

Machine learning algorithms applied to the blowout susceptibility estimation around pressurized cavities in drained soil

Fernando Patino-Ramirez

Georgia Institute of Technology, Atlanta, USA

Chloé Arson

Georgia Institute of Technology, Atlanta, USA

ABSTRACT: The need to solve complex phenomena that still do not have a closed form analytical solution is one of the challenges of current practice, including the susceptibility estimation of cavity blowout during horizontal directional drilling (HDD). Computational tools and high frequency data acquisition in conjunction with machine learning open up new opportunities to create tools that can aid in the understanding and design of these problems. We propose the use of a predictor function, calibrated with a pool of numerical simulations that can predict the susceptibility of blowout of a cavity, given a fixed geometrical and stress configuration based on the mechanical parameters of the soil the cavity is embedded in; results using k-means clustering (for classification) and support vector machines (SVM) to create the predictor function, showed an accuracy of about 87% in predicting the blowout susceptibility.

1 INTRODUCTION

The study of the response of geomaterials under complex loading conditions is an important portion of the geotechnical research and practice. Even though significant breakthroughs have been made towards the understanding and solution of many of these problems, oftentimes they become too complex to be solved analytically. Fortunately, advances in computational science and increase of computing power have made possible to adopt numerical methods to solve these complex problems with exceptional detail and within reasonable computing time.

This same increase in computational power has generated unprecedented amounts of data, both because of the increased number of simulations and because of the level of detail achieved in these simulations (Diebold *et al.*, no date). The necessity to analyze and interpret all this data has deemed conventional tools as insufficient to a level such that data science has become a leading field both in scientific research and in industrial practice (Provost and Fawcett, 2013; van der Aalst, 2016). These new (or reinvigorated) data science tools have been encased into the keyword category of Machine learning, the

use of which is ubiquitous to any research discipline, including geotechnical engineering and its goal of understanding the response of soils to complex loading conditions (Shahin, Jaksa and Maier, 2009; Puri, Prasad and Jain, 2018).

In the present study, we explore and propose the use of machine learning, specifically supervised and unsupervised learning, as tools that can aid both scientific research, giving insights towards the understanding of controlling mechanisms; and the industry, by creating reliable and fast tools and functions that help on projects design.

One particular problem that lacks an analytical, accurate solution is finding the maximum internal pressure that can be applied inside the cavity before causing unconfined shear failure (blowout) of the soil surrounding the borehole during horizontal directional drilling (HDD). In this paper, we focus on the problem of a pressurized, (initially) cylindric cavity embedded inside a soil subjected to biaxial far field stresses under drained loading conditions. This problem is not only pertinent to HDD (6-8), but

also, to micro tunneling, in situ testing (Mair and Wood, 1987; Li, Li and Sun, 2016; Zhou, Kong and Liu, 2016) (pressuremeter testing under anisotropic conditions) and resource storage and withdrawal (Wang *et al.*, 2009).

The role of the drilling fluid is crucial during every stage of the HDD process, including the boring of the pilot hole, its enlargement to the objective diameter (reaming) and the installation of the product pipe into the bored cavity (pullback) (Baumert, Allouche and Moore, 2005). This fluid acts as a cooling agent for the drill head, suspends and transports the cuttings from the excavation front to the surface for an effective drilling and it can even be responsible for rotating the drill bit (American Society for Testing and Materials, 2011); therefore, having an appropriate fluid pressure inside the borehole is paramount for HDD operation at any of its stages.

However, if the minimum required pressure of the drilling fluid exceeds the maximum allowable for the soil at a given point in the borehole, mud can return inadvertently or drilling fluid can be lost into the surrounding soil, which can affect the integrity of surrounding structures (e.g. pavements, foundations), water bodies (subterranean aquifers, ponds and rivers) or cause significant surface settlements (Kennedy, Skinner and Moore, no date; Keulen, 2001). The current state of the practice relies mainly on two equations to find this maximum allowable pressure value; the first is commonly known as the Delft equation (Keulen, 2001) developed in 2001; it is the first closed form solution that was proposed to quantify the maximum allowable pressure of the drilling fluid inside the bore cavity, nevertheless, it assumes isotropic far field stresses and cylindrical deformation of the cavity, and it is limited to the calculations in the elastic range until the yield point is reached.

