XXII Congresso Brasileiro de Mecânica dos Solos e Engenharia Geotécnica
XI Simpósio Brasileiro de Engenheiros Geotécnicos Jovens
24 a 28 de agosto de 2026 – Brasília/DF
DOI: 10.47094/COBRAMSEG2026.2/24

Previsão Baseada Em Machine Learning De Resistência Não Drenada (Su) A Partir De Ensaios DSS Para Aplicações Offshore

Gleyce de Souza Baptista, M.Eng
Engª Geotécnica, UTEC-Geoservices, Rio de Janeiro, Brasil, gleycesouzaa@gmail.com

Ana Carolina Viana, M.Eng
Engª Geotécnica, UTEC-Geoservices, Rio de Janeiro, Brasil, ana.viana@benthic.com

Rogeane Seixas
Técnica em Estradas, UTEC-Geoservices, Rio de Janeiro, Brasil, rogeane.seixas@benthic.com

Caroline Juliette Barbosa Fernandes
Engª Geotécnica, UTEC-Geoservices, Rio de Janeiro, Brasil, carolinejuliette@poli.ufrj.br

Felipe Barreto Ribeiro Santos Macedo
Engenheiro, Petrobras, Rio de Janeiro, Brasil, felipe.b.macedo@petrobras.com.br

RESUMO: A estimativa da resistência não drenada (Su) é um elemento central em análises de estabilidade e no projeto de estruturas *offshore* em argilas marinhas. Este trabalho avalia a aplicação de modelos baseados em *machine learning* para a previsão da resistência não drenada obtida em ensaios de *Direct Simple Shear* (DSS), comparando seus resultados com correlações empíricas consagradas na engenharia geotécnica. A base de dados é composta por resultados de ensaios de laboratório realizados em argila marinha proveniente de uma campanha *offshore* na Bacia de Santos. A metodologia adotada fundamenta-se principalmente em regressão simbólica via algoritmo genético (AG), permitindo a obtenção de equações analíticas explícitas e fisicamente interpretáveis. O desempenho dos modelos foi avaliado por meio de métricas estatísticas usuais e confrontado com correlações empíricas tradicionais. Os resultados indicam que os modelos baseados em *machine learning* apresentam desempenho superior na reprodução dos valores medidos de Su,DSS, com redução significativa da dispersão e dos erros médios, sem prejuízo da interpretabilidade. Os achados reforçam o potencial dessas abordagens como ferramentas complementares às correlações empíricas na caracterização geotécnica de argilas marinhas *offshore*.

PALAVRAS-CHAVE: Resistência não drenada, Machine Learning, Ensaio DSS, Argila marinha, Offshore.

ABSTRACT: The estimation of undrained shear strength (Su) is a key parameter in stability analyses and in the design of offshore structures founded on marine clays. This study evaluates the application of machine learning–based models for predicting the undrained shear strength obtained from Direct Simple Shear (DSS) tests, comparing their performance with well-established empirical correlations used in geotechnical engineering. The database comprises laboratory test results from marine clay samples collected during an offshore campaign in the Santos Basin. The adopted methodology is primarily based on symbolic regression using genetic algorithms, enabling the derivation of explicit and physically interpretable analytical equations. Model performance was assessed using standard statistical metrics and benchmarked against traditional empirical correlations. The results indicate that the machine learning–based models provide superior agreement with measured Su,DSS values, with a significant reduction in data dispersion and average errors, without loss of interpretability. These findings highlight the potential of such approaches as complementary tools to empirical correlations for the geotechnical characterization of offshore marine clays.

KEYWORDS: Undrained shear strength, Machine Learning, DSS test, Marine clay, Offshore.

