Robust Assessment of Shear Parameters in Geotechnics

Wolfgang Fellin1) and Michael Oberguggenberger2)

1) Geotechnical and Tunnel Engineering, University of Innsbruck, wolfgang.fellin@uibk.ac.at
2) Engineering Mathematics, University of Innsbruck, michael.oberguggenberger@uibk.ac.at

Abstract. In the geotechnical determination of the cohesion \( c \) and the angle of internal friction \( \varphi \) of a soil from shear tests, a linear regression model is fitted to normal and shear stress data, and confidence bounds are computed. The applicability of standard linear regression is limited by the physical requirement of nonnegative cohesion and the statistical requirement of normality. We propose two methods from computational statistics that are able to overcome both obstacles: a bootstrap resampling method in case the experimental data set is sufficiently large, and a Bayesian approach for small samples. The methods are demonstrated at the hand of a real data set for glacial till.

Keywords: Shear parameters; non-normal regression; Bayesian methods; bootstrap confidence intervals; computational statistics.

1. Introduction

One of the principal tasks in geotechnical engineering is the determination of characteristic values of the shear strength \( \tau_f \), expressed by the cohesion \( c \) and the angle of internal friction \( \varphi \). Eurocode 7 requires that laboratory and in situ experiments have to be supplemented by experience in comparable situations and, in case statistical procedures are used, 5% and 95% confidence limits have to be observed for the characteristic values (Eurocode 7, 2004).

The shear strength of soil can be displayed in a stress diagram, where the shear stress \( \tau \) is plotted versus the effective normal stress \( \sigma \) (Figure 1-left). The shear strength is a generally curved line separating stress states which can be sustained by the soil (gray region in Figure 1) from stress states that cannot appear. Stress states on the line will cause local failure. A linearisation in a specific stress range is usually used for design

\[
\tau_f = c + \sigma \tan \varphi,
\]

(1)

with the cohesion \( c \) and the angle of internal friction \( \varphi \) (Figure 1-right).

The shear strength \( \tau_f \) can be determined by direct shear tests for example. A linear regression with the resulting pairs of normal stress and shear stress at failure of the sample yield the mean values of cohesion and friction angle. The confidence region for this estimated shear strength line can be calculated by standard statistical procedures (Figure 2-left). In civil engineering a 90% confidence level is of interest, since the probability of the shear strength being less than the lower limit of the region is only 5%, which is in agreement with the requirements of Eurocode 7.
A further linearisation of the lower confidence limit in a specific stress range (Figure 2-right) gives the stress line $\tau_{f,5\%}$, which is given by the equation

$$\tau = c_{5\%} + \sigma \tan \varphi_{5\%}$$

in the $(\sigma, \tau)$-plane. The stress line $\tau_{f,5\%}$ can serve for assessing the characteristic shear strength (Fellin et al., 2009). We call the pair $(c_{5\%}, \varphi_{5\%})$ the 5%-fractiles of the parameters $c$ and $\varphi$. A linearisation of the upper confidence limit yields $(c_{95\%}, \varphi_{95\%})$.

There are two obstacles for the standard linear regression model: First, in cases where the mean value of the cohesion is close to zero, the standard approach may result in negative $c_{5\%}$ values. Second, the regressor variables and hence the regression coefficients may fail to be normally distributed.

In principle, the first problem can be dealt with by replacing the quadratic optimisation problem in the regression analysis by a quadratic optimisation with constraints. This leads to subcases of regression with or without intercepts as well as to definitely not normally distributed variables. In this paper, we propose two approaches to address non-normality and to obtain geotechnically acceptable confidence limits.