These assumptions are not realistic in practice, and may become unconservative as far field stresses become anisotropic (Xia and Moore, 2006; Shu, Zhang and Liang, 2018). The second widespread equation used to quantify the blowout pressure, known as the Queens equation (Xia and Moore, 2006), releases some of the assumptions made in the Delft equation, by using an elasto-plastic material and anisotropic stress conditions. Nevertheless, it still assumes that the plastic region is circular and encompasses the cavity, and it assumes that the plastic point that is the farthest away from the cavity is above the crown of the cavity.

Several other authors have attempted to propose new, improved solutions including (Xia, 2009; Rostami *et al.*, 2016; Shu, Zhang and Liang, 2018). Still, the lack of consensus and limited practicality of new solutions has caused the state of the practice to still rely on the Delft and Queens equations, as can be seen on the Best practices manual from the NASTT on its latest version to this date (2017).

In this study, we create a numerical, finite element (FE) model, which is run using a wide range of combinations of the mechanical parameters from the soil that control its response. Then, we analyze the soil response by describing the geometry of the cavity and that of the plastic zone after the internal cavity pressurization. Geometric parameter combinations are then classified (using k-means algorithm) into 3 categories defined as low, mid and high susceptibility of blowout; finally, we make use of supervised learning (SVM) to generate a function that can predict the blowout susceptibility based on the mechanical parameters of the soil surrounding the expanding cavity.

2 FINITE ELEMENT MODEL

A finite element model is built in Abaqus is used to generate the source data that will later feed the classification/prediction algorithms. This model consists of a plane strain (2D) domain which approximates the cross section of the cavity since the drill length is significantly larger than the diameter of the cavity. Additionally, the existence of two planes of symmetry allowed us to use a quarter of the domain. The resulting square domain is shown in Figure 1.

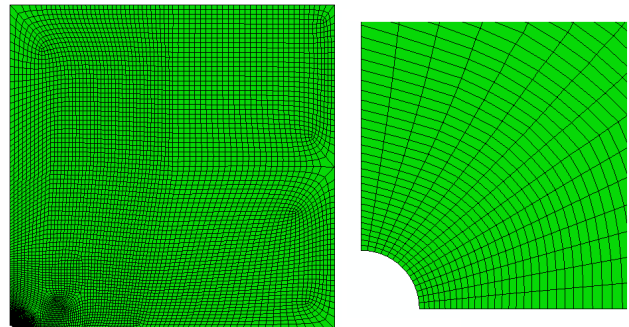


Figure 1. FEM domain of the simulation, the length of the edges of the square domain are 100 times larger than the cavity radius in order to avoid boundary effects. Left: full domain, Right: Close up of the cavity at the lower left corner.

The radius of the cavity was set to 1 m and, in order to avoid any possible boundary effects, the width (and height) of the model were set to 100 m. The corresponding symmetry boundary conditions were applied at the bottom and left edges of the domain; the right edge of the domain was fixed in order to generate a horizontal reaction to the later imposed vertical stress; during the first step, when the far field stress was applied, the displacement of the nodes along the edge of the cavity was fixed in order to prevent unrealistic and excessive deformations, this boundary condition was later removed when the outwards radial pressure was applied.

The vertical far field stress was fixed to 200 kPa throughout the simulations and similarly, the outwards radial pressure applied at the cavity nodes was fixed to 500 kPa. The horizontal far field stress is controlled by the at-rest lateral earth pressure coefficient (K_0) which is a function of the friction angle of the material. We used the Drucker-Prager (DP) constitutive model for the elasto-plastic soil, because it ensures better convergence than the more popular Mohr-Coulomb (MC) model.

