Bulletin of Engineering Geology and the Environment (2026) 85:206
https://doi.org/10.1007/s10064-026-04849-8

ORIGINAL PAPER

Integrated experimental and statistical evaluation of hydraulic and strength behavior in compacted high-plasticity fine-grained soils

Murat Türköz[1] · **Burak Yılmaz**[1] · **Hakan Bilici**[1]

Received: 28 July 2025 / Accepted: 16 February 2026
© The Author(s) 2026

Abstract
This study investigates the hydro-mechanical behavior of twenty compacted high-plasticity fine-grained soils by integrating laboratory testing with water retention modeling and statistical analysis. Soil–water characteristic curves (SWCCs) were generated using the Perera and Zapata empirical formulations, and air-entry values (AEVs) were determined through a reproducible tangent-line procedure. Undrained shear strength behavior was evaluated using connected bilinear failure envelopes in p–q space, from which mechanical transition thresholds (P_{break}) were identified. The results show that the Perera model consistently predicts higher and more physically representative AEVs for plastic soils compared with the Zapata model. AEV was found to relate closely to P_{break} for many samples, indicating that hydraulic desaturation and mechanical strength transitions originate from similar microstructural mechanisms, although the two thresholds do not always coincide. Principal component analysis and K-means clustering revealed three distinct soil behavior groups governed by plasticity, compressibility, suction retention, and strength characteristics. Soils with high fines content and high plasticity formed a cluster with the largest AEV and P_{break} values, reflecting suction-controlled mechanical behavior. In contrast, denser and less plastic soils transitioned earlier to friction-dominated response with lower hydraulic and mechanical thresholds. The findings demonstrate that incorporating both hydraulic and mechanical transition parameters provides a unified perspective for interpreting the behavior of compacted unsaturated clays. This dual-threshold framework enhances understanding of hydro-mechanical coupling and offers a practical basis for classifying fine-grained soils in engineering design.

Keywords Compacted fine-grained soils · Soil water characteristic curve (SWCC) · Air entry value (AEV) · Bilinear shear strength envelope · Principal component analysis (PCA) · K-means clustering

Introduction

High-plasticity fine-grained soils are commonly encountered in road embankments, foundations, earth dams, slopes, and landfill liners, where they frequently change water content and stress state. Their microstructure, dominated by active clay minerals, results in a high water retention capacity, pronounced shrink–swell behavior, and significant variations in shear strength and stiffness under environmental wetting and drying conditions (Fredlund and Rahardjo 1993; Shafqat et al. 2025). These hydro-mechanical responses are strongly influenced by suction, which evolves in response to climatic or load-induced moisture changes and governs both volumetric and strength behavior in unsaturated soils. Suction contributes to apparent cohesion at low moisture contents, modifies stiffness and compressibility, and affects shear strength through changes in effective stress pathways.

The soil water characteristic curve (SWCC) provides the fundamental relationship between matric suction and water content and is widely used for modeling permeability, volume change, and unsaturated shear strength. A key feature of the SWCC is the air entry value (AEV), which marks the onset of desaturation as air first invades the largest pores. This hydraulic transition often coincides with changes in

✉ Murat Türköz
mturkoz@ogu.edu.tr

Burak Yılmaz
brkylmz186@gmail.com

Hakan Bilici
hbilici@ogu.edu.tr

[1] Civil Engineering Department, Eskisehir Osmangazi University, Eskisehir 26040, Turkey

Published online: 16 March 2026

stiffness and load-bearing behavior, making AEV a critical index for interpreting the engineering performance of compacted clays (Quevedo et al. 2024). Numerous empirical SWCC models, including those derived from the Fredlund and Xing (1994) formulation, have been proposed to estimate AEV and other retention parameters from simple index properties such as plasticity index (PI), liquid limit (LL), and percent fines (P200). Among these, the formulations of Zapata et al. (2000) and Perera et al. (2005) allow rapid estimation of SWCC parameters in the absence of suction-controlled laboratory measurements. This makes them especially attractive for practical applications. While these models provide practical AEV estimates, their potential link to mechanical transitions such as breakpoints in bilinear shear strength envelopes has received limited attention. Recent modeling advances, including the bimodal SWCC function proposed by Bello et al. (2025) and the void-ratio-dependent SWCC predictions of Zhai et al. (2023), demonstrate the growing importance of capturing microstructural effects in SWCC formulation. Recent studies have increasingly combined laboratory characterization with numerical approaches to better capture the microstructural and suction related mechanisms governing unsaturated soil behavior. For instance, Zhang et al. (2022) employed a coupled experimental and discrete element modeling (DEM) framework to analyze both the SWCC and suction stress characteristic curve of reticulated red clay, demonstrating how pore scale interactions influence macroscopic suction and strength behavior. Their results highlight the importance of microstructure sensitive modeling in accurately predicting hydraulic transitions. Integrating such insights with classical SWCC formulations further emphasizes the need to examine whether hydraulic transition points such as the AEV also align with mechanical thresholds observed in shear strength envelopes.

Parallel to hydraulic characterization, the mechanical behavior of compacted fine-grained soils is often nonlinear under undrained loading. Several experimental studies have corroborated the influence of suction on shear strength transitions. Niu et al. (2024) demonstrated that the shear strength envelope of weakly expansive soils shifts markedly as suction varies, showing nonlinear transitions around AEV-level suction. Zhai et al. (2019) similarly emphasized the need to accurately identify AEV to capture the transition between saturation-controlled and suction-controlled mechanical behavior. Lin et al. (2022) further showed that soils with bimodal SWCCs may exhibit bilinear strength envelopes due to dual capillary regimes. These results collectively support the hypothesis that hydraulic transitions reflected in SWCC may correspond to mechanical transition points in strength envelopes.

While the classical Mohr–Coulomb framework assumes a single linear failure envelope, experimental evidence shows that compacted clays, especially at different degrees of saturation, frequently exhibit bilinear shear strength envelopes in p–q space. These envelopes reflect a transition from a low stress regime dominated by structure and apparent cohesion to a high stress regime governed more strongly by frictional response and particle rearrangement. The stress level at which the slopes of these segments change is referred to here as the mechanical breakpoint (P_{break}), representing a strength transition threshold analogous in conceptual role to the hydraulic AEV.

Although AEV is well established as a hydraulic transition parameter and P_{break} is routinely observed in undrained strength interpretation, very few studies have examined these two thresholds together. The fundamental question of whether hydraulic desaturation thresholds (AEV) correspond to mechanical transition thresholds (P_{break}) in compacted high-plasticity soils remains largely unexplored. Identifying such relationships would provide new insights into hydro-mechanical coupling and improve the predictive framework for the behavior of compacted clays used in geotechnical design.

Temperature also affects water retention and shear strength of high-plasticity clays due to changes in surface tension, viscosity, and the thickness of adsorbed water layers. Although temperature-dependent constitutive models such as adjusted SWCC formulations and thermo–hydro–mechanical coupling approaches exist, the present study focuses exclusively on isothermal conditions.

Beyond experimental characterization, data-driven statistical tools offer complementary insights into the multivariate nature of unsaturated soil behavior (Chiu and Ng 2012; Johari et al. 2006; Tripathy et al. 2014; Zhang et al. 2022). Bekdaş et al. (2025) utilized machine learning models to predict the static and cyclic properties of extremely high plasticity clays treated with lime and fly ash, achieving high performance in their predictions. These advancements highlight the potential of combining traditional geotechnical methods with modern computational tools for enhanced soil analysis. Multivariate analysis and machine learning (ML) are becoming indispensable in geotechnical research for identifying underlying patterns and improving predictive models (Gao 2024). Harle and Wankhade (2025) reviewed ML applications in seismic geotechnics, showcasing the value of algorithms like support vector machines (SVM) and random forests (RF). Hsiao et al. (2024) applied explainable AI using SHAP values to predict liquefaction-induced lateral spreading, enhancing model transparency. Ensemble ML models have shown superior accuracy in predicting vertical swelling potential, and graph neural

networks have improved modeling of granular flow behavior. Additionally, Sharma et al. (2025) explored the use of machine learning models for SWCC prediction, employing grain size distribution and suction data to estimate retention curves, demonstrating high accuracy. Their work supports the potential for integrating data-driven SWCC estimation with experimental or graphical AEV methods. In addition, it highlights promising opportunities for automated, regression-based curve fitting that could complement or validate graphical AEV determination. Recent applications of machine learning and multivariate clustering in geotechnics demonstrate the potential of such methods to identify hidden patterns, reduce dimensionality, and classify soils based on integrated hydraulic–mechanical indicators. However, few studies have incorporated classification methods such as principal component analysis (PCA) and K-means clustering to evaluate compacted clay behavior in terms of both AEV and strength-transition parameters.

The present study addresses these gaps by integrating hydraulic, mechanical, and statistical analyses of 20 compacted high-plasticity clay soils. SWCCs were generated using the Perera et al. (2005) and Zapata et al. (2000) models, and AEVs were determined using a structured tangent-line method. Shear strength behavior was characterized using connected bilinear p–q envelopes optimized by the maximum R^2 criterion. The resulting hydraulic (AEV) and mechanical (P_{break}) thresholds were compared to investigate potential coupling. Finally, PCA and K-means clustering were applied to identify multivariate behavioral patterns among the samples. The findings offer new insights into the relationship between hydraulic and mechanical transitions in compacted plastic soils and provide a unified framework for interpreting unsaturated soil behavior.

Materials

Soil characterization

The physical properties of the soils were characterized through sieve analysis, hydrometer analysis, and Atterberg limit testing. Particle size distribution results showed that all samples were dominated by fine-grained material, with the percentage passing the No. 200 sieve (P200) ranging from 71.35% to 96.17%. Clay content varied from 33.13% to 64.51%, while silt content ranged between 28.02% and 56.55%. Gravel and sand fractions were low for most samples, except Samples 2 and 3, which contained higher gravel contents (15.36% and 12.41%, respectively) (Fig. 1).