In the case of availability of a sufficiently large sample size ($\geq 15$), we propose a computational approach that provides bootstrap confidence intervals by repeated resampling of the data. This approach does not require normality and the $c \geq 0$ constraint can be naturally guaranteed.
In the case of small sample sizes (< 5) we find a Bayesian approach appropriate that replaces the confidence intervals by high probability density regions of the posterior distribution. Here we face the problem of assigning the prior distribution. We show how diffusive priors (interval parameters) can be plausibly chosen. We briefly indicate how the method could be extended to families of priors, that is, a robust Bayesian approach.

In both approaches, current methods from computational statistics have to be employed, resulting in credible confidence limits for the characteristic values.

2. General idea of proposed methods

As outlined above, standard regression gives an estimate of the 90% confidence range as depicted in Figure 2-left. This estimate is based on the assumption of normality. In this geotechnical problem, the assumption is questionable and definitely violated if the $c \geq 0$ constraint is imposed. In order to arrive at an estimate of a specific stress range as in Figure 2-right, we need to know the joint distribution of $c$ and $\tan \varphi$, or at least a bivariate sample $(c_i, \tan \varphi_i)$, $i = 1, \ldots, M$, thereof. Each data pair $(c_i, \tan \varphi_i)$ determines a shear stress line; our task will be to determine the population of lines that lie within a 90% confidence range. The actual limiting stress line $\tau_{f,5\%}$ will be inferred by a geometrical enveloping procedure.

2.1. Determination of 90% confidence regions

In the sequel, we will produce samples $(c_i, \tan \varphi_i)$, $i = 1, \ldots, M$, following the joint distribution of $c$ and $\tan \varphi$ by two methods: by a Bayesian approach and by bootstrap resampling. In the Bayesian approach, the 90% confidence range will be obtained as a high probability density region. In the bootstrap resampling approach, we apply and compare two different methods. The first one is probabilistic: a bivariate normal distribution is fitted to the sample $(c_i, \tan \varphi_i)$, using second moments. For further processing (see Section 2.2), those sample points $(c_i, \tan \varphi_i)$ are selected that lie inside the 90% elliptic region of the bivariate normal distribution.

Of course, this introduces again a normality assumption. That is why we prefer the second method, which is also more adapted to an engineering point of view. We want to determine shear strength parameters such that the probability of a smaller shear strength is less than 5%. First we must define what is a smaller shear strength. Consider the two possible shear strength lines in Figure 3. Remember that $\tau_f$ is considered as linearisation of the generally curved shear strength line in a specific stress range. The areas below each line in this stress range are the stress states that can be sustained by the soil. The area below line B is smaller. We state therefore that the parameters $c$ and $\varphi$ of line B are related to a smaller shear strength than those of line A for the given stress range.

We sort all parameter pairs of the possible shear strength lines with respect to the area below these lines in the given stress range. If we then exclude the upper and lower 5% of the sorted lines we end up with a fan of lines that occur with the desired estimated probability of 90%.

For brevity, we shall refer to the approach using the fitted bivariate normal distribution for determining the 90% confidence region as standard method and to the approach using the comparison of the areas as area method.
2.2. Definition of fractile values

Consider now the remaining fan of shear strength lines with a probability of 90%, displayed as gray lines in Figure 4. The two intersection points of the lowest shear strength lines with the limits of the stress range gives what we define to be \( \tau_{f,5\%} \). The same procedure with the upper lines gives \( \tau_{f,95\%} \). This definition replaces the linearisation of the hyperbolas bounding the confidence region in Figure 2 obtained from classical linear regression.

3. Methods

3.1. Bootstrap resampling

The starting point is a set of laboratory measurements of shear and normal stress pairs \((\sigma_j, \tau_{f,j})\), \(j = 1, \ldots, N\), of sufficiently large size, say \(N \geq 15\). The cohesion \(c\) and the friction coefficient \(\tan \varphi\) are obtained as coefficients in a linear regression. A single set of measured data \((\sigma_j, \tau_{f,j})\), \(j = 1, \ldots, N\), gives rise to a single estimate of \(c\) and \(\tan \varphi\). It tells nothing about the joint distribution of the pair of parameters \((c, \tan \varphi)\).