Because the limits and meaning of MC resistance parameters are more broadly known than the DP ones, we chose MC parameters and used the matching plane strain DP parameters in the FEM model. The match between the plane strain response of the two models was obtained from Abaqus' documentation and is explained in Equation 1 and Equation 2 below:

$$\sin \phi = \frac{\tan \beta \sqrt{3(9 - \tan^2 \psi)}}{9 - \tan \beta \tan \psi} \quad (1)$$

$$c \cdot \cos \phi = \frac{\sqrt{3(9 - \tan^2 \psi)}}{9 - \tan \beta \tan \psi} \cdot d \quad (2)$$

where ϕ , ψ and c are the Mohr-Coulomb friction angle, dilation angle and cohesion respectively; and β and d are the Drucker Prager friction angle and cohesion respectively.

3 DATA GENERATION – POOL OF SIMULATIONS

In order to study the blowout potential of a cavity under the loading conditions described, we run a pool of simulations with mechanical parameters varying over a range

of values realistic for soil. Two elastic parameters are varied: Young's modulus (E) and Poisson's ratio (ν). The other two parameters varied thorough the tests are the Drucker-Prager friction angle (β) and dilation angles (ψ) of the material; a constant shear yield stress (also known as Drucker-Prager cohesion) of 5kPa is assigned to the material, since completely cohesionless materials result in early localized plastic strains around the cavity, resulting in lack of convergence. Non-associated, dilatant soil conditions are always ensured by setting a dilation angle (ψ) within the range $0 < \psi < \beta$.

We assign five different values to each of the four mechanical parameters, which results in a total of 625 (5^4) different combinations simulated. The same combinatorial strategy was used by the authors in a previous study done with a feature selection algorithm (Patino-Ramirez and Arson, 2018) that allowed to determine which mechanical parameters had a higher influence on the cavity elasto-plastic boundary shape.

The present study covers a broad range of values applicable to each parameter, going from fine to coarse granular materials and from loose to dense configurations (Kulhawy and Mayne, P.W. (Cornell Univ., Ithaca, 1990; Terzaghi, Peck and Mesri, 1996). The intervals of variations of the parameters are indicated in Table 1 below:

Table 1. Ranges of variation of soil mechanical parameters

Parameter	Units	Lower Bound	Upper Bound
Young's Modulus (E)	kPa	10	100
Poisson Ratio (ν)	-	0.15	0.45
MC friction angle (ϕ)	Deg ($^\circ$)	20	45
MC Dilation angle (ψ)	% of ϕ	5%	95%

4 RESPONSE VARIABLES – OUTPUT CHARACTERIZATION

After the simulations finished, the elastoplastic (EP) boundary was found by interpolating the values of the plastic strain stored at each element integration points. In order to get a smooth boundary, the boundary was defined as the contour corresponding to a level curve with an interpolated plastic strain of $1e-3$. The deformed cavity shape was also retrieved from the output database, following the final coordinates of the nodes along the cavity.

One of the simulations corresponding to a combination of very low stiffness and resistance parameters could not converge and aborted because of excessive distortion of the nodes when the internal pressure was applied, therefore, a total of 624 scenarios were considered in our analysis.

Contrary to the common assumption that the EP boundary is ellipsoidal, EP boundary shapes varied widely, progressing towards a localized plastic region that forms a band at a given angle towards the surface. The extent and orientation of these plastic regions are reminiscent of the behavior observed, for instance, in the shear bands developed on trap door tests (Ladanyi and Hoyaux, 1969).

Once the boundaries were extracted, they were characterized by a set of shape indexes. In the case of the EP boundary, four different indexes were defined: (i) The area of the plastic region plus the cavity (calculated from the polygon that encloses the EP boundary and the center point of the cavity); (ii) The solidity of the region (a convexity metric defined as the ratio between the area of the region to the area of its convex hull -- the smallest convex polygon that encloses the region), (iii) The distance between the center of the cavity to the point along the EP boundary that is the furthest away from the cavity center; (iv) The angle from the horizontal of the line joining the cavity center to the point that is the furthest away from the cavity center.