Consistency limit results confirmed the highly plastic nature of the soils. Liquid limits (LL) ranged from 52% to 72%, while plasticity indices (PI) varied from 25% to 39%,

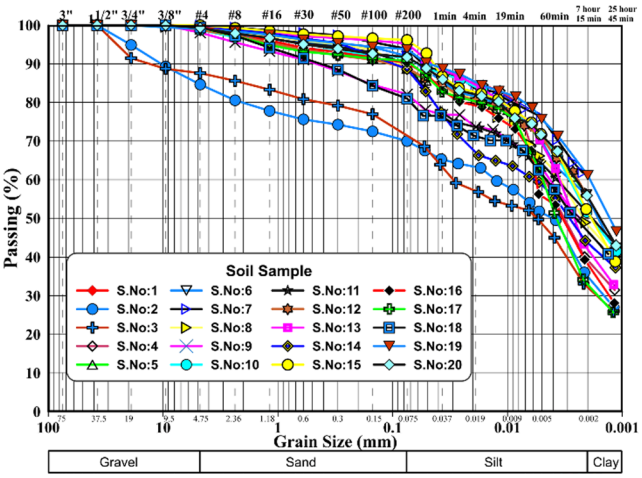


Fig. 1 Grain size distribution curves of the soil samples

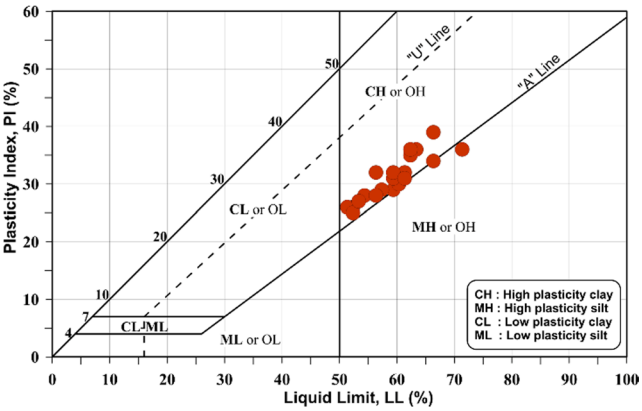


Fig. 2 Plasticity chart (Casagrande chart) showing the distribution of the soil samples

classifying 19 of the 20 samples as CH (high-plasticity clay) and one sample as MH (high-plasticity silt) on the Casagrande chart (Fig. 2).

Mineralogical composition

Mineralogical analysis was conducted using X-ray diffraction (XRD). Each diffraction pattern was normalized to its maximum peak intensity and vertically offset to enable clear visual comparison (Fig. 3). The mineralogical composition was generally characterized by the presence of kaolinite and illite as the major clay minerals. Smaller but detectable amounts of smectite/montmorillonite were identified in several samples. Quartz was consistently observed as the dominant non-clay mineral.

The predominance of kaolinite and illite explains the moderate-to-high plasticity and relatively stable fabric observed in many samples, while the occurrence of smectitic minerals accounts for the elevated swelling potential

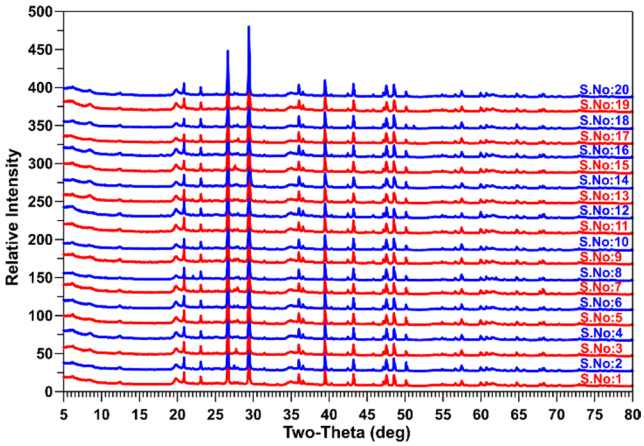


Fig. 3 Vertically offset normalized XRD patterns of the 20 clay soil samples

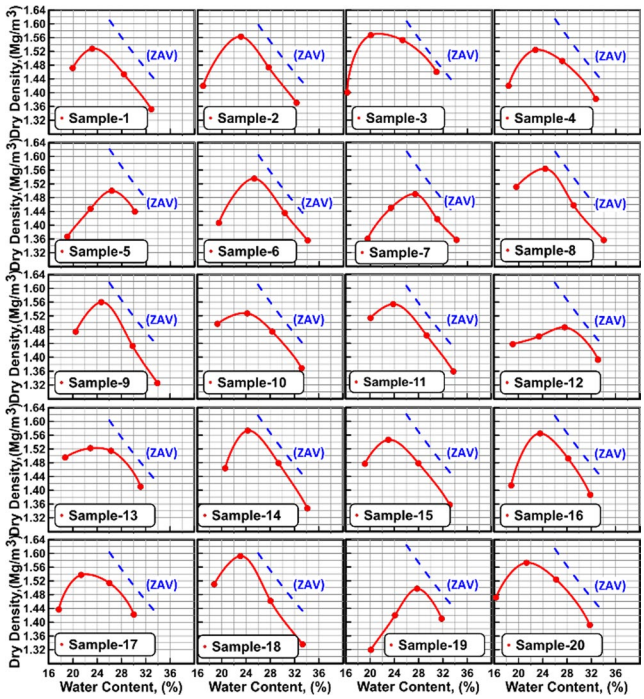


Fig. 4 Standard Proctor compaction curves for all soil samples

and suction sensitivity in others. These mineralogical distinctions provide important context for interpreting variations in swelling behavior, compressibility, and shear strength across the soil dataset.

Compaction characteristics

Standard Proctor compaction tests (ASTM D698 2012) were conducted to determine the optimum water content (w_{opt}), maximum dry density (qdmax), specific gravity (Gs), and degree of saturation (S) at optimum conditions (Fig. 4; Table 1).

Table 1 Summary of standard Proctor compaction test results

Sample ID	q_{dmax} (Mg/m^3)	w_{opt} (%)	Gs	S (%)
1	1.528	23.10	2.77	78.72
2	1.563	23.50	2.73	85.92
3	1.578	21.60	2.76	79.59
4	1.528	23.60	2.78	80.07
5	1.500	26.40	2.80	85.29
6	1.535	25.60	2.75	88.94
7	1.491	27.40	2.78	85.46
8	1.565	24.00	2.78	85.94
9	1.560	24.40	2.79	86.34
10	1.527	24.20	2.77	82.35
11	1.556	23.40	2.76	83.47
12	1.489	27.00	2.79	86.22
13	1.526	24.80	2.75	85.03
14	1.573	24.50	2.79	88.35
15	1.546	23.20	2.80	80.09
16	1.568	24.00	2.79	85.92
17	1.542	22.40	2.75	78.63
18	1.590	22.80	2.74	89.82
19	1.497	27.50	2.81	82.04
20	1.574	22.00	2.78	79.82

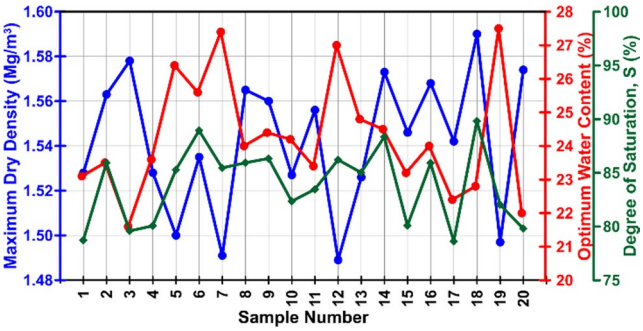


Fig. 5 Comparative plot of maximum dry density, optimum water content, and saturation degree

The q_{dmax} values ranged from 1.489 to 1.590 Mg/m³, with w_{opt} values between 21.6 and 27.5%, consistent with the high plasticity and elevated moisture demand of clay-rich soils. Degrees of saturation at optimum conditions ranged between 78.63 and 89.82%, confirming that specimens compacted under Proctor energy were still in an unsaturated state. In Fig. 5, the changes in compaction characteristics and saturation degree for each soil sample are shown. These results allow a comparative evaluation of how density, water content, and saturation vary among the samples.

Methods

Nomenclature

P200 Percent finer than No. 200 sieve (%)

q_{dmax} Maximum dry density (Mg/m³)
w_{opt} Optimum water content (%)
Swell (%) Vertical swelling strain (%)
Swell (kPa) Swell pressure (kPa)
Cc Compression index (–)
Cs Swell index (–)
m_v Coefficient of volume change (m²/MN)
q_{un}/UCS Unconfined compressive strength (kPa)
c Apparent cohesion (kPa)
φ Friction angle (°)
p Mean total stress (kPa)
q Deviator stress (kPa)
P_{break} Mechanical breakpoint stress (kPa)
h,Ψ Matric suction (kPa)
θ(h) Volumetric water content at suction h (–)
θs Saturated volumetric water content (–)
AEV Air entry value (kPa)
a_z, n_z, m_z Zapata model parameters (–)
a_p, b_p, c_p Perera model parameters (–)
hr Residual suction (kPa)
C(h) Correction function in Fredlund and Xing model (–)
PC1, PC2, PC3 Principal components (–)
μi Cluster centroid (–)
Ci Cluster i (–)
k Number of clusters (–)

All laboratory tests were conducted under controlled isothermal conditions of 20 ± 1 °C, as temperature significantly influences water retention, suction behavior, and shear strength of fine grained soils. Specimens were prepared using the maximum dry density (q_{dmax}) and optimum water content (w_{opt}) obtained from Standard Proctor compaction to ensure consistent initial conditions across tests.

In this study, a series of laboratory experiments were conducted to investigate the geomechanical and volumetric behavior of high plasticity soils. The experimental program consisted of four main test groups: (1) Swelling percentage and swelling pressure tests, (2) One–dimensional consolidation tests, (3) Unconfined compression tests, and (4) Unconsolidated–undrained (UU) triaxial compression tests.