A large sample of the joint distribution of \((c, \tan \varphi)\) can be produced by means of an empirical procedure called bootstrapping cases (Montgomery et al., 2001). A bootstrap sample of size \(N\) is obtained from the original sample \((\sigma_j, \tau_{f,j})\) by drawing randomly \(N\)-times with replacement (thus a bootstrap sample may have duplicated values among the \((\sigma_j, \tau_{f,j})\)). This procedure is repeated \(M\) times, say \(M = 1000\). Each time a linear regression is performed. In this way, a large sample of the
desired cohesion and friction coefficients \((c_i, \tan \varphi_i), i = 1, \ldots, M\) is obtained. It can be shown (Shao, 1995) that the bootstrap distribution coincides asymptotically with the distribution of \((c, \tan \varphi)\) that is defined via regression by the empirical distribution of the original sample \((\sigma_j, \tau_{f,j})\). For the choice of the size \(M\) we refer to the literature; in the absence of detailed estimates, \(M = 1000\) is generally recommended.

The generated pairs cohesion/friction coefficient \((c_i, \tan \varphi_i), i = 1, \ldots, M\), determine a fan of stress lines, from which the 90\% regions are computed as described in Section 2.

What concerns the geotechnical requirement of \(c \geq 0\), we have various options. First, the regression coefficients \((c, \tan \varphi)\) are obtained by minimising the mean square error, or equivalently, the sum of squares \(SS_E\) of the errors. We can adjoin the constraint \(c \geq 0\) in this quadratic minimisation problem. When \(c\) is estimated to be less than zero, this just means that a regression without \(y\)-intercept \((c = 0)\) is performed instead. A second option is offered by the bootstrap method. We can simply reject all bootstrap samples yielding \(c < 0\). It turned out that in our examples about 10\% of the bootstrap samples were rejected, leaving sufficiently many samples for the further analysis. Finally, one can do nothing and accept negative values for \(c\), hoping that the 5\% bound for \(c\) is still nonnegative. This is generally not the case, but included in our presentation for comparison.

We remark that there is an alternative resampling method for producing the empirical distribution of \((c_i, \tan \varphi_i)\). This is the \textit{jackknife} (Shao, 1995). In the jackknife, the new samples are not drawn by replacement, but rather as subsamples of smaller size \(N - 1, N - 2, \ldots\) The method is combinatorially more expensive. For this reason we will not present the jackknife technique here (though we undertook all calculations with the jackknife method as well).

3.2. \textsc{Bayesian approach}

When the size \(N\) of the original set of measured data is small, say \(N < 5\), the standard regression technique loses credibility and the bootstrapping becomes impossible (because it produces too many bootstrap samples with identical entries). In such a situation, a Bayesian approach is recommendable. We start with a brief introduction to Bayesian methods.

Suppose a random variable \(X\) is given whose statistical parameters we want to estimate. We concentrate here on the case that \(X\) has a probability density function \(p(x|\theta)\) that depends on certain statistical parameters \(\theta\). For example, when \(X\) is normally distributed, \(\theta\) comprises the mean and the variance and \(p(x|\theta)\) is the normal probability density function.

Prior knowledge is encoded in the probability density function \(p(\theta)\) of the parameter \(\theta\), the so-called a priori distribution. Given a set of measurements, data = \((x_1, \ldots, x_N)\), one may get an improved assessment of the statistical properties of the parameter \(\theta\) through the a posteriori distribution \(p(\theta|\text{data})\) according to Bayes’ formula

\[
p(\theta|\text{data}) = C p(\text{data}|\theta)p(\theta)
\]

with the normalising constant \(C\) and the \textit{likelihood function} \(p(\text{data}|\theta)\) given by