The area of the region is a measure of the extent of the plastic zone, not taking into account its shape. The solidity of the region is a value between 0 and 1 that decreases as the shape becomes less convex, where 1 corresponds to a convex region (an ellipse section for instance). Index 3 tracks the localized development of narrow elasto-plastic bands that appear and quantifies their extension, distancing from the center of the cavity. Lastly, index 4 measures the orientation of this fingering region. If the case the EP boundary is a perfect ellipse, indexes 3 and 4 correspond to the length of the major axis and a 90° angle, respectively.

The shape of the deformed cavity was fitted to an ellipse with great accuracy in every case and its area and eccentricity were calculated to describe the cavity size and shape respectively. The eccentricity of an ellipse is

defined as the ratio of the distance between the foci of the ellipse and its major axis; a value of 0 corresponds to a circle and the maximum value of 1 corresponds to the degenerate case of a line.

5 UNSUPERVISED LEARNING – BLOWOUT SUSCEPTIBILITY

From the extracted elasto-plastic boundaries, it becomes evident that the common assumption of an elliptical plastic region around the cavity is only valid when the soil is able to resist the internal pressure with little plastic deformation, nevertheless, when as the soil becomes weaker (in relation to the internal pressure), the extent of the plastic region increases rapidly, deviating from an elliptical shape and causing the blowout as it loses the capacity to withstand the load. Since the relation between the strength and deformability of the soil and the internal pressure is highly complex, and no comprehensive analytical solution is available, we will classify the elastoplastic boundaries in terms of their susceptibility to blowout.

We defined three categories: low, medium and high susceptibility to blowout. Making use of the geometric indexes defined for the boundaries, we use an automated classification technique, formally known as an unsupervised learning algorithm, in order to assign a category to each one of our performed simulations. We used a K-means implementation with the 4 different indexes as the features, 3 clusters (classes to classify into) and sample correlation as the similarity metric (Dhanachandra, Manglem and Chanu, 2015). Figure 2 shows all the EP boundaries and the corresponding blowout susceptibility category they were classified to.

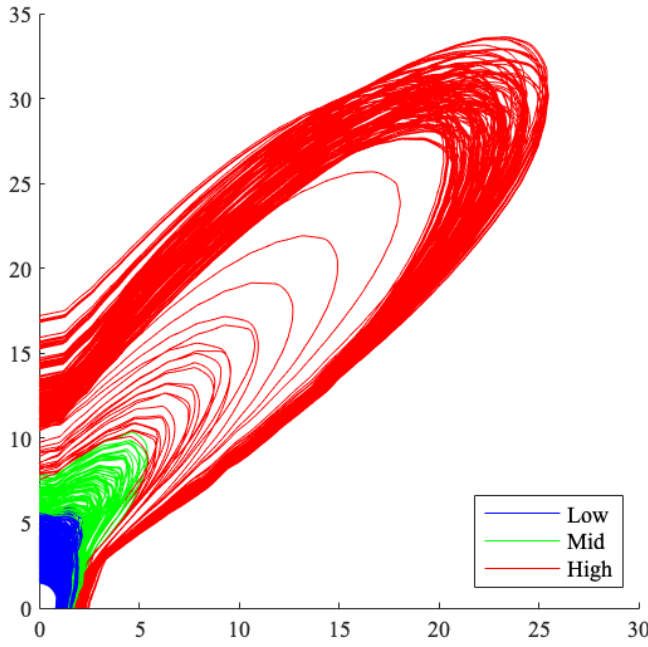


Figure 2. Elasto-plastic boundaries for all the numerical simulations (624) classified into susceptibility to blowout categories.

We then analyzed the distribution of the shape indexes by class. A good index should be able to split clearly the different categories, meaning that it should show very little overlap between the histograms of the different classes. The frequency histograms are shown on Figure 3.