1 Introdução

A resistência não drenada, Su, é um parâmetro fundamental em análises de estabilidade e projetos geotécnicos, sendo usualmente obtida por ensaios de laboratório, como o *Direct Simple Shear*, ou estimada por meio de correlações empíricas baseadas em propriedades físicas e parâmetros de estado do solo. Embora amplamente empregadas devido à sua simplicidade, essas correlações apresentam limitações na representação da variabilidade e da complexidade do comportamento de argilas marinhas, especialmente em contextos *offshore* e bases de dados específicas. Nesse cenário, técnicas de *machine learning* têm sido investigadas como ferramentas complementares às abordagens convencionais, com potencial para aprimorar a estimativa da resistência não drenada sem perda de coerência física.

Este trabalho tem como objetivo avaliar o desempenho de modelos baseados em algoritmo genético (AG) na previsão da resistência não drenada obtida em ensaios de DSS, comparando-os com correlações empíricas tradicionais e discutindo seu potencial de aplicação prática na engenharia geotécnica.

1.1 Métodos empíricos para previsão de Su

Os ensaios de laboratório são essenciais para a compreensão do comportamento dos solos, pois permitem investigar, sob condições controladas, efeitos como anisotropia, histórico de tensões e dependência do caminho de carregamento, que não podem ser plenamente avaliados por ensaios *in situ*. Quando a execução de ensaios avançados, como DSS, triaxial de compressão ou extensão, não é viável, a literatura indica que a resistência não drenada *in situ* pode ser estimada de forma consistente a partir de ensaios edométricos, por meio da determinação da razão de sobreadensamento (OCR) e de correlações empíricas que relacionam Su à tensão vertical efetiva inicial, conforme expresso na equação 1. Essa abordagem tem se mostrado particularmente adequada para algumas análises, embora estudos clássicos ressaltam que, para projetos de maior complexidade, a realização de ensaios com diferentes modos de cisalhamento permanece fundamental para capturar os efeitos da anisotropia e da ruptura progressiva na seleção de parâmetros de resistência (Jamiolkowski et al, 1985).

$$\frac{S_u}{\sigma'_v} = (0.23 \pm 0.04) OCR^{0.8} \quad (1)$$

A correlação empírica para a estimativa da resistência não drenada proposta por Ahmed & Agaiby (2020), equação 2, baseia-se na integração de propriedades índice do solo e parâmetros de estado de tensões, buscando representar de forma consistente o comportamento resistente de argilas sob condições não drenadas. Os autores desenvolveram uma expressão para Su,DSS a partir de dados obtidos em ensaios de caracterização, limites de Atterberg e ensaios de adensamento, incorporando explicitamente a influência do índice de plasticidade (PI), do teor de umidade natural (wn), da massa específica dos sólidos (Gs), da tensão vertical efetiva inicial (σ'v0) e da razão de sobreadensamento (OCR). A formulação reflete o entendimento de que a resistência não drenada é dependente do caminho de tensões, sendo particularmente adequada para representar o modo de cisalhamento associado ao ensaio de cisalhamento direto simples (DSS). Dessa forma, a correlação permite estimar Su,DSS de maneira prática e fundamentada, utilizando parâmetros rotineiramente disponíveis em campanhas geotécnicas *offshore* e *onshore*, sem a necessidade de ensaios mecânicos avançados.

$$S_{u,DSS} = \left[\frac{0.624 G_s PI^{0.32}}{6 + G_s(0.416 PI^{0.32} + 0.06\, W_n)}\right] \sigma'_{v0} OCR^{0.8} \quad (2)$$

1.2 Aplicações de Machine Learning em Geotecnia

De forma geral, técnicas de *machine learning* (ML) têm sido amplamente aplicadas na engenharia para identificar padrões e relações complexas em bases de dados geotécnicas, sobretudo quando abordagens empíricas tradicionais se mostram limitadas (Goodfellow *et al.*, 2016). Entre os modelos mais utilizados destacam-se o *Random Forest*, a *Support Vector Regression* e as Redes Neurais Artificiais, que aprendem a partir de dados observados, mas diferem quanto à estrutura, capacidade de generalização e requisitos de