\[
\frac{1}{C} = \int p(\text{data}|\theta)p(\theta) \, d\theta, \quad p(\text{data}|\theta) = \prod_{i=1}^{N} p(x_i|\theta).
\] (3)
In order to explain how this is applied in regression analysis, we have to complete the linear model (1) by the random error term \( \varepsilon \)

\[
\tau_f = c + \sigma \nu + \varepsilon
\]

where we have written \( \nu = \tan \varphi \) for clarity. The error term \( \varepsilon \) has mean zero; we denote its variance by \( s_\varepsilon^2 \). The regressor variables are the stresses \( \tau_f \) and \( \sigma \) and play the role of the random variable \( X \) above. The parameter \( \theta \) is comprised by the regression coefficients \( c, \nu = \tan \varphi \) and the standard deviation \( s_\varepsilon \) of the error term. In the standard regression model, given \( \sigma \), the regressor variable \( \tau_f \) is assumed to be normally distributed with mean \( c + \sigma \nu \) and variance \( s_\varepsilon^2 \), thus it has probability density function

\[
p(x|\theta) = \frac{1}{\sqrt{2\pi s_\varepsilon}} \exp \left( -\frac{(x - c - \sigma \nu)^2}{2s_\varepsilon^2} \right).
\]

Given a set of measurement data \( \{\sigma_j, \tau_{f,j}\}, j = 1, \ldots, N \), the likelihood function becomes

\[
p(\text{data}|\theta) = \left( \frac{1}{\sqrt{2\pi s_\varepsilon}} \right)^N \exp \left( -\sum_{i=1}^N \frac{(\tau_{f,j} - c - \sigma_j \nu)^2}{2s_\varepsilon^2} \right).
\]

Finally, we have to choose the a priori distributions for \( c, \nu \) and \( s_\varepsilon \). For the sake of presentation, we work with so-called diffusive priors here, that is, we enter in our calculations with coefficients \( c, \nu \) and \( s_\varepsilon \) which are uniformly distributed on certain intervals, denoted by \([c_{\min}, c_{\max}]\), \([\nu_{\min}, \nu_{\max}]\), and similarly for \( s_\varepsilon \). A main advantage of the diffusive assumption is that one can enforce the \( c \geq 0 \) constraint in a simple way. The choice of the other interval bounds will be specified in the next section; a discussion of other types of priors is in Section 3.4.

The a posteriori distribution \( p(\theta|\text{data}) \) has to be determined by a threefold numerical integration with respect to \( c, \nu \) and \( s_\varepsilon \). Finally, the desired joint a posteriori distribution of \( (c, \nu = \tan \varphi) \) is obtained by a further integration with respect to \( s_\varepsilon \), and this is the result to be processed further. Actually, we employed a simplified version assuming that the variance \( s_\varepsilon \) was known exactly (see below for the choice of \( s_\varepsilon \)). This way we could drop the \( s_\varepsilon \)-integration and arrived at the a posteriori distribution \( p(c, \nu|\text{data}) \) in one double integration step.

The task of producing a sample of values \( (c_i, \nu_i) \) remains to be explained. The numerical integration delivered the values \( p(c_i, \nu_i|\text{data}) \) of the a posteriori density on a chosen grid with coordinates \( (c_i, \nu_i) \). From there we computed the level sets of \( p(c, \nu|\text{data}) \) by a search algorithm, and we calculated the respective total probability mass allocated in each level set. The level set containing 90% of all values \( (c, \nu) \) was chosen as the 90% confidence region. The sample \( (c_i, \nu_i) \) was defined as the set of grid points belonging to the 90% confidence region. The bounding stress lines \( \tau_{f,5\%} \) and \( \tau_{f,95\%} \) were then computed as in Section 2.
3.3. Estimation of variance and parameter interval limits

It turned out that leaving the regression variance \( s_2^2 \) diffusive would introduce too much variability in the Bayes estimators. If \( s_2^2 \) is known from geotechnical experience, this value should be taken. If it is not known, we propose here to estimate it from the data by means of the formula

\[
s_\varepsilon \approx \frac{q}{\sqrt{N}} \sqrt{\frac{SS_E}{N-2}}
\]

(5)

where \( q \) is the 95%-quantile of the Student distribution with \( N - 2 \) degrees of freedom and \( SS_E \) is the sum of squares of the residuals of the classical regression. \( SS_E/(N-2) \) actually is the standard estimate for the regression variance. The scaling factor \( q/\sqrt{N} \) was obtained empirically by numerical experiments and proved to produce plausible results in the geotechnical applications.