Swell percentage and swell pressure tests

Swelling behavior was evaluated using a computer-controlled swell apparatus developed by Türköz (2019), which allows simultaneous measurement of both swell percentage and swell pressure under controlled loading. Cylindrical specimens were compacted into thin-walled rings with an internal diameter of 7 cm and a height of 2 cm at their respective q_{dmax} and w_{opt} values. In the swell percentage test, a constant vertical stress of 7 kPa was applied to simulate light overburden. Vertical deformation was recorded at fixed intervals using a high-resolution digital displacement sensor. The swell percentage was calculated as the ratio of vertical deformation to the initial specimen height. For swell pressure, specimens were restrained vertically and allowed to absorb water freely. The resulting swelling force was measured using a calibrated proving ring and converted to stress. The swelling tests were used primarily for characterizing the soils' hydro-reactivity; however, they are not employed in the central dual-threshold (AEV–P_{break}) analysis.

One-dimensional consolidation test

Compressibility behavior was evaluated using a standard fixed-ring oedometer apparatus following ASTM D2435 (2020). Specimens were compacted into rigid steel rings with a diameter of 5 cm and a height of 2 cm, ensuring that testing conditions matched the compaction state used in all other experiments. A loading sequence of 100, 200, 400, 800, and 1600 kPa was applied, followed by unloading to 800, 400, and 100 kPa. Vertical deformation was recorded continuously and interpreted using the logarithm-of-time method. From the results, the following parameters were determined: compression index (Cc), swelling index (Cs), and coefficient of volume change (m_v).

Unconfined compression test

The unconfined compression tests followed ASTM D2166 (2016). Cylindrical specimens (5 cm diameter × 10 cm height) were compacted into stainless steel molds and extruded immediately prior to testing. Axial load was applied under displacement control at a rate of 1 mm/min until specimen failure or approximately 15–20% axial strain. The peak axial stress was recorded as the unconfined compressive strength (q_{un}).

Because the specimens were unsaturated, q_{un} was not converted to undrained shear strength using $Su = q_{un}/2$, which is only valid for fully saturated clays. Instead, q_{un} is treated as a direct indicator of compressive strength and incorporated into the statistical analyses.

Unconsolidated–Undrained (UU) triaxial compression test

UU triaxial compression tests were conducted to evaluate the shear strength behavior of the soils under total stress conditions. All UU tests followed ASTM D2850 (2024). Specimens were statically compacted into 5 cm diameter × 10 cm high steel molds, identical to the UC test setup. Initial tests were performed at 50, 100, 200, and 400 kPa confining pressures. However, to capture the nonlinear strength

behavior due to unsaturated conditions, a connected bilinear envelope approach in p–q space was adopted instead of a single linear envelope. To clearly identify the breakpoint and extend both segments of the envelope, additional tests were conducted at 25 kPa and 600 kPa.

In classical soil mechanics, the Mohr–Coulomb failure criterion has been widely used to characterize the shear strength behavior of soils, often assuming linearity across all stress levels. However, numerous studies have shown that the shear strength envelope of fine-grained soils, especially under undrained conditions, may not remain linear across all confining pressures. Instead, a bilinear or nonlinear failure envelope often emerges due to changes in the governing failure mechanisms at low and high stress levels (Lade and Duncan 1973; Houlsby 1997).

In the analysis of undrained triaxial (UU) test results, the stress trace method (p–q space) is an effective way to interpret shear strength behavior of soils. Rather than relying on a single linear Mohr–Coulomb failure envelope, a connected bilinear model allows for a more accurate representation of strength parameters in cases where a single envelope is insufficient due to changes in mechanical behavior at different confining pressures. The optimized R^2 approach aims to identify the point of transition (breakpoint) between two linear segments such that the combined coefficient of determination (R^2) of the two linear regressions is maximized.

To capture the nonlinear strength response, connected bilinear envelopes were constructed using an optimized coefficient of determination (R^2) approach, as follows:

1. Plot all test points in p–q space for each soil sample.
2. Evaluate all possible breakpoints, dividing the dataset into a low-stress and high-stress segment.
3. Fit linear regressions to each segment:

– Segment 1: $q = m_1 p + b_1$.
– Segment 2: $q = m_2 p + b_2$.

4. Compute R^2 values for both segments.
5. Select the breakpoint that maximizes the combined R^2 ($R^2_1 + R^2_2$).
6. Ensure continuity, connecting the two segments smoothly at the breakpoint.
7. Convert the slopes to Mohr–Coulomb parameters if required: $\varphi = \tan^{-1}(m)$, $c = b/\tan(\varphi)$.

The stress corresponding to the optimal break point is defined as the mechanical transition threshold (P_{break}). This parameter is a central component of the dual-threshold framework and is later compared with the AEV from SWCC modeling. The overall methodology provides a more nuanced understanding of shear strength development, especially in compacted soils influenced by suction. A typical connected bilinear failure envelope derived using this approach is shown in Fig. 6.

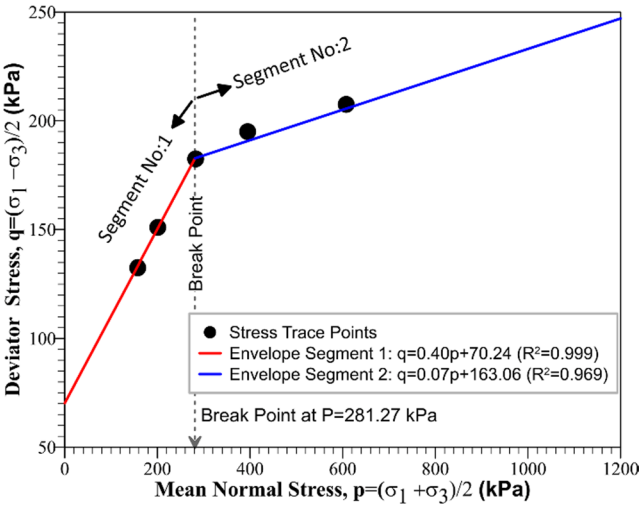


Fig. 6 A typical connected bilinear failure envelope in p-q space

Soil-water characteristic curve (SWCC) modeling and AEV estimation

SWCCs were estimated using the Fredlund and Xing (1994) equation as the foundational model (Eq. 1). Two empirical parameter estimation approaches were used: one proposed by Zapata et al. (2000) and the other by Perera et al. (2005). Both models relate SWCC parameters to the plasticity index (PI) and percent finer than No. 200 sieve (P200), enabling prediction of SWCC behavior without suction-controlled tests.

$$\theta(\psi) = C(\psi) \cdot \left[\frac{\theta_s}{\left[\ln\left(e + \left(\frac{\psi}{a} \right)^n \right) \right]^m} \right] = \left[1 - \frac{\ln\left(1 + \frac{\psi}{C_r} \right)}{\ln\left(1 + \frac{10^6}{C_r} \right)} \right] \cdot \left[\frac{\theta_s}{\left[\ln\left(e + \left(\frac{\psi}{a} \right)^n \right) \right]^m} \right] \quad (1)$$

where:

$\theta(\Psi)$ volumetric water content at suction Ψ,

θ_s saturated volumetric water content,

Ψ or h: suction (kPa),

a, n, m fitting parameters controlling air entry, slope, and residual shape,

$C(\Psi)$ Correction factor to handle residual behavior,

C_r suction value corresponding to residual conditions (same as h_r).

The parameter estimation equations used in this study for the Zapata et al. (2000) model were adopted exactly as originally proposed, with only a mathematical conversion applied to express base-10 logarithmic terms in natural logarithm (ln) form for consistency. No modifications or recalibrations were made to the original coefficients.

Both Zapata et al. (2000) and Perera et al. (2005) proposed empirical correlations to estimate the SWCC parameters based on the plasticity index (PI) and the percent finer than the No. 200 sieve (*P200*). These were often combined into the product $wPI = PI \times P200$, which served as a single input variable.

For Zapata et al. (2000), the empirical equations using natural logarithms (ln) are:

$$\ln(a_z) = [-1.0846 + 0.01335\,(wPI)]\,2.3026 \tag{2}$$

$$n_z = 1.412 + 0.00148\,(wPI) \tag{3}$$

$$m_z = 0.2049 + 0.00277\,(wPI) \tag{4}$$

$$\ln(h_{rz}) = [2.4121 + 0.00466\,(wPI)]\,2.3026 \tag{5}$$

For Perera et al. (2005), the fitted expressions were:

$$\ln(a_p) = -1.253 + 0.01878\,(wPI) \tag{6}$$

$$b_p = 1.396 + 0.00052\,(wPI) \tag{7}$$

$$c_p = 0.202 + 0.00341\,(wPI) \tag{8}$$

$$h_{rfp} = 500 \;\; (assumed\ cons\tan t\ across\ samples) \tag{9}$$

These equations were applied to each of the 20 soil samples to compute the corresponding Fredlund and Xing (1994) model parameters. The Perera model was plotted as saturation (%) versus matric suction (kPa), whereas the Zapata model was plotted as volumetric water content (θ) versus matric suction (kPa), in accordance with the representation used in the original studies. SWCC curves were generated in the suction range from h = 0.1 to 1,000,000 kPa.

Although several studies have explored optimization-based approaches to recalibrate the empirical equations of Zapata et al. (2000), including polynomial, exponential, or ratio-based formulations tailored to specific soil datasets, the present study deliberately retained the original form and coefficients of the Zapata model without modification.

Recent work by Uzundurukan (2023) employed a particle swarm optimization (PSO) algorithm to recalibrate the empirical parameters of the Zapata et al. (2000) model for plastic soils. The study demonstrated that advanced optimization techniques could improve the prediction accuracy of the SWCC, particularly for fine-grained soils. While such approaches are promising, the present study chose to retain the original form and coefficients of the Zapata model without modification, in order to ensure generalizability and enable direct comparison with other standardized models.

The Fredlund and Xing (1994) equation forms the basis for both Zapata et al. (2000) and Perera et al. (2005) models, with parameter estimation strategies being the main difference. Table 2 presents outlines how the original Fredlund and Xing parameters map to those used in the Zapata and Perera approaches.