calibração (James, 2023; Li *et al.*, 2025). O desempenho desses modelos depende diretamente das características da base de dados e do objetivo da análise, sendo empregados com sucesso na previsão de propriedades e parâmetros geotécnicos, como a resistência não drenada Su (Ahmad, 2025; Nguyen *et al.*, 2025). Nesse contexto, a Regressão Simbólica destaca-se por buscar automaticamente uma forma funcional livre para representar a relação entre variáveis, utilizando a Programação Genética como estratégia de busca (Pacheco, s.d.). O núcleo desse método são os Algoritmos Genéticos, formalizados por John Holland (1975), que exploram o espaço de soluções por meio de uma população de candidatos e de operadores de seleção, cruzamento e mutação, permitindo a identificação simultânea da estrutura do modelo e de seus parâmetros. Essa abordagem evolutiva é reconhecida por sua robustez em problemas não lineares e complexos, reduzindo a probabilidade de convergência para ótimos locais e dispensando hipóteses rígidas sobre o comportamento dos dados (Oliveira, 2017; Pacheco, s.d.).

1.3 Ensaio de DSS

O ensaio de *Direct Simple Shear* (DSS) é um ensaio de laboratório empregado para determinar a resistência ao cisalhamento dos solos, no qual o corpo de prova é confinado lateralmente, impedindo deslocamentos horizontais e permitindo apenas a deformação por cisalhamento sob comprimento horizontal constante (Dyvik *et al.*, 1987). Executado, em geral, conforme a norma ASTM D6528-17, o ensaio é realizado em duas etapas principais: carregamento vertical sob confinamento lateral e posterior aplicação do cisalhamento simples. O DSS pode ser conduzido em condições verdadeiramente não drenadas ou drenadas a volume constante, sendo demonstrada a equivalência entre essas condições em termos de tensões efetivas (Dyvik *et al.*, 1987). Apesar de as tensões serem conhecidas apenas no plano horizontal durante o cisalhamento, o ensaio é amplamente reconhecido por representar de forma mais realista, quando comparado ao ensaio de cisalhamento direto, determinadas condições de campo, incluindo a rotação dos eixos das tensões principais (Villet *et al.*, 1985; Airey e Wood, 1987).

2 Materiais e métodos

2.1 Base de dados

A base de dados utilizada neste estudo foi obtida a partir de uma campanha geotécnica *offshore* realizada em uma área marítima da costa brasileira, na Bacia de Santos. A investigação de campo foi conduzida por meio de um sistema de perfuração submarina operado remotamente (PROD), permitindo a execução controlada de ensaios *in situ* e a recuperação de amostras no fundo do mar. As amostras indeformadas foram coletadas com sistemas de amostragem apropriados e, posteriormente, acondicionadas, seladas e preservadas conforme procedimentos técnicos. Então essas amostras foram transportadas para o laboratório e armazenadas até a realização dos ensaios *onshore* da UTEC Geoservices. A Figura 1 ilustra o sistema PROD em diferentes fases operacionais, bem como a embarcação de apoio responsável pela execução da campanha geotécnica.

Após a execução dos ensaios e consolidação dos resultados em planilha única, iniciou-se uma etapa de preparação e inspeção preliminar, cujo objetivo foi avaliar a consistência das medições e compreender a variabilidade inerente às propriedades geotécnicas analisadas. Foram geradas estatísticas descritivas contendo média, mínimo e máximo, conforme Tabela 1, permitindo identificar faixas típicas de valores esperados para argilas marinhas da região e apontar possíveis registros discrepantes.



(a) (b) (c)

Figura 1. (a) PROD sendo preparado para mergulho; (b) Sistema PROD emergindo após a operação; (c) Embarcação de apoio responsável pela execução da campanha geotécnica.

Tabela 1. Dados estatísticos do banco de dados.

