied to the residuals. For the data from CPT3, this gives s_X = 1.88 MPa.

4.4 *Step 4 – determine the appropriate degree of caution*

Eurocode 7 requires the characteristic value of a spatially-averaged parameter to be selected as a 95% confident mean value. As explained by Bond and Harris (2008, §5.5.2), the lower (or 'inferior') characteristic value $X_{k,inf}$ of a geotechnical parameter X is given by:

$$X_{k,\inf} = m_X - k_n s_X \tag{4}$$

where m_X is that parameter's mean value (i.e. as predicted by the trend line from Step 1); s_X is the standard deviation calculated in Step 3; and k_n is a statistical coefficient that depends on number of data points available, n. For cases where the standard deviation is not known *a priori*, this statistical coefficient is given by:

$$k_n = t_{n-1}^{95\%} \times \sqrt{1/n} \tag{5}$$

where $t_{n-1}^{95\%}$ is Student's t-value for $(n - 1)$ degrees of freedom at a confidence level of 95%, as shown in **Figure 5** (from Bond and Harris, 2008).

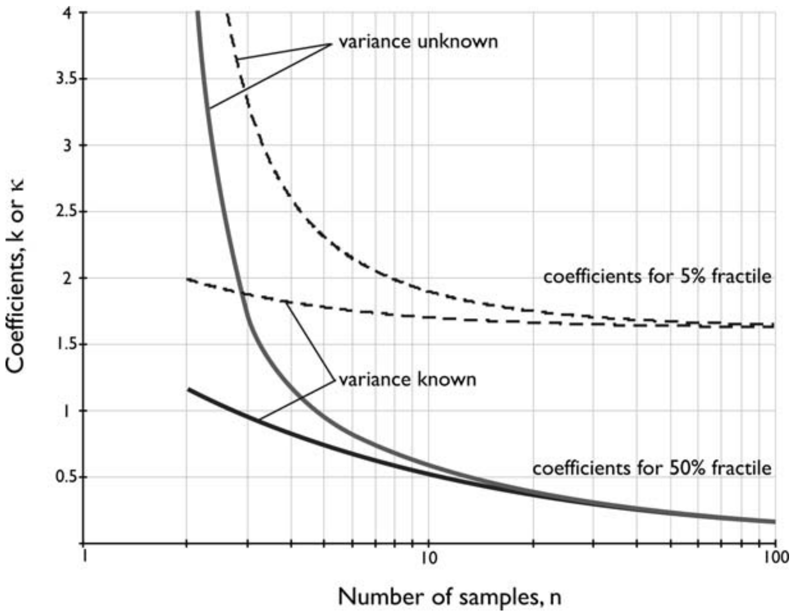

Figure 5, statistical coefficients for determining the 5% and 50% fractiles with 95% confidence

Hence, to determine the appropriate degree of caution, we need to look up the value of k_n from **Figure 5** for the 50% fractile, with variance unknown. The 50% fractile is chosen because we are seeking the *mean* value of cone resistance; the 'variance unknown' curve is chosen because we rarely know the degree of scatter our tests results are likely to have.

For the n = 60 data points shown in **Figure 2** (right) above 6 m, k_n = 0.216. Hence the mean value of q_c predicted by the trend line given by Equ. (3) must be reduced by an amount Δq_c given by:

$$\Delta q_c = -k_n \times s_X = -0.216 \times 1.88 = -0.406 MPa \tag{6}$$

where s_X = 1.88 MPa was calculated in Step 3.

4.5 *Step 5 – plot the resulting characteristic line*

The final step in the procedure for determining the characteristic value is to adjust the best-fit line by the amount calculated in Step 4.

Figure 6 shows predictions of the characteristic values (solid lines) – and compares them with the corresponding best-fit lines for the reduced data set (dashed lines) – for all four cone tests.