A strict, reproducible tangent-intersection procedure was implemented:

Perera-based SWCC.

- Tangent in the near-saturation region: h = 0.1–10 kPa.
- Tangent in the desaturation region: h = 10^4–10^6 kPa.
- AEV = intersection point of the two tangents.

Zapata-based SWCC.

- Numerical derivative computed at all suction points.
- "Flat zone": local minima in |dθ/dh|.
- "Steepest zone": global maximum in |dθ/dh|.
- Tangents fitted automatically using a sliding-window algorithm.
- AEV = intersection point.

Statistical analyses

A structured multivariate statistical workflow was implemented to synthesize the geotechnical, hydraulic, and strength parameters obtained from the 20 compacted fine-grained soil samples. The analysis consisted of three stages: (i) data screening and normalization, (ii) dimensionality reduction using principal component analysis (PCA), and (iii) unsupervised classification using K-means clustering.

Table 2 Parameter equivalency between Fredlund and Xing (1994), Zapata et al. (2000), and Perera et al. (2005)

Fredlund and Xing (1994) Parameter	Zapata et al. (2000)	Perera et al. (2005)	Description
a	a_z	a_p	Controls the air-entry region (location of curve drop)
n	n_z	b_p	Controls slope in the transition zone
m	m_z	c_p	Controls the tail of the curve, the residual approach
h_r	h_{rz}	h_{rfp} = 500 kPa (fixed)	Residual suction – controls correction function *C(h)*

Data screening and normalization

All variables were initially examined for redundancy using Pearson correlation coefficients. A representative set of 14 parameters was retained, including plasticity indices (LL, PI, P200), compaction characteristics (maximum dry density, optimum water content), compressibility indices (Cc, Cs, m_v), swelling behavior, shear strength parameters (UCS, cohesion, friction angle), and hydro-mechanical thresholds (Perera based AEV and bilinear breakpoint stress from the UU envelopes).

To ensure equal weighting across variables with different physical units, all data were standardized using z-score normalization. This step prevents scale dominance and promotes balanced variable contributions to subsequent analyses, as recommended in recent methodological reviews of geotechnical data science workflows (e.g., Bozorgzadeh and Feng 2024).

Principal component analysis

Principal component analysis (PCA) was performed on the standardized dataset to reduce dimensionality and identify the dominant independent modes of variance. Component retention was guided by the Kaiser criterion (eigenvalue > 1), cumulative explained variance, and inspection of the scree plot. Based on these criteria, three principal components were selected for further analysis.

The retained components each represent an orthogonal axis of soil behavior.

- PC1 primarily reflects plasticity, fines content, and compressibility-related processes.
- PC2 is dominated by strength-related variables and hydro-mechanical transition indicators.
- PC3 is chiefly governed by friction-controlled mechanical response.

These components served as the reduced dimensional feature space for clustering and provided an interpretable structure linking index properties, suction behavior, and shear strength mechanisms.

K-means clustering

Clustering was conducted in the PCA-transformed space to ensure orthogonality among input variables. The optimal number of clusters (k = 3) was determined using the elbow method and silhouette scores. Cluster solutions were stabilized using 10 random initializations. Convex hulls were applied only for visualization of cluster extents, not as a statistical criterion. The clustering provides a data-driven classification of soils based on integrated hydraulic–mechanical behavior.

The K-means algorithm partitions the dataset by minimizing the total within-cluster sum of squared distances:

$$Minimize:\ J = \sum_{i=1}^{k} \sum_{x \in C_i} \|x - \mu_i\|^2 \tag{10}$$

Where:

- x is a soil sample described in d-dimensional space,
- C_i is the i-th cluster,
- μ_i is the centroid of cluster C_i, and.
- $\|x - \mu_i\|^2$ is the squared Euclidean distance.

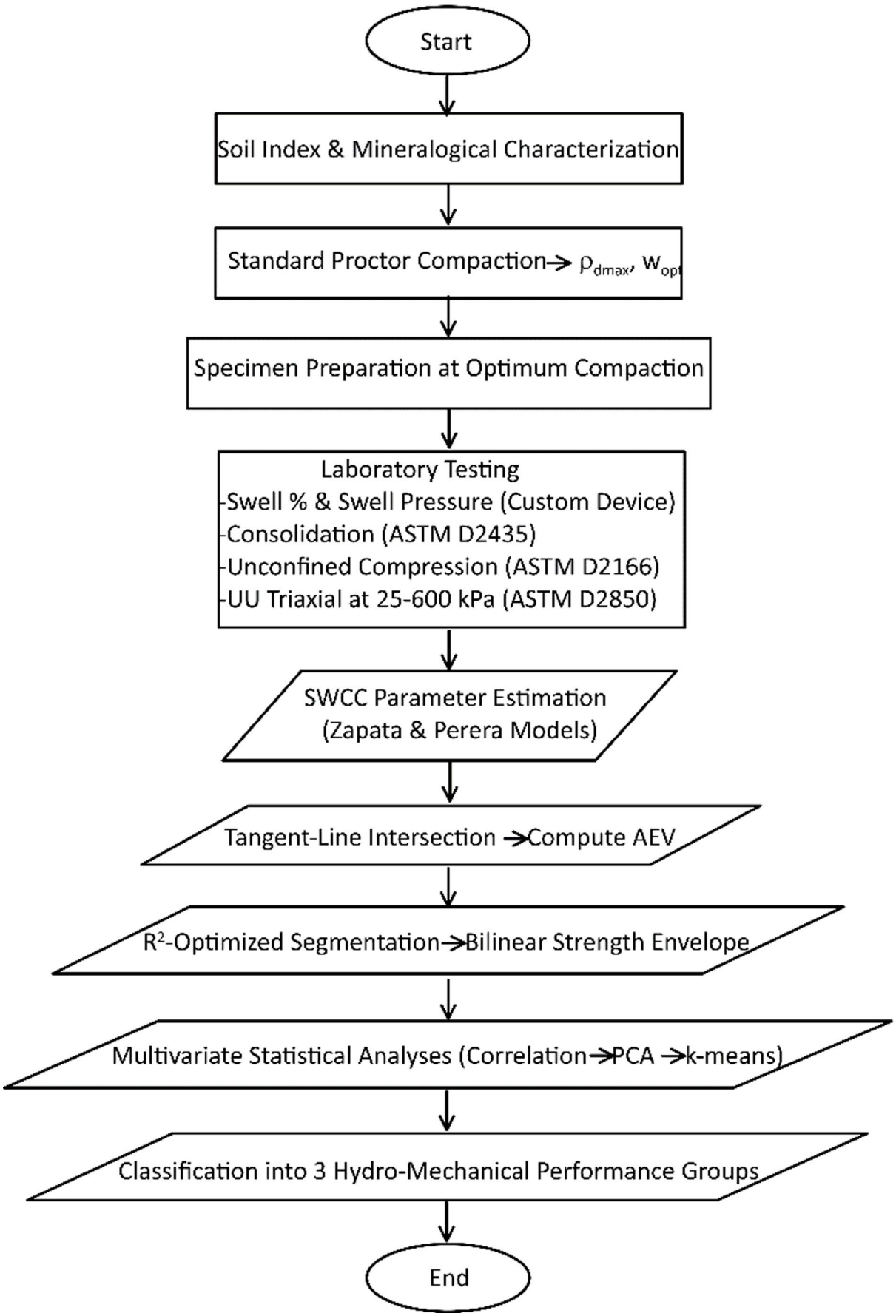
The algorithm proceeds iteratively:

1. Randomly initialize k centroids;
2. Assign each observation to the nearest centroid;
3. Recalculate centroids as the mean of assigned samples;
4. Repeat steps 2–3 until assignments converge.

This methodological structure was intentionally designed to align with recent recommendations promoting transparent, reproducible, and data-centric evaluation frameworks in geotechnical machine-learning applications. In particular, the workflow follows the guidance of Bozorgzadeh and Feng (2024), who emphasize the importance of clearly justifying algorithm selection, documenting preprocessing steps, and ensuring that model interpretation remains grounded in geotechnical domain knowledge. This approach produced a partition of the soil dataset that highlights differences and similarities among samples based on their engineering behavior. The use of PCA integrated with K-means clustering is also supported by prior geotechnical studies. For example, Goktepe et al. (2005) employed K-means to classify stabilized soils based on their mechanical properties, while Ozcoban et al. (2018) used clustering to group compacted clays according to their permeability and suction characteristics. Together, these applications underscore the suitability of unsupervised statistical methods for identifying behavioral patterns in fine-grained soils.

A graphical summary of the complete experimental and analytical workflow used in this study is presented in Fig. 7 to support methodological transparency and reproducibility.

Fig. 7 Overall methodological workflow used in this study


Start
Soil Index & Mineralogical Characterization
Standard Proctor Compaction→ ρ_{dmax}, w_{opt}
Specimen Preparation at Optimum Compaction
Laboratory Testing
-Swell % & Swell Pressure (Custom Device)
-Consolidation (ASTM D2435)
-Unconfined Compression (ASTM D2166)
-UU Triaxial at 25-600 kPa (ASTM D2850)
SWCC Parameter Estimation
(Zapata & Perera Models)
Tangent-Line Intersection →Compute AEV
R^2-Optimized Segmentation→Bilinear Strength Envelope
Multivariate Statistical Analyses (Correlation→PCA →k-means)
Classification into 3 Hydro-Mechanical Performance Groups
End

Results and discussion

Swell test results

Figures 8 and 9 present the swell percentage–time and swell pressure–time relationships for the compacted soil specimens. Although all samples were prepared at their respective optimum water contents, the swelling response exhibits substantial variation across the dataset, governed primarily by fines content and plasticity. The measured swell percentages range from 1.1% to 4.9%, with the highest values observed in soils containing more than 90% fines and a plasticity index between 36% and 39%. These results reflect the well-established influence of clay activity and surface area on wetting-induced volume change in compacted fine-grained soils. The corresponding swell pressures, shown in Fig. 9, follow a similar pattern. Samples with higher plasticity mobilize significantly larger pressures during inundation. Notably, Samples 3, 17, and 20 developed swell pressures exceeding 45 kPa, whereas soils with PI≤28% generated considerably lower pressures, consistent with their reduced clay activity. This behavior can be attributed to the reduction of matric suction following wetting, which induces diffuse double-layer expansion and interparticle repulsion within the clay-rich samples. Even under optimum compaction, these soils retain appreciable suction, enabling the mobilization of demonstrable swelling stresses.