	Prof (m)	σ'_{v0} (kPa)	σ'_{vm} (kPa)	OCR	γ_n (kN/m³)	e_0	W (%)	IP (%)	Su_{DSS}
min	0.56	2.84	7.94	1.00	13.89	1.31	37.61	27.00	4.34
média	32.05	190.39	225.78	1.74	15.64	2.00	71.26	56.77	61.00
max	98.99	664.84	870.4	25.35	17.72	2.98	117.42	96	182.62

2.2 Metodologia

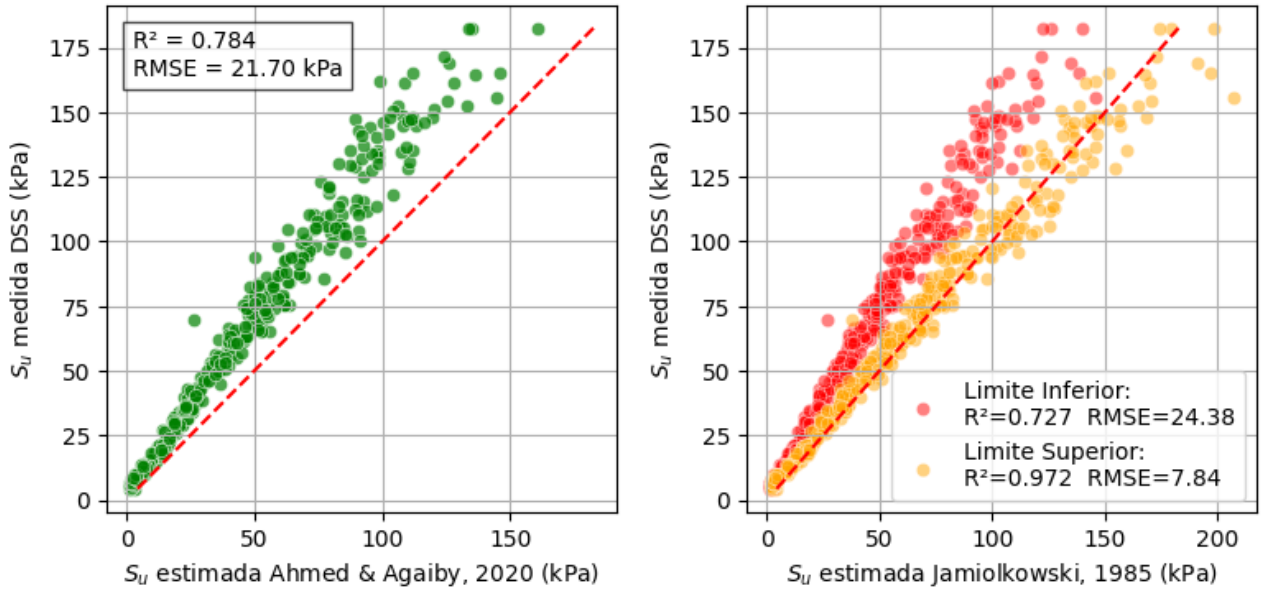
Os dados deste estudo foram consolidados em uma base única composta por resultados de ensaios laboratoriais, passando por etapa de pré-processamento para remoção de registros incompletos e garantia de consistência. A resistência não drenada obtida em ensaios de *Direct Simple Shear* (Su,DSS) foi definida como variável de interesse, enquanto parâmetros físicos e de estado do solo, como tensão vertical efetiva inicial, índice de vazios, teor de umidade, índices de plasticidade e liquidez, razão de sobreadensamento e massa específica dos sólidos, foram utilizados como variáveis de entrada. As variáveis de entrada foram padronizadas segundo $z = (x - \mu)/s$, com média μ e desvio-padrão s calculados no conjunto de treinamento; para as variáveis retidas pelos modelos finais. As equações apresentadas na Seção 3.2 foram reescritas algebricamente em termos das variáveis físicas, de modo a permitir sua aplicação direta em unidades de engenharia. O conjunto de dados foi dividido em subconjuntos independentes de treinamento (70%) e teste (30%). A metodologia baseou-se na aplicação de regressão simbólica via programação genética, visando obter equações analíticas explícitas e interpretáveis que relacionassem as variáveis de entrada à Su,DSS, a partir de um processo evolutivo que privilegia bom ajuste, simplicidade e coerência física. Para fins comparativos, correlações empíricas consagradas da literatura geotécnica, como as propostas por Ahmed & Agaiby (2020) e Jamiolkowski *et al.* (1985), foram aplicadas ao mesmo conjunto de dados. O desempenho dos modelos foi avaliado no conjunto de teste por meio de métricas estatísticas usuais, como o coeficiente de determinação (R^2) e o erro quadrático médio (RMSE), considerando-se adicionalmente a interpretabilidade e o potencial de aplicação prática das equações obtidas. Todo o processamento e as análises foram realizados em ambiente computacional baseado em Python, utilizando a plataforma Google Colab.