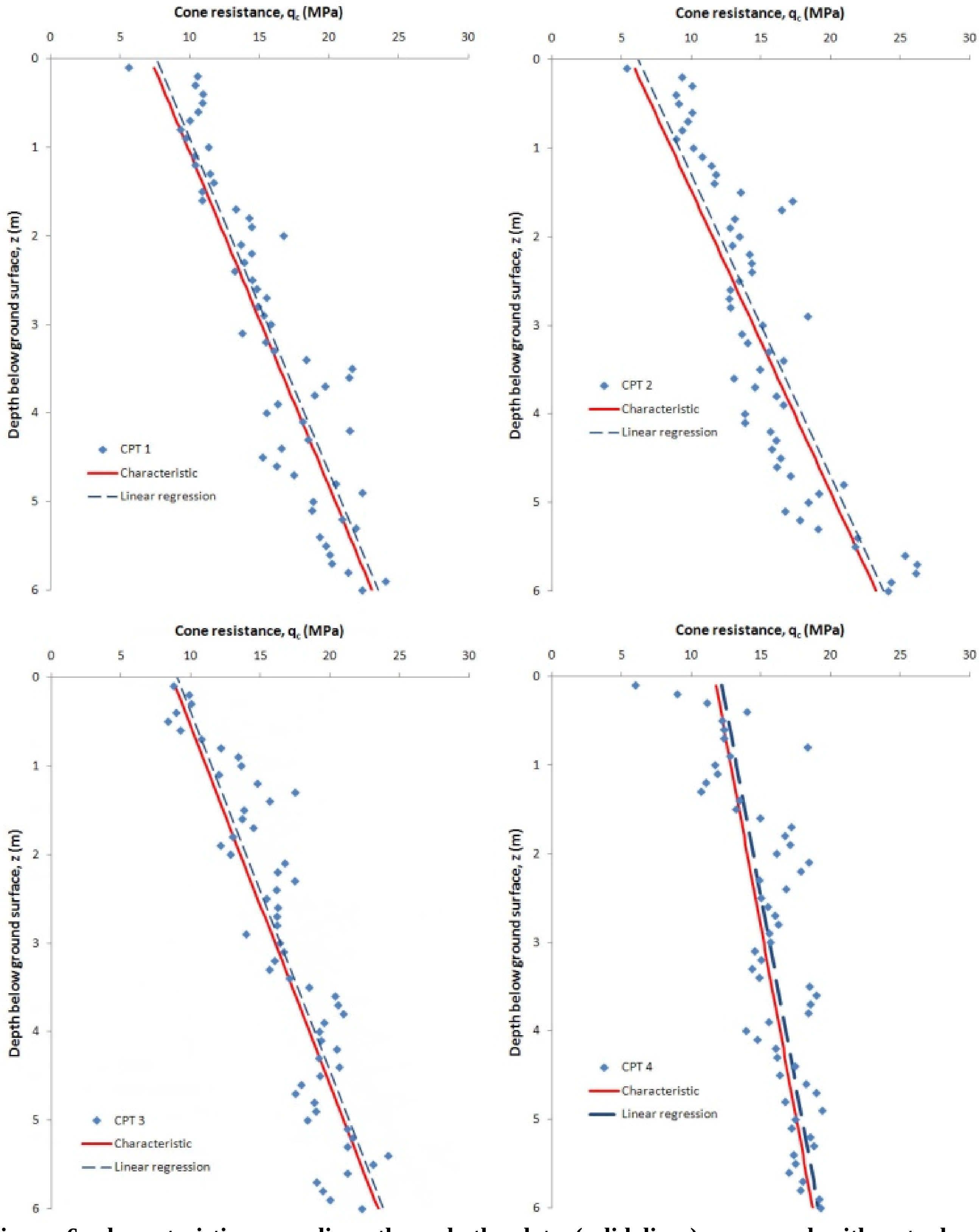

Figure 6, characteristic mean lines through the data (solid lines) compared with actual mean (dashed lines) for (top-left) CPT1; (top-right) CPT2; (bottom-left) CPT3; and (bottom-right) CPT4

It is remarkable how small the separation between the solid and the dashed lines is, which is a consequence of the large number of data points available (60) and the reasonably small degree of scatter in the data (which was obvious from **Figure 2**, left). In other circumstances, the separation would be much greater.