Figure 10 shows the direct relationship between swell pressure and swell percentage. The clear positive correlation indicates that soils undergoing larger volumetric expansion also mobilize higher swelling stresses. This relationship highlights the role of microstructure: soils with a finer pore network and more dispersed clay fabric characteristic of high-PI samples at w_{opt} undergo greater wetting-induced expansion and mobilize higher repulsive forces upon saturation. Conversely, lower plasticity soils form denser, more aggregated structures during compaction, limiting both swelling strain and swelling pressure.

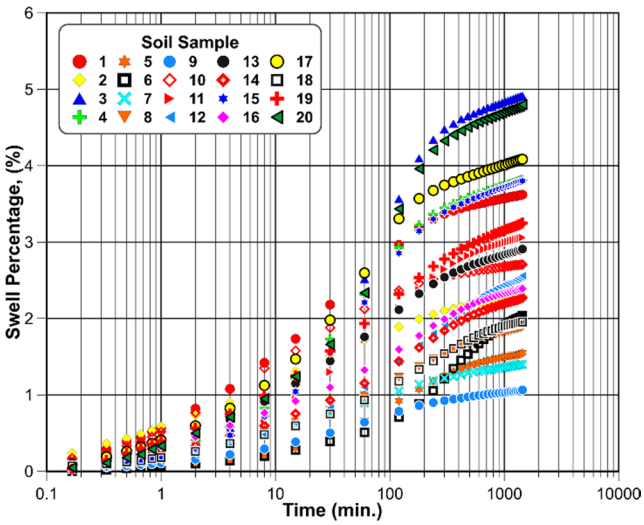


Fig. 8 Swell percentage vs. time for 20 fine-grained soil samples

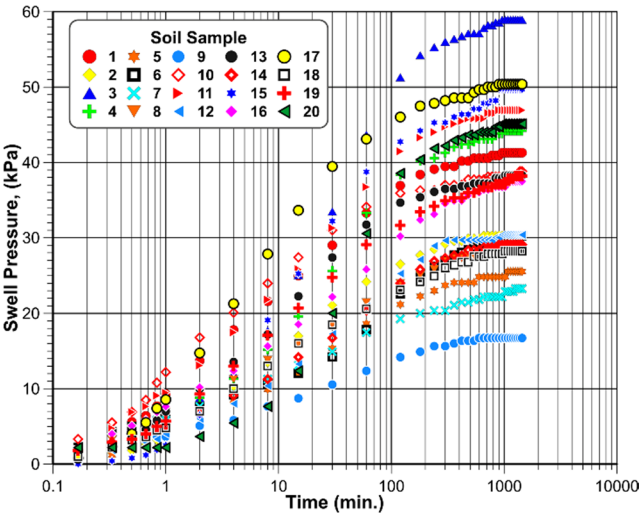


Fig. 9 Swell pressure vs. time for 20 fine-grained soil samples

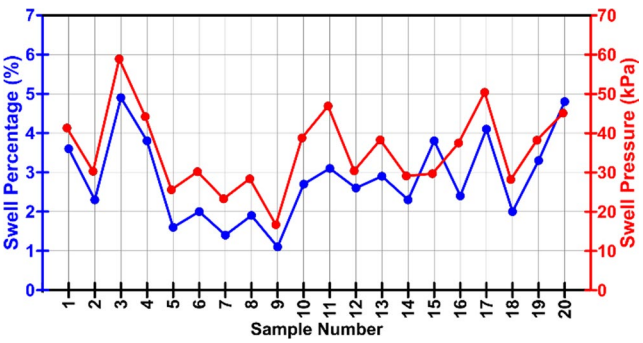


Fig. 10 Relationship between final swell pressure and swell percentage

The swelling behavior provides fundamental insight into the soils' hydro-mechanical evolution. Samples exhibiting greater swelling typically correspond to higher compression indices (Cc) and larger coefficients of volume change (m_v) in the consolidation tests (Sect. 4.2), reflecting their more deformable microstructure. The same samples also tend to exhibit lower undrained shear strength in the UU tests (Sect. 4.4), attributable to the suction loss and microstructural softening associated with wetting. Furthermore, the expansive samples generally exhibit higher air entry values (AEV) in the SWCC results (Sect. 4.5), consistent with a finer pore size distribution that sustains suction during desaturation.

In summary, the swell test results demonstrate that key intrinsic soil properties, particularly fines content, plasticity, and micro fabric established at optimum water content, govern both the magnitude of wetting-induced deformation and the stresses developed during swelling. These findings

form an essential basis for interpreting consolidation behavior, shear strength, and hydraulic response in the subsequent sections.

Consolidation test results

The consolidation parameters obtained from the oedometer tests are presented in Figs. 11, 12 and 13. The compression index (Cc) values (Fig. 11) range from approximately 0.18 to 0.32, with higher values generally associated with soils containing greater fines content and higher plasticity. These soils possess more deformable microstructures and are more susceptible to particle reorientation under applied loads.

The swell index (Cs) values (Fig. 12) exhibit a similar pattern, varying between 0.04 and 0.07. Soils with higher PI and higher clay fractions consistently exhibit larger Cs values, reflecting increased susceptibility to recompression following unloading. The similarity in trends between Cc and Cs indicates that a common set of microstructural controls, primarily clay content, plasticity, and pore-size distribution, governs both the initial swelling response (Sect. 4.1) and the subsequent compressibility during one-dimensional loading.

The coefficient of volume change, m_v, was evaluated over four pressure ranges (100–200, 200–400, 400–800, and 800–1600 kPa) and is shown in Fig. 13. Across these ranges, m_v typically falls between 0.03 and 0.14 m²/MN, with the highest values observed for samples with larger Cc and Cs. These soils demonstrate greater sensitivity to structure breakdown under loading, consistent with their higher swelling strains and swelling pressures. Even under higher stresses (800–1600 kPa), the compressibility of these samples remains elevated, indicating a more open or weaker clay fabric formed during compaction at optimum water content.

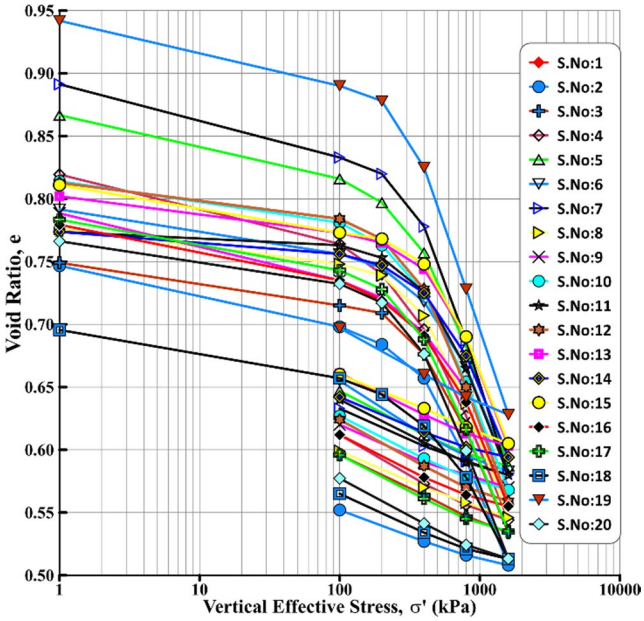


Fig. 11 Void ratio–log effective stress consolidation curves for 20 soil samples

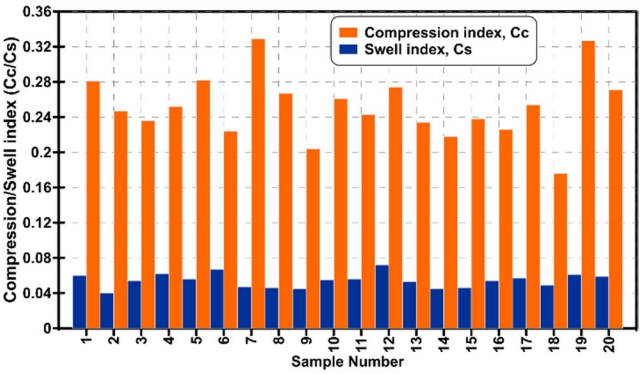


Fig. 12 Compression index (Cc) and swell index (Cs) variation by sample number

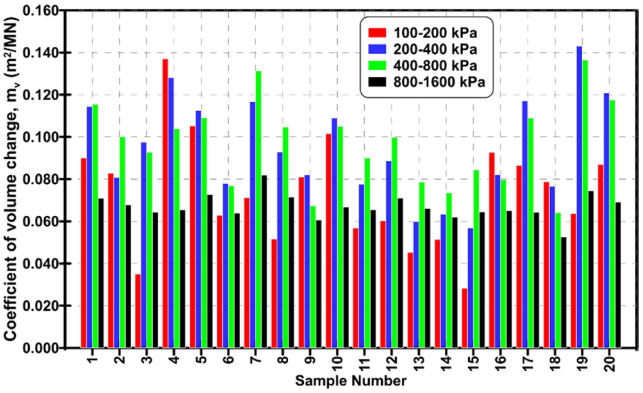


Fig. 13 Coefficient of volume change (m_v) across stress intervals for 20 soil samples

The consolidation behavior is strongly linked to the swelling response described in Sect. 4.1. Samples that exhibited higher swell percentages, such as those with swell strains approaching 5%, also showed among the highest Cc and m_v values. This correspondence aligns with fundamental suction mechanics: soils that undergo greater wetting-induced expansion typically possess finer pore networks and more active clay minerals, which not only swell upon saturation but also compress more readily under load due to suction loss and microfabric softening.