3 Resultados e discussões

3.1 Desempenho dos modelos tradicionais empíricos

A Figura 2 apresenta a comparação entre os valores de resistência não drenada medidos em ensaios de cisalhamento direto simples (Su,DSS) e aqueles estimados por correlações empíricas consagradas na literatura, no formato observado *versus* estimado, com a inclusão da linha de referência 1:1. A correlação de

Jamiolkowski et al (1985), apresentou desempenho global moderado (R2=0,727; RMSE = 24,38 kPa), com aumento significativo da dispersão para valores mais elevados de Su e tendência sistemática de subestimação e superestimação em diferentes faixas de tensão, indicando limitação na representação da variabilidade associada às propriedades físicas do solo; embora o limite superior da correlação apresente bom ajuste (R2=0,972; RMSE = 7,84 kPa). Em contraste, a correlação proposta por Ahmed & Agaiby (2020) apresentou desempenho superior (R2=0,784; RMSE = 21,70 kPa), com menor dispersão e melhor aderência à linha 1:1 ao longo da faixa analisada, atribuída à inclusão de parâmetros adicionais relacionados ao estado e à composição do solo; ainda assim, observa-se tendência de subestimação para valores elevados de Su, evidenciando que, apesar de úteis como ferramentas preliminares, as correlações empíricas apresentam limitações intrínsecas, reforçando a motivação para o uso de abordagens complementares baseadas em *machine learning*.



(a) (b)

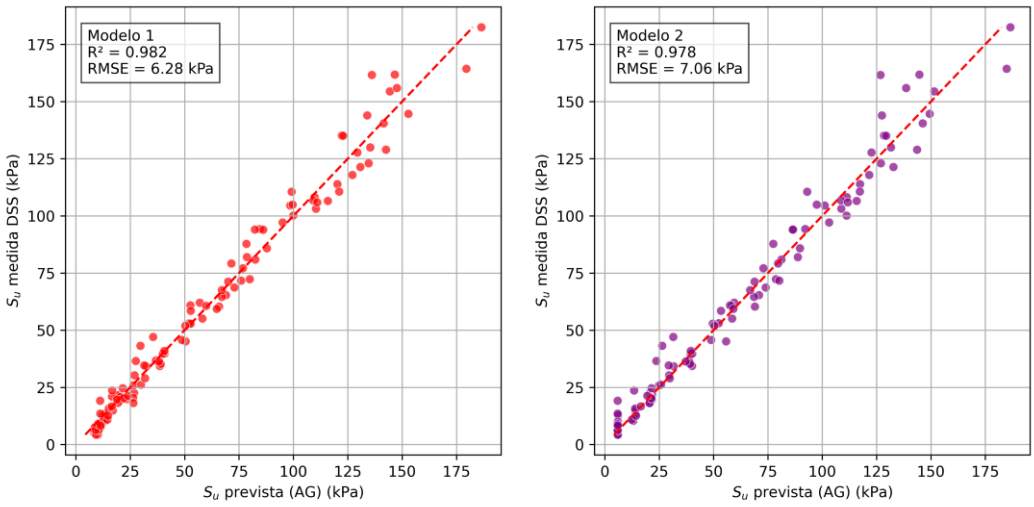
Figura 2. Comparação entre os valores de Su medidos em ensaios de DSS e estimados por correlações empíricas: (a) correlação de Jamiolkowski, 1985; (b) correlação de Ahmed & Agaiby, 2020.