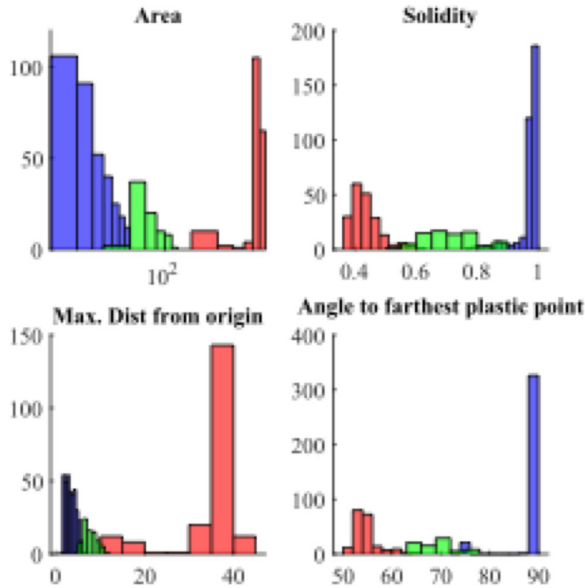


Figure 3. Frequency histograms by blowout susceptibility category, shown for the four EP boundary indexes: Area (m²), solidity, max distance from origin (m) and angle to farthest plastic point (degrees).

From these histograms, we can observe that there is a clear separation between the classes, suggesting a good

choice of indexes, and confirming the expected relationship between the blowout susceptibility formulated. High susceptibility simulations show significantly higher plastic region areas (the horizontal axis on Figure 3 is shown in log scale) in addition to highly non-convex shapes, due to the generation of a preferred plastic region extending far from the cavity at a relatively constant angle from the horizontal. This “fingering” phenomenon is supported by the distance to the cavity center and the angle to the furthest point (index 4).

Conversely, low susceptibility regions are elliptical, showing the angle to the furthest point is consistently around 90 degrees (major axis), very high solidity (convex shape) and small plastic extent. The mid susceptibility region falls in between the extremes, showing the transition from an ellipse to a plastic band of preferred orientation (fingering effect).

Lastly, we tested the same strategy with the deformed cavity indexes in order to test whether their distribution was directly linked to the blowout susceptibility. The results are shown on Figure 4.

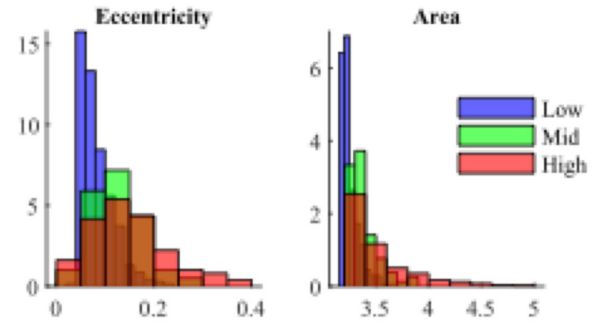


Figure 4. Frequency histograms by blowout susceptibility category for deformed cavity shape indexes: eccentricity and area (m²).

From these results it becomes apparent that the eccentricity (sometimes referred to as the ovality) and extent of the deformed cavity are not explained by the plastic region that surrounds it. The high overlap between categories shows that two given cavities of similar extent and shape can underlie completely different plasticity states.

6 SUPERVISED LEARNING – SUSCEPTIBILITY ESTIMATION FROM SOIL PARAMETERS

We now make use of a supervised learning algorithm in order to generate a predictor function that returns the predicted blowout susceptibility category (classification class) based on the soil mechanical parameters (acting as predictors variables)

To this end, we tried different learning algorithms categories: classification trees, naïve Bayes classifiers, support vector machines (SVM) and neural networks. After evaluating the performance of each one of the variations within these algorithms, we found out that the best performing algorithms were the neural networks and the cubic SVM, both with accuracies of between 82 and 87%.