These consolidation trends also have implications for shear strength behavior. Soils with higher compressibility indices tend to exhibit lower undrained shear strength in both unconfined compression and UU triaxial tests. Conversely, samples with lower Cc and m_v values, indicative of denser and more aggregated microstructures demonstrate greater resistance to deformation and correspondingly higher shear strength.

Finally, the consolidation characteristics contribute to the interpretation of hydraulic behavior observed in the soil water characteristic curves (Sect. 4.5). Soils with higher compressibility often exhibit higher air entry values (AEV), consistent with finer effective pore sizes that sustain matric suction over a broader suction range. This reinforces the interconnected hydraulic–mechanical behavior explored further in Sect. 4.5, where the relationship between AEV and the mechanical breakpoint stress is examined.

Unconfined compression test results

Figure 14 illustrates the variation in unconfined compressive strength (UCS) for the compacted soil samples. The measured UCS values range from 107 kPa to 250 kPa, indicating considerable differences in undrained shear resistance among the specimens. The lowest strength was recorded in Sample 7 (107 kPa), while Sample 20 (250 kPa) exhibited the highest strength. This range reflects variations in fines content, plasticity index, and the soil fabric formed during compaction at optimum water content.

The UCS trends are consistent with suction-controlled behavior in compacted fine-grained soils. Zhang et al. (2018) demonstrated that the undrained strength of stiff clays is influenced by the void ratio corresponding to the main air entry value, with the controlling mechanism transitioning depending on saturation. When the degree of saturation exceeds the AEV threshold, the void ratio of water governs strength; when saturation is lower, the void ratio of air becomes dominant. This interpretation helps explain the plateauing of strength in soils where suction gradually dissipates during loading. Chae et al. (2010) further highlighted that suction stress arising from both meniscus and bulk-water effects acts as an internal confining pressure that enhances apparent strength. These findings support the interpretation that suction plays a significant role in stabilizing the more plastic and fine-grained samples tested in this study.

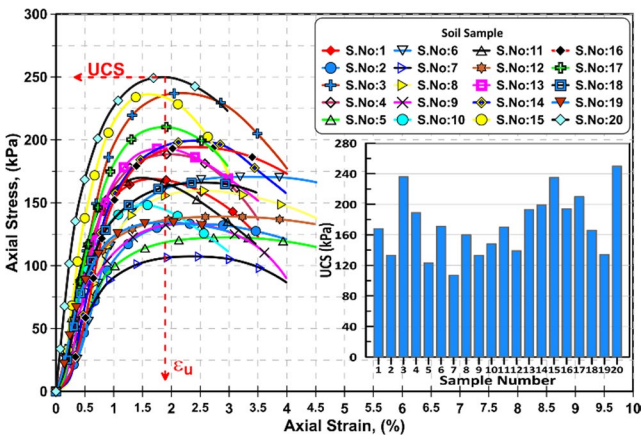


Fig. 14 Variation of unconfined compressive strength (UCS) across 20 soil samples

The UCS characteristics also reflect differences in compaction-derived microstructure. Specimens with higher UCS values such as Samples 15, 17, and 20 tend to correspond to soils with higher dry densities and moderate water contents at molding, conditions that promote a denser and more aggregated structure capable of resisting undrained loading. Lower UCS values are typically associated with higher optimum moisture contents or higher plasticity, which result in a softer and less cohesive fabric under short-term shear.

Interactions between UCS and deformation-related properties are also evident. For instance, Samples 3 and 15, which achieved relatively high UCS values and moderate swelling characteristics, suggest a favorable combination of stiffness and volumetric stability. In contrast, soils such as Sample 7, which combine low UCS with elevated swell pressure or high m_v values, represent materials with limited strength and heightened deformation susceptibility.

The unconfined compression results provide a direct measure of short-term undrained shear resistance, complementing the volumetric responses identified in the swelling (Sect. 4.1) and consolidation (Sect. 4.2) tests. These UCS values will subsequently be integrated into the clustering analysis to classify soils according to their combined strength, compressibility, and hydraulic characteristics.

UU Triaxial test results

Figure 15 presents the p–q failure envelopes obtained from unconsolidated–undrained (UU) triaxial tests. A connected bilinear model was fitted to each envelope to represent two distinct strength regimes corresponding to low and high confining pressures. These two linear segments capture the transition in shearing mechanism as pore pressure effects intensify under increasing confinement. The bilinear approximation yielded excellent agreement with experimental data across the dataset, with R^2 values exceeding 0.95 for nearly all samples. Breakpoint stresses, marking the transition between the two linear regimes, varied considerably. For example, Sample 3 exhibited a clear breakpoint at 665.5 kPa, whereas Sample 2 displayed a more continuous envelope without a pronounced inflection, suggesting a gradual rather than abrupt shift in failure mechanism.

These transitions reflect suction-controlled behavior, consistent with previous findings on unsaturated soil strength. Abd et al. (2020) reported that shear strength–suction relationships commonly deviate from linearity at a critical matric suction value, coinciding with a shift in failure mode as capillary forces diminish. Pande et al. (2024) similarly demonstrated that the air entry value marks the onset of nonlinearity in strength development and corresponds to the stress state at which suction-driven

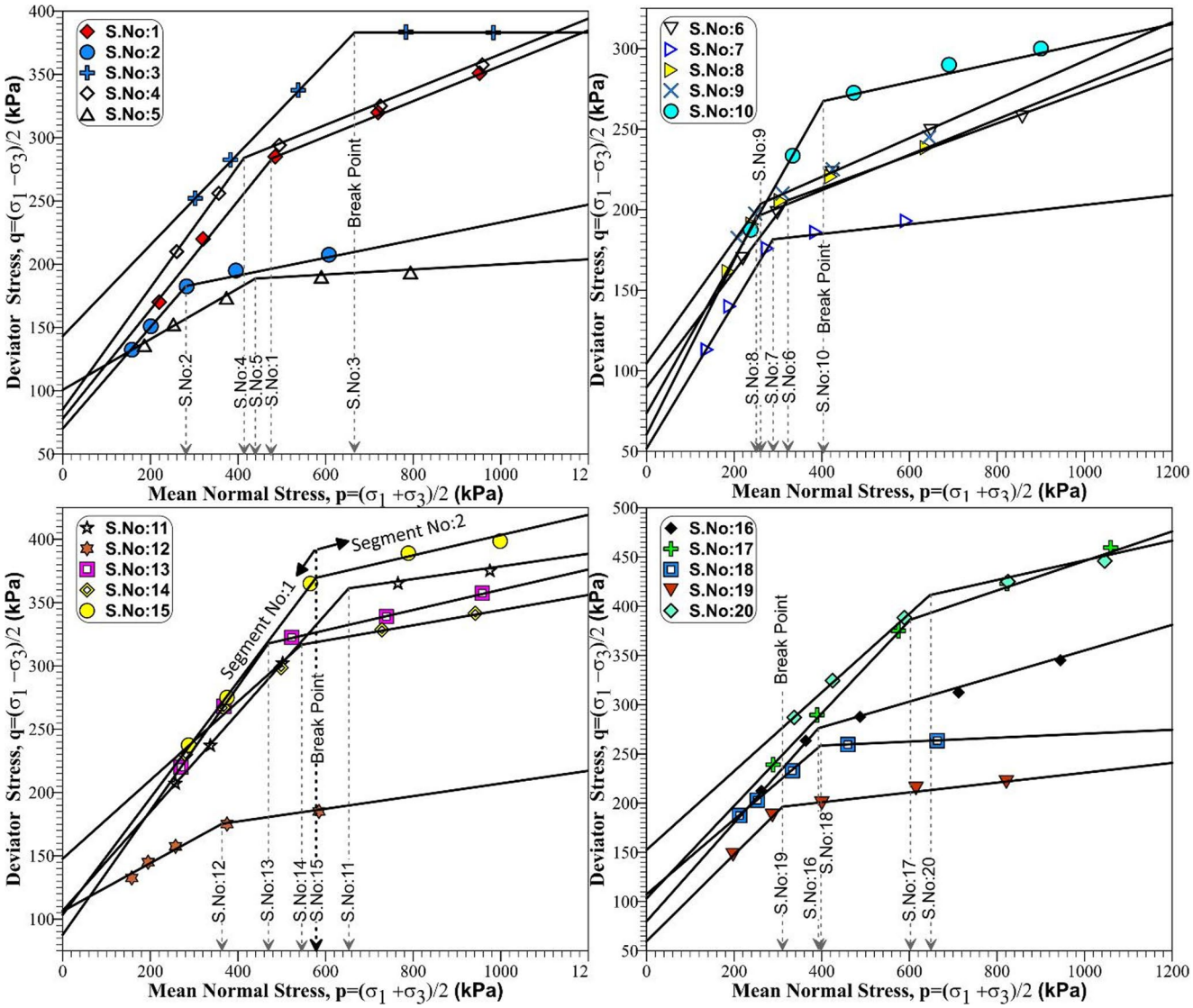


Fig. 15 p–q bilinear failure envelopes for 20 soil samples based on UU triaxial tests

stiffening begins to weaken. The behavior observed in the present study aligns closely with these findings: sharper breakpoints were typically observed in denser specimens with more structured fabrics, whereas more gradual transitions appeared in soils exhibiting less microstructural cohesion.

Figure 16 illustrates cohesion values for the low and high pressure segments of the bilinear envelopes. Most samples demonstrated substantial increases in apparent cohesion with increasing confinement. For example, Sample 1 exhibited an increase from 86.45 kPa to 219.03 kPa, and Sample 3 increased from 153.52 kPa to 383.00 kPa. This trend reflects the elevated pore pressure and suction stress effects generated under undrained loading, which increase the apparent effective cohesion as drainage becomes increasingly restricted.