3.2 Desempenho dos modelos de ML

A Figura 3 apresenta a comparação entre os valores de resistência não drenada medidos em ensaios de cisalhamento direto simples (Su,DSS) e aqueles estimados por modelos de *machine learning* baseados em regressão simbólica por programação genética, evidenciando desempenho significativamente superior às correlações empíricas tradicionais. Dois modelos foram desenvolvidos: o Modelo 1, que privilegia equações mais simples e interpretáveis ao restringir as operações matemáticas, resultando na equação 3, apresentou excelente ajuste nos dados de teste (R2=0,982; RMSE = 6,28 kPa), com elevada concentração dos pontos ao longo da linha 1:1; já o Modelo 2, obtido com um conjunto ampliado de operadores, resultou em uma relação linear direta, equação 4, também com desempenho elevado (R2=0,978; RMSE = 7,06 kPa). Nas equações 3 e 4, Su e σ'v0 são expressos em kPa e w em porcentagem. A comparação entre os modelos evidencia o equilíbrio entre complexidade da equação e desempenho preditivo, mostrando que a regressão simbólica é capaz de capturar relações não lineares relevantes com equações explícitas, transparentes e fisicamente interpretáveis, reforçando o potencial de algoritmos genéticos como uma alternativa complementar entre correlações empíricas clássicas para aplicações em engenharia geotécnica.

$$S_u = (0.1381\, w - 21.886)(-0.01914\, \sigma'_{v0} - 0.9001) \quad (3)$$

$$S_u = 0.2902\, \sigma'_{v0} + 4.905 \quad (4)$$



(a) (b)

Figura 3. Comparação entre os valores de Su nos dados de teste pelo (a) Modelo 1; (b) Modelo 2

3.3 Comparação direta entre ML × modelos empíricos

A Tabela 2 compara o desempenho das correlações empíricas tradicionais e dos modelos baseados em *machine learning* na estimativa da resistência não drenada obtida em ensaios DSS. As correlações empíricas apresentaram desempenho moderado (R2≈0,73) e erros médios da ordem de 20–25 kPa, com elevada dispersão, especialmente para valores mais elevados de Su. Em contraste, os modelos de regressão simbólica alcançaram desempenho significativamente superior nos dados de teste (R2≥0,98; RMSE<8kPa), mantendo equações analíticas explícitas e de fácil interpretação. Esses resultados indicam que os modelos de *machine learning* constituem uma alternativa complementar às correlações empíricas, oferecendo maior precisão sem perda de transparência para aplicações em engenharia geotécnica.

Tabela 2. Comparação entre os modelos de ML e tradicionais

Abordagem	Modelo	Variáveis	Forma da equação	R^2	RMSE (kPa)
Empírico tradicional	Jamiolkowski et al (1985)	σ'_{v0}, OCR	Lei potencial simples	≈ 0,73	≈ 24,4
Empírico tradicional	Ahmed & Agaiby (2020)	G_s, IP, w, σ'_{v0}, OCR	Expressão empírica multivariada	0,78	21,7
ML – Algoritmo Genético	Modelo 2	σ'_{v0}	Linear	0,98	7,06
ML – Algoritmo Genético	Modelo 1	w, σ'_{v0}	Não linear explícita	0,98	6,28

4 Conclusões

Este trabalho avaliou a estimativa da resistência não drenada obtida em ensaios de DSS a partir de correlações empíricas tradicionais e de modelos baseados em ML, utilizando uma base de dados composta por argila marinha de uma campanha *offshore* da Bacia de Santos. As análises evidenciaram que as correlações empíricas, embora amplamente utilizadas na prática geotécnica e de fácil aplicação, podem apresentar limitações na reprodução dos valores medidos de Su,DSS, especialmente em faixas mais elevadas de resistência, refletidas em dispersão significativa e erros médios relativamente elevados. Em contraste, os modelos desenvolvidos por regressão simbólica, algoritmo genético, apresentaram desempenho estatisticamente superior, com redução expressiva do erro e melhor aderência aos resultados de laboratório, mantendo coerência com as tendências físicas observadas.