The decision about the a priori bounds for \( c \) and \( \nu = \tan \varphi \), if not based on hard geotechnical knowledge about the soil in question, should take into account the following points. First, the lower bound \( c_{\text{min}} \) for \( c \) should be nonnegative. Second, the intervals \([c_{\text{min}}, c_{\text{max}}]\) and \([\nu_{\text{min}}, \nu_{\text{max}}]\) should be large enough so that the 90% confidence region calculated from the a posteriori distribution lies inside the rectangle \([c_{\text{min}}, c_{\text{max}}] \times [\nu_{\text{min}}, \nu_{\text{max}}]\) which they form (too small intervals would lead to an artificial truncation). Finally, the intervals should be in a plausible range around the presumed mean values of \( c \) and \( \nu \). As there was no sufficient reason to exclude small values of \( c \) and \( \nu \), we always chose the intervals with zero lower bounds, that is, of the form \([0, c_{\text{max}}]\) and \([0, \nu_{\text{max}}]\). Excluding friction angles greater than 45\(^\circ\) led to the upper bound \( \nu_{\text{max}} = \tan \varphi_{\text{max}} = 1 \). Finally, we chose \( c = 100 \) kPa in all examples to follow in order to render them comparable, although this may not have been the most plausible choice in some cases.

The following three generic examples underline our claim that the proposed estimates for the a priori distributions lead to plausible results. Generic outcomes of \( N = 3 \) shear tests are listed in Table I.

<table>
<thead>
<tr>
<th>( \sigma )</th>
<th>( \tau ) series 1</th>
<th>( \tau ) series 2</th>
<th>( \tau ) series 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>66.6</td>
<td>68.3</td>
<td>22.7</td>
</tr>
<tr>
<td>150</td>
<td>120.7</td>
<td>116.1</td>
<td>73.4</td>
</tr>
<tr>
<td>250</td>
<td>161.6</td>
<td>160.0</td>
<td>116.6</td>
</tr>
</tbody>
</table>

We compare the results of the Bayesian approach with the results of a standard linear regression. Series 1 in Table I leads to a linear regression with moderately large variance, Series 2 produces a small regression variance, and Series 3 leads to a case where the standard lower confidence limit has a negative intercept, see Fig. 5. The shaded regions in Fig. 5 mark the 90% confidence regions resulting from the Bayesian approach. It is seen that they do not deviate much from the confidence regions obtained by standard regression. However, negative intercepts – as may happen when the cohesion is small – are now avoided, as can be seen in Fig. 5-right.
To complete the picture, we display the 90% confidence regions of the a posteriori distribution, approximated numerically as outlined in Section 3, referring to the generic data of Table I in Fig. 6. Each data point in the shaded truncated elliptic regions gives rise to a pair \((c, \nu = \tan \varphi)\) which in turn corresponds to a stress line in Fig. 5, the totality of which forms the shaded confidence regions. The lower truncations of the elliptic regions visible in Fig. 6 reflect the prior information \(c \geq 0\) that was encoded in the prior distributions.

3.4. Remarks on robust Bayesian methods

The same procedure offers the possibility to drop the assumption of normality in the likelihood function. In fact any other family of probability density functions that can be parametrised by the mean value and the variance will do, for example, a lognormal distribution. Likewise, the a priori densities need not be uniform; one may think of beta distributions. By repeating the same calculations leading to the 90% confidence region with different families of distributions, one may arrive at more conservative estimates. Since the methods for deriving such robust estimates are just a replication, we will not present them here.