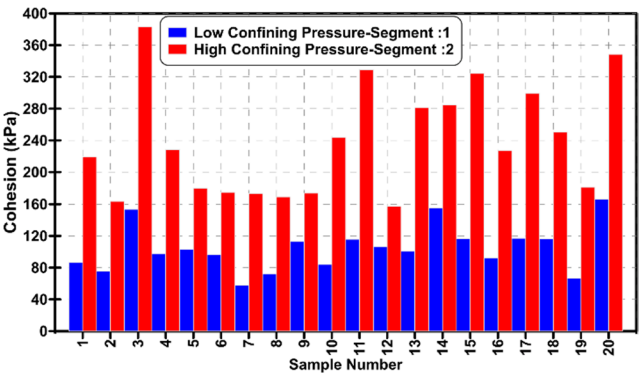


Fig. 16 Variation in cohesion values for low and high confining pressure segments

The variation in friction angle (φ) between the two confining pressure regimes is shown in Fig. 17. In contrast to cohesion, φ values generally decreased with increasing confining pressure. Some samples experienced substantial reductions; for instance, Sample 8 decreased from 32.68° to 6.32°, and Sample 3 reached 0° in the high-pressure segment. This behavior indicates a shift toward a cohesion-dominated failure mechanism as pore pressures increase, diminishing frictional resistance and suppressing dilation. Such reductions in φ are typical for fine-grained soils undergoing undrained shear under elevated confinement.

These observations confirm the ability of the Optimized R^2 Connected Bilinear Envelope Approach to represent the dual strength mechanisms active under UU loading conditions. The method effectively identifies:

- the transition from friction-dominated to cohesion-dominated behavior,
- the stress levels associated with microstructural yielding or pore pressure–induced softening, and.
- the variation in strength parameters (c, φ) across confining pressure regimes.

Furthermore, the strength characteristics revealed through the bilinear envelopes are consistent with the findings of Niu et al. (2024), who reported a suction-driven increase in apparent cohesion in the post-break regime, supporting the interpretation that suction-induced stiffening plays a critical role in the observed transitions.

The connected bilinear envelopes developed in this study also align with the theoretical understanding of nonlinear shear strength in unsaturated soils. Although the bilinear model simplifies strength behavior into two linear segments, it serves as a practical and physically meaningful approximation of a fundamentally continuous nonlinear relationship. Classical power-type formulations such as those proposed by Khalili and Khabbaz (1998) and Lu and Likos (2004) describe shear strength as a smooth function of both mean effective stress and

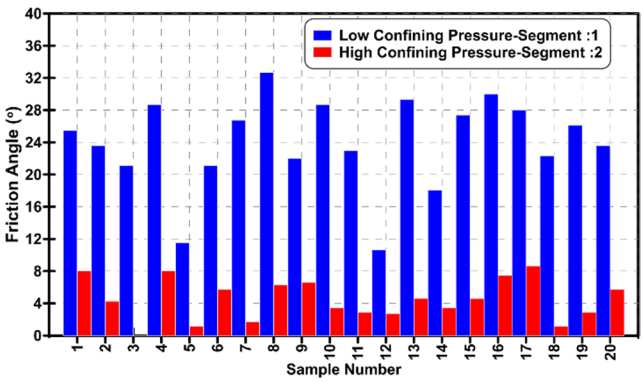


Fig. 17 Variation in friction angle values for low and high confining pressure segments

suction-dependent effective stress, capturing suction induced apparent cohesion through curvature rather than segmentation. Within this framework, the bilinear representation used here can be interpreted as a simplified means of identifying regime transitions: the first segment represents the friction-dominated state where suction incrementally contributes to shear resistance, while the second segment represents a suction stabilized, cohesion-controlled regime under elevated pore pressures. Thus, the connected bilinear model provides a transparent and nonparametric alternative that remains consistent with widely accepted suction driven strength theories.

SWCC results

Figure 18 shows the soil water characteristic curves (SWCCs) fitted using the Perera et al. (2005) model. The Perera formulation successfully captured the suction-degree of saturation relationship for these plastic, fine-grained soils, and AEVs were determined using the tangent-line intersection method. This graphical approach is theoretically consistent with recent formulations proposed by Zhai et al. (2019) and Quevedo et al. (2024), who highlighted the importance of identifying suction thresholds to characterize unsaturated shear response. The AEVs obtained from the Perera curves ranged from 263.26 kPa (Sample 2) to 632.78 kPa (Sample 17), reflecting the strong suction retention capacity typical of high-plasticity clays. Samples 17, 19, and 8, all with AEVs exceeding 590 kPa, correspond to soils with high fines content and plasticity indices, typical of microstructures capable of sustaining large matric suction before air intrusion.

Figure 19 presents the SWCCs fitted using the Zapata et al. (2000) model. Although the general curve shapes were similar, the AEVs were systematically lower, ranging from 52.11 kPa (Sample 2) to 205.68 kPa (Sample 17). Table 3 summarizes the comparison between the two modeling approaches. The mean AEV for the Perera model was 488.3 kPa, whereas the mean for the Zapata model was 123.8 kPa, indicating a fourfold reduction on average. These lower AEVs reflect the more empirical nature of the Zapata formulation and its tendency to underpredict suction thresholds in active clay soils where microstructural effects dominate water retention. Consequently, the Perera model is more appropriate for the soils tested in this study, consistent with findings from prior work on compacted plastic clays.

A key observation arises when comparing AEVs with the mechanical break points obtained from the UU triaxial tests (Table 3). Several samples exhibited close correspondence between the Perera-based AEV and the breakpoint stress (p or σ_n(eq), implying a potential hydro mechanical coupling. For example, Sample 20 showed a breakpoint stress of 647.07 kPa and an AEV of 589.93 kPa, while Sample 3 demonstrated a breakpoint of 665.47 kPa and an AEV of 290.61 kPa. These

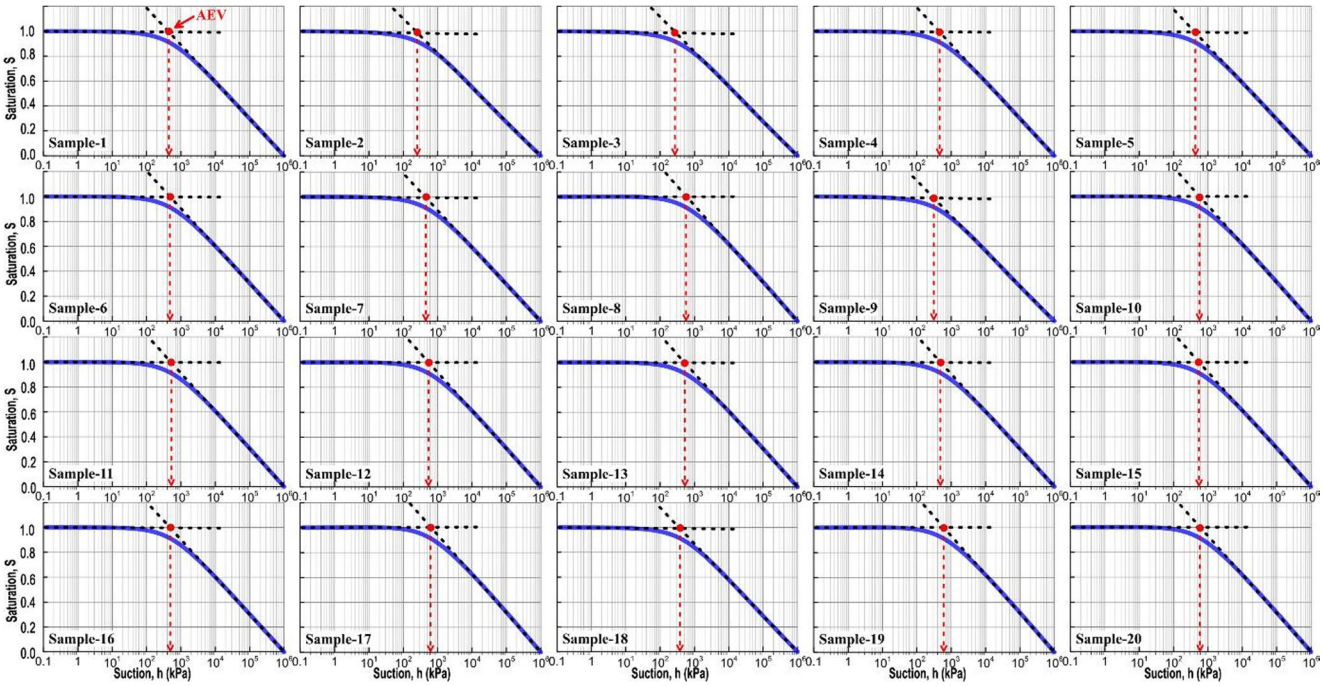


Fig. 18 SWCC curves for 20 soil samples modeled using the Perera et al. (2005) equation