Neural networks gave us a maximum accuracy of 86.7%, corresponding to a network with 10 hidden layers trained using scaled conjugate gradient backpropagation. We partitioned the data as: 70% of as the training set, and 15% for the validation and test sets each. On the other hand, the best performing SVM algorithm used a cubic boundary and 10 fold cross validation. The final accuracy of the algorithm was 82.5%.

Nevertheless, the performance of the neural network seemed to be highly variable depending on the tuning parameters and the separation of data into training and validations sets. This fact, added to the increased training time diverted us towards choosing the SVM algorithm as the most convenient, which achieved similar accuracy without the mentioned caveats. Figure 5 shows the confusion matrix for the predictor function gotten from the cubic SVM algorithm.

The confusion matrix shows that the class with the worst prediction accuracy is the medium susceptibility class, which is not surprising since it acts as the transition buffer between the two extreme classes. From the misclassified cases, 2.4% correspond to low susceptibility, which overestimates the blowout potential. 4% correspond to the medium class, from which 2.9% correspond to underestimation of blowout (unsafe scenario). Finally, the high susceptibility class underestimates 6.9% of the cases, from which 5.3% correspond to samples of high blowout susceptibility classified as low.

Output Class	Low	334 53.5%	18 2.9%	33 5.3%	86.8% 13.2%
	Mid	9 1.4%	53 8.5%	10 1.6%	73.6% 26.4%
	High	6 1.0%	7 1.1%	154 24.7%	92.2% 7.8%
		95.7% 4.3%	67.9% 32.1%	78.2% 21.8%	86.7% 13.3%
		Low	Mid	High	Target Class

Figure 5. Total confusion matrix for the susceptibility to blow-out prediction model using a cubic SVM learning algorithm.

7 CONCLUSIONS

The developed strategy showed promising results that could be implemented as a readily available tool to assess blowout potential or any other complex phenomena that has to be addressed even though it is not completely understood yet. These tools build upon the best capabilities of FEM (capability to model complex geometries and loading scenarios), unsupervised algorithms (capacity to find similarities and cluster large sets of data) and supervised learning algorithms (capacity to find correlations and causality relationships that explain and predict the interactions within the data).

Nevertheless, further research is needed in order to increase the reliability of these algorithms, especially to prevent unsafe predictions that may compromise the stability of structures. It is important to mention as well that this type of tool should not be seen as a replacement to traditional, analytic methods, but rather as an extra tool that can hint and guide towards the development of new comprehensive methods.

8 REFERENCES

van der Aalst, W. (2016) 'Data Science in Action', in *Process Mining*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 3–23. doi: 10.1007/978-3-662-49851-4_1.

American Society for Testing and Materials (2011) 'F1962-11

Standard Guide for Use of Maxi-Horizontal Directional Drilling for Placement of Polyethylene Pipe or Conduit Under Obstacles, Including river crossings', 04, pp. 1–18. doi: 10.1520/F1962-11.1.

Baumert, M. E., Allouche, E. N. and Moore, I. D. (2005) 'Drilling Fluid Considerations in Design of Engineered Horizontal Directional Drilling Installations', *International Journal of Geomechanics*, 5(4), pp. 339–349. doi: 10.1061/(ASCE)1532-3641(2005)5:4(339).

Dhanachandra, N., Manglem, K. and Chanu, Y. J. (2015) 'Image Segmentation Using K-means Clustering Algorithm and Subtractive Clustering Algorithm', *Procedia Computer Science*. Elsevier, 54, pp. 764–771. doi: 10.1016/J.PROCS.2015.06.090.

Diebold, F. X. *et al.* (no date) 'A Personal Perspective on the Origin(s) and Development of "Big Data": The Phenomenon, the Term, and the Discipline'. Available at: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.29.7.4593> (Accessed: 13 September 2019).

Kennedy, M. J., Skinner, G. D. and Moore, I. D. (no date) *elastic calculations of limiting mud pressures to control hydro-fracturing during hdd*. Available at: <https://pdfs.semanticscholar.org/92e4/c6656e0c05670297a437af00747952dab0e2.pdf> (Accessed: 13 September 2019).