The method is also quite insensitive with respect to the choice of the upper bounds for the diffusive priors, as long as these bounds are not too far from the respective mean values (say not
more than three times the mean). Choosing very large a priori intervals may lead to substantial changes in the resulting estimates for $c_{5\%}$, $c_{95\%}$, $\phi_{5\%}$ and $\phi_{95\%}$. Further investigations are needed to arrive at a general, geotechnically oriented proposal for the choice of the a priori intervals.

4. Application: ring shear test results

We present the behaviour of the proposed methods on a real test series. We use the results of 14 direct ring shear tests on 4 similar samples of glacial till (Schuppener, 1999), see Table II. One engineering task may be to assess one shear parameter set for all samples, the other to deduce the shear parameters for each sample, e.g. for the first four columns in Table II, which are the result of the first test series.

Table II. Results of direct ring shear tests (Schuppener, 1999, data reconstructed from Bild 1)

<table>
<thead>
<tr>
<th></th>
<th>Series 1</th>
<th></th>
<th>Series 2, 3 and 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$ (kPa)</td>
<td>100 150 200 300</td>
<td>150 250 300 100 150 250 200 150 200 250</td>
<td></td>
</tr>
<tr>
<td>$\tau$ (kPa)</td>
<td>68 127 135 206</td>
<td>127 148 197 76 78 168 123 97 124 157</td>
<td></td>
</tr>
</tbody>
</table>

The first task can be done with bootstrapping due to the relatively large data set. The second task will be performed with the proposed Bayesian approach.

4.1. Bootstrap

The result of 1000 bootstrap realisations without any constraint is given in Fig. 7 together with the standard regression information, i.e. the regression line ($c = 12$ kPa, $\phi = 31^\circ$) and the 90\% confidence region. The two methods as outlined in Section 2 of estimating the 90\% confidence region of the bootstrap sample are shown in the two other plots of Fig. 7.

Figure 7. Bootstrap sample with 1000 realisations.

Applying the fitted bivariate normal distribution of the shear parameters (method standard, Fig. 7-middle) or the minimum area criterion below the shear strength (method area, Fig. 7-right) yield different remaining fans of shear strength lines, see Fig. 8. Both fans lie approximately within...
the standard confidence region, which shows that bootstrap without constraint is similar to a standard regression analysis.

Fig. 9 shows the results obtained by bootstrapping where in each linear regression the constraint \( c \geq 0 \) was imposed. The selection of the 90% confidence region was again done by method standard and method area. In this case, small differences occur that lead to the distinct linearised bootstrap confidence regions in Fig. 9.

The results of all methods are summarised in Table III. For comparison we included the estimation with the Bayesian approach.

It should be pointed out that the bootstrap is a numerical simulation method and therefore comes with a numerical error: different bootstrap samples of size \( M \) lead to different results – obeying the general Monte Carlo simulation error bound of magnitude \( 1/\sqrt{M} \). The scatter in the case of using method standard turned out to be smaller than in the case of using method area. In this example with 1000 realisations the resulting friction angles may vary by approximately 1 degree and the cohesions by 3 kPa. Clearly, increasing \( M \) leads to smaller scatter; generally, \( M \) should be chosen so that the numerical error remains in an acceptable range from the point of view of the intended application.
Table III. Fractiles of shear parameters estimated with various methods: BS bootstrap, nc no constraint, pc positivity constraint, nr negative c-values rejected, area minimum area below shear strength line, standard fitted bivariate normal distribution of shear parameters.