5 SECOND EXAMPLE

The procedure described above has been applied the results of standard penetration tests performed at a site comprising highly variable Boulder Clay. The data is taken from ETC 10 Design Example 2.2 (ETC 10, 2009) and is shown in **Figure 7** together with the best-fit linear regression line (dashed) and the predicted characteristic line (solid). The two very high blow counts (> 90) at 6 m have been ignored.

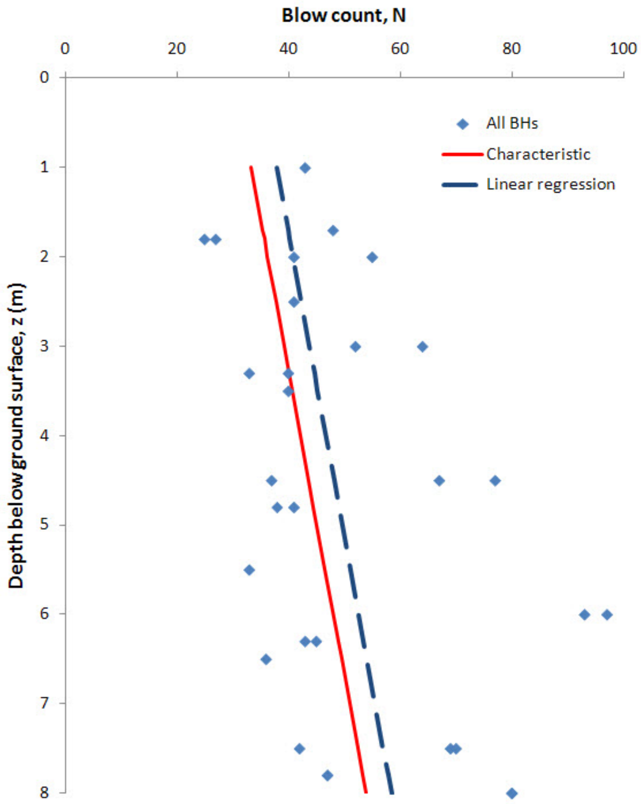

Figure 7, characteristic mean line (solid) compared with actual mean line (dashed) for standard penetration test results from ETC 10 Design Example 2.2

In this example, there are only 28 usable data points and so, from **Figure 5**, $k_n \approx 0.322$ (compared with 0.216 used previously for CPT3, i.e., 50% higher). With a larger deviation in data as well, this results in greater separation between the best-fit and characteristic lines.

6 LIMITATIONS OF THE PROCEDURE

The procedure described in this paper has a number of important limitations. First, it has been assumed that the data values increase linearly with depth; second, that the scatter in the data is random (i.e. there is no systematic influence affecting the data points); and third, that the differences between the data points and the trend-line (the residuals) fit a normal distribution.

The first limitation may be overcome by adopting a non-linear trendline, based perhaps on geological and geotechnical knowledge of the site. The second limitation is more difficult to overcome, since it is rare that we have sufficient knowledge to understand any systematic relationship between successive data points. In the absence of that knowledge, this is a limitation that we must live with. The third assumption can be checked during the procedure and data points omitted (as was done earlier) to rectify if possible.

7 CONCLUSION

Predictions of the characteristic value of a geotechnical parameter have been shown to vary greatly from one engineer to another, particularly when the data on which those predictions are made is highly scattered. Unfortunately, for many sites that is often the case. A procedure has been proposed to enable a consistent prediction of the characteristic value to be made using simple statistical techniques that follow the principles of the Eurocodes and are relatively easy to put into practice. The results of this procedure have been illustrated with data from two sites, one in sand and the other in clay.

REFERENCES

Bond, A. J. and Harris, A.J. (2008). Decoding Eurocode 7, Taylor and Francis, London, 608pp.

EN 1997-1, Eurocode 7: Geotehnical design – Part 1: General rules.

ETC 10 (2009). ISSMGE European Technical Committee 10, Evaluation of Eurocode 7. Design Examples 2 can be found at www.eurocode7.com/etc10.

Wikipedia, http://en.wikipedia.org/wiki/Pearson_product-moment_correlation_coefficient