Destaca-se que a melhoria de desempenho obtida com os modelos de ML não ocorreu à custa da perda de interpretabilidade, uma vez que o algoritmo genético forneceu equações analíticas explícitas, passíveis de análise crítica por engenheiros e de aplicação direta com variáveis em unidades de engenharia. A comparação entre modelos de diferentes níveis de complexidade indicou que formulações relativamente simples já são capazes de capturar grande parte da variabilidade observada, reforçando a importância de critérios de semelhança para aplicações práticas. De forma geral, os resultados indicam que abordagens baseadas em ML podem atuar como ferramentas complementares às correlações empíricas tradicionais, contribuindo para a redução de incertezas na estimativa da resistência não drenada quando aplicadas à bases de dados representativas e bem caracterizadas.

REFERÊNCIAS BIBLIOGRÁFICAS

Ahmad, M.; Mohammad A. Z.; Hamad, A.; Mohanad, M. S. S.; Jawad, B. M.; Shay H.; Tariq, O.; Abdullah, A. Improved prediction of soil shear strength using machine learning algorithms: interpretability analysis using SHapley Additive exPlanations. Frontiers in Earth Science, v. 13, art. 1542291, 2025.

Ahmed, S.M.; Agaiby, S.S. (2020). Strength and stiffness characterization of clays using Atterberg limits. Transportation Geotechnics, 25, 100420. Elsevier. https://doi.org/10.1016/j.trgeo.2020.100420

FATEHNIA, Milad; PARAN, Sanaz. Applying machine learning for SPT-N prediction from raw CPT data. In: CONFERENCE ON DEEP FOUNDATIONS, 50., 2025, Nashville (TN). Proceedings… Nashville (TN): Deep Foundations Institute, 2025. p. 345–355.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. Deep Learning. Cambridge: MIT Press, 2016.

JAMES, Gareth et al. An Introduction to Statistical Learning: with Applications in Python. Cham: Springer, 2023. (Springer Texts in Statistics).

Jamiolkowski, M.; Ladd, C.C.; Germaine, J.T.; Lancellotta, R. (1985). New developments in field and laboratory testing of soils. Proceedings of the International Conference on Soil Mechanics and Foundation Engineering, 1985.

LI, Kai; JIA, Pengfei; CHEN, Zihao; WANG, Yong. Fine-Scale Stratigraphic Identification Using Machine Learning Trained on Multi-Site CPTU Data. Geosciences, Basel, v. 15, n. 11, art. 437, 2025.

NGUYEN, Ho-Hong-Duy; NGUYEN, Thanh-Nhan; PHAN, Thi-Anh-Thu; HUYNH, Ngoc-Thi; HUYNH, Quoc-Dat; TRIEU, Tan-Tai. Explainable artificial intelligence model for the prediction of undrained shear strength. Theoretical & Applied Mechanics Letters, v. 15, n. 3, p. 100578, 2025.

OLIVEIRA, N. C. P. Aplicabilidade dos Algoritmos Genéticos na Engenharia Civil. Trabalho de Conclusão de Curso (Graduação em Engenharia Civil) – Universidade Federal da Paraíba, João Pessoa, 2017.

PACHECO, M. A. C. Algoritmos Genéticos: Princípios e Aplicações. Laboratório de Inteligência Computacional Aplicada (ICA), PUC-Rio, [s.d.].

AIREY, D. W.; WOOD, D. M. An evaluation of direct simple shear tests on clay. Geotechnique, v. 37, n. 1, p. 25–35, 1987.

VILLET, W. C. B.; SITAR, N.; JOHNSON, K. A. Simple Shear Tests on Highly Overeonsolidated Offshore Silts. 17th Offshore Technology Conference, 1985.

DYVIK, R. et al. Comparison of truly undrained and constant volume direct simple shear tests. 1987. Geotechnique, v. 37, n. 1, p. 3–10.