<table>
<thead>
<tr>
<th>Approach</th>
<th>c5%</th>
<th>c95%</th>
<th>ϕ5%</th>
<th>ϕ95%</th>
<th>Approach</th>
<th>c5%</th>
<th>c95%</th>
<th>ϕ5%</th>
<th>ϕ95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical regression</td>
<td>0.0</td>
<td>24.2</td>
<td>31.0</td>
<td>31.3</td>
<td>BS, pc, 0.9 area</td>
<td>11.7</td>
<td>25.1</td>
<td>28.3</td>
<td>31.0</td>
</tr>
<tr>
<td>BS, nc, 0.9 area</td>
<td>4.2</td>
<td>25.0</td>
<td>29.4</td>
<td>31.0</td>
<td>BS, pc, 0.9 standard</td>
<td>5.3</td>
<td>28.3</td>
<td>30.5</td>
<td>30.3</td>
</tr>
<tr>
<td>BS, nc, 0.9 standard</td>
<td>-5.1</td>
<td>29.6</td>
<td>31.9</td>
<td>30.1</td>
<td>BS, nr, 0.9 area</td>
<td>12.8</td>
<td>26.5</td>
<td>28.1</td>
<td>30.8</td>
</tr>
<tr>
<td>Bayesian</td>
<td>6.0</td>
<td>18.5</td>
<td>31.0</td>
<td>31.1</td>
<td>BS, nr, 0.9 standard</td>
<td>5.9</td>
<td>28.8</td>
<td>30.2</td>
<td>30.1</td>
</tr>
</tbody>
</table>

4.2. Bayesian approach

The first series of the ring shear tests in Table II comprises $N = 4$ experiments. Standard linear regression yields $c = 13$ kPa and $\phi = 33^\circ$. Assessing characteristic shear parameters for this data set is possible with the Bayesian approach. We used the settings proposed in Section 3.3, that is, we fixed the regression variance $s^2_\varepsilon$ to the value given by Formula (5) with $SS_E$ from the standard linear regression. The a priori distributions were taken to be uniform on the interval $[0, 100]$ kPa for the cohesion $c$ and on the interval $[0, 1]$ for the friction coefficient $\nu = \tan \phi$. The result is shown in Figure 10, with the confidence and linearised confidence limits for the shear strength lines on the left and the underlying a posterior 90% confidence region for $c$ and $\nu = \tan \phi$ on the right.

![Figure 10. Result of the Bayesian approach with the first series in Table II: regression (left), 90% confidence region (right). The stress range for the geometrical enveloping is depicted by vertical dashed lines.](image-url)

Table IV. Fractiles of shear parameters of the first series in Table II estimated with two methods.

<table>
<thead>
<tr>
<th>Approach</th>
<th>c5%</th>
<th>c95%</th>
<th>ϕ5%</th>
<th>ϕ95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical regression</td>
<td>-17.8</td>
<td>43.2</td>
<td>31.7</td>
<td>34.1</td>
</tr>
<tr>
<td>Bayesian</td>
<td>7.5</td>
<td>52.0</td>
<td>26.8</td>
<td>33.4</td>
</tr>
</tbody>
</table>
Linearisation of the standard confidence limits would give a negative $c_{5\%}$ for this example, whereas our proposed Bayesian approach yields a reasonable positive value, see Table IV. Setting the minimum of the stress range (dashed lines in Fig. 10-left) to zero would give $c_{5\%} = 0$.

5. Conclusion

This article addresses the issue of determining characteristic soil parameters in geotechnical engineering. Eurocode 7 requires the provision of 5% and 95% confidence limits. We propose two methods for credibly determining such limits from direct shear tests, incorporating the constraint that the cohesion must be nonnegative and handling non-normality of distributions. The first method is applicable for sufficiently large sample sizes and employs bootstrap resampling to approximate the joint distributions of the cohesion and the friction angle. The second method can be used with small samples as well and is based on a Bayesian approach. Both methods require tools from computational statistics and are computer based.

The proposed methods should provide the geotechnician with a working procedure for arriving at the required characteristic values, using all available information in a systematic and transparent way.

References