Keulen, B. (2001) 'Maximum allowable pressures during horizontal directional drillings focused on sand'. Available at: <https://repository.tudelft.nl/islandora/object/uuid:ad91dad8-b958-481b-82d8-c7395d1a3874> (Accessed: 13 September 2019).

Kulhawy, F. H. and Mayne, P.W. (Cornell Univ., Ithaca, N. (USA). G. E. G. (1990) 'Manual on estimating soil properties for foundation design'.

Ladanyi, B. and Hoyaax, B. (1969) 'A STUDY OF THE TRAP-DOOR PROBLEM IN A GRANULAR MASS', *Canadian Geotechnical Journal*. NRC Research Press Ottawa, Canada, 6(1), pp. 1–14. doi: 10.1139/t69-001.

Li, L., Li, J. and Sun, D. (2016) 'Anisotropically elasto-plastic solution to undrained cylindrical cavity expansion in K0-consolidated clay', *Computers and Geotechnics*. Elsevier Ltd, 73, pp. 83–90. doi: 10.1016/j.compgeo.2015.11.022.

Mair, R. J. and Wood, D. M. (1987) *Pressuremeter Testing : Methods and Interpretation*. Elsevier Science.

Patino-Ramirez, F. and Arson, C. (2018) 'Numerical analysis of the plastic zone in drained soil around a pressurized cylindrical cavity under biaxial stress', in *IS Atlanta*. Atlanta, GA.

Provost, F. and Fawcett, T. (2013) 'Data Science and its Relationship to Big Data and Data-Driven Decision Making', *Big Data*. Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA, 1(1), pp. 51–59. doi: 10.1089/big.2013.1508.

Puri, N., Prasad, H. D. and Jain, A. (2018) 'Prediction of Geotechnical Parameters Using Machine Learning Techniques', *Procedia Computer Science*. Elsevier, 125, pp. 509–517. doi: 10.1016/J.PROCS.2017.12.066.

Rostami, A. *et al.* (2016) 'Estimation of Maximum Annular Pressure during HDD in Noncohesive Soils'. doi:

10.1061/(ASCE)GM.1943-5622.0000801.

Shahin, M. A., Jaksa, M. B. and Maier, H. R. (2009) 'Recent Advances and Future Challenges for Artificial Neural Systems in Geotechnical Engineering Applications', *Advances in Artificial Neural Systems*. Hindawi Publishing Corp., 2009, pp. 1–9. doi: 10.1155/2009/308239.

Shu, B., Zhang, S. and Liang, M. (2018) 'Estimation of the maximum allowable drilling mud pressure for a horizontal directional drilling borehole in fractured rock mass', *Tunnelling and Underground Space Technology*. Pergamon, 72, pp. 64–72. doi: 10.1016/J.TUST.2017.11.016.

Terzaghi, K., Peck, R. B. (Ralph B. and Mesri, G. (1996) *Soil mechanics in engineering practice*. Wiley.

Wang, S. Y. *et al.* (2009) '2D-numerical analysis of hydraulic fracturing in heterogeneous geo-materials', *Construction and Building Materials*. Elsevier, 23(6), pp. 2196–2206. doi: 10.1016/J.CONBUILDMAT.2008.12.004.

Xia, H. (2009) 'Investigation of maximum mud pressure within sand and clay during horizontal directional drilling', p. 272.

Xia, H. W. and Moore, I. D. (2006) 'Estimation of maximum mud pressure in purely cohesive material during directional drilling', *Geomechanics and Geoengineering*. Taylor and Francis, 1(1), pp. 3–11. doi: 10.1080/17486020600604024.

Zhou, H., Kong, G. Q. and Liu, H. L. (2016) 'Pressure-controlled elliptical cavity expansion under anisotropic initial stress: Elastic solution and its application', *Science China Technological Sciences*, 59(7), pp. 1100–1119. doi: 10.1007/s11431-016-6023-4.