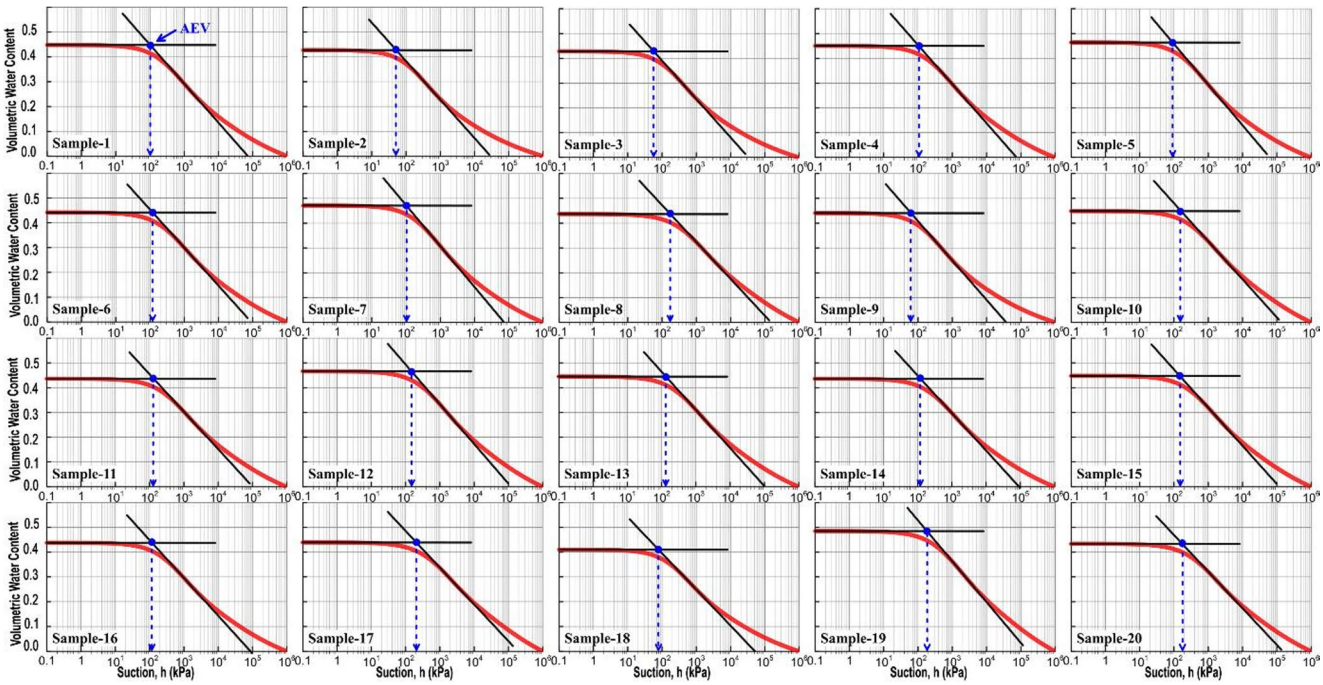


Fig. 19 SWCC curves for 20 soil samples modeled using the Zapata et al. (2000) equation

relationships suggest that the suction threshold at which air begins to enter the soil pores may coincide with the stress level at which shear strength transitions from friction-dominated to suction-dominated behavior during undrained loading. Quevedo et al. (2024) reported similar coupling between suction evolution and structural transitions, supporting the interpretation that AEV can serve as an indicator of mechanical regime change.

This relationship is shown explicitly in Fig. 20, which plots AEV against the bilinear break point stress (P_{break}) together with a 1:1 reference line. Most samples fall close to or slightly above the reference line, indicating that P_{break} generally exceeds AEV,

Table 3 Comparison of Air Entry Values for 20 soil samples using the Perera and Zapata models, with corresponding triaxial breakpoint pressures

Sample No	(q=mp+*n*)	(τ=c+σ·tan(φ))	Zapata et al. (2000)	Perera et al. (2005)
	Break point *p* (kPa)	Break point σ_n(eq) (kPa)	AEV (kPa)	AEV (kPa)
1	478.55	396.02	101.63	448.80
2	281.27	243.24	52.11	263.26
3	665.47	594.50	57.45	290.61
4	413.76	323.41	105.25	458.28
5	439.50	419.18	91.16	419.14
6	322.23	273.99	119.76	493.01
7	288.33	242.02	102.69	451.62
8	250.09	183.45	175.20	592.34
9	260.46	209.83	62.41	313.86
10	403.60	328.51	161.37	571.27
11	652.65	571.19	131.56	517.97
12	363.64	364.34	147.99	548.86
13	469.24	375.17	133.29	521.42
14	545.20	486.56	121.51	496.90
15	577.71	474.40	146.71	546.58
16	392.24	301.52	117.80	488.63
17	602.31	478.89	205.68	632.78
18	397.39	344.62	79.19	380.30
19	310.62	259.93	183.04	603.47
20	647.07	542.60	173.55	589.93

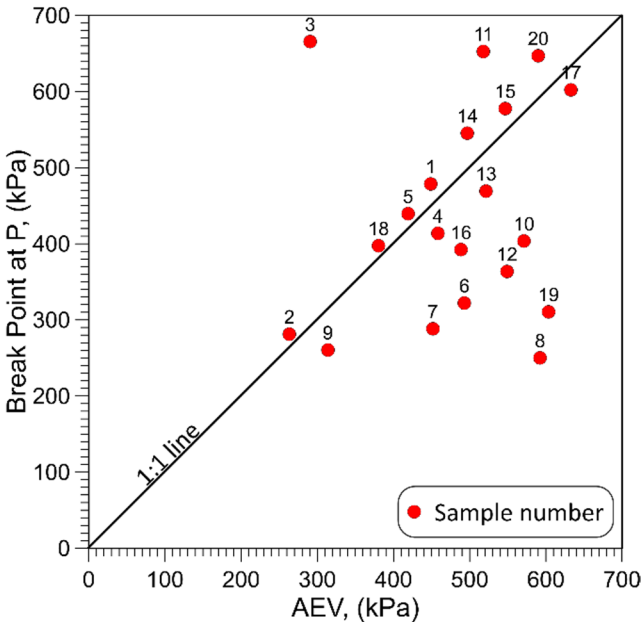


Fig. 20 Relationship between AEV (Perera model) and bilinear break point stress (P_{break}) from UU triaxial tests

meaning that desaturation typically occurs before the mechanical transition captured by the bilinear p–q response. High-plasticity soils such as Samples 17, 20, 11, and 15 cluster near the upper right portion of the plot, showing strong agreement between hydraulic and mechanical thresholds. In contrast, Sample 3 lies well above the 1:1 line, suggesting that structural rearrangement and pore-pressure buildup dominated its mechanical response even after the onset of air entry. This pattern reinforces the hydro mechanical interpretation: the onset of desaturation and the onset of mechanical regime transition are governed by closely related microstructural processes, but do not necessarily coincide perfectly for all soils.

This interpretation is reinforced by Xiao et al. (2023), who demonstrated that the shear strength of unsaturated soils exhibits nonlinear dependence on matric suction, with distinct transitions in failure mode at critical suction levels. Their findings showed that suction contributes disproportionately to apparent cohesion at high suction states and that the onset of nonlinearity in shear strength often corresponds to a hydraulic transition similar to the AEV. By framing suction as an internal state variable that governs both water retention and strength evolution, Xiao et al. (2023) provided theoretical justification for the observed alignment between AEVs and bilinear breakpoint stresses in the present study. The correspondence between these thresholds therefore reflects a broader hydro mechanical behavior where desaturation onset and failure-mechanism transition share a common microstructural origin.

The influence of fines and plasticity on AEV is consistent with the findings of Fondjo et al. (2021), who reported that fines content, PI, and shrinkage parameters exhibit predictive capacity exceeding 90% for AEV estimation. In the present dataset, soils with high PI and high clay fractions (e.g., Samples 17 and 19) exhibited the strongest water retention capacity, validating both the tangent-line estimation method and the suitability of the Perera model. Outliers in the dataset, where AEV did not scale directly with clay content, are consistent with Fondjo et al.'s observation that mineralogical variability, particularly illite content, plays a secondary role in AEV control.

The combined SWCC findings support three important conclusions. First, the Perera model provides reliable and physically grounded AEV estimates for plastic clays, whereas the Zapata model may be more suitable for silty or less active soils. Second, the correspondence between AEVs and UU breakpoint stresses indicates that hydraulic and mechanical transitions may stem from the same microstructural thresholds. Third, the AEV and suction-related parameters derived from the SWCCs are essential inputs for the clustering and classification presented later in this study, which evaluate hydro mechanical behavior in an integrated framework.

Statistical results

A comprehensive multivariate statistical analysis was performed to integrate the geotechnical, hydraulic, and strength parameters of the 20 compacted fine-grained soils. Principal Component Analysis (PCA) was applied to the z-score–standardized dataset, and the first three components, which

Table 4 Principal component loading matrix for the first three components (PC1–PC3), calculated from standardized geotechnical and hydraulic variables

Variable	PC1	PC2	PC3
LL (%)	0.765	0.511	0.100
PI (%)	0.565	0.717	0.221
P200 (%)	0.654	0.400	0.191
q_{dmax} (Mg/m^3)	−0.859	0.236	0.347
W_{opt} (%)	0.756	−0.524	−0.208
Compression index, Cc	0.830	0.006	−0.166
Swell index, Cs	0.487	0.359	−0.557
m_v (m^2/MN)_400–800 kPa	0.722	0.127	−0.126
Unconfined compressive strength (kPa)	−0.489	0.846	0.029
Breakpoint at p (kPa)	−0.381	0.789	−0.324
Perera- AEV (kPa)	0.665	0.670	0.231
Cohesion (kPa)-Low Confining Pressure	−0.639	0.567	−0.413
Friction Angle (°)-Low Confining Pressure	0.082	0.251	0.893
Swell (%)	−0.141	0.849	−0.205

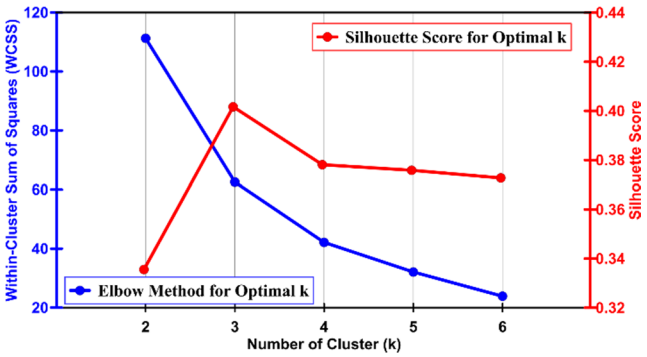


Fig. 21 Elbow method and Silhouette score for optimal cluster number

collectively accounted for 77.6% of the total variance, were retained for clustering.

The loading matrix (Table 4) demonstrates that PC1 is dominated by plasticity, fines content, and compressibility parameters (LL, PI, P200, Cc, m_v), forming a plasticity–compressibility–suction axis. PC2 is governed primarily by strength-related variables (UCS, cohesion, swell, P_{break}), representing a strength–transition axis, while PC3 is heavily influenced by friction angle, indicating a friction-controlled mechanical axis. This structure provides a transparent basis for dimensionality reduction and aligns with reviewer expectations for reproducibility.

The optimal number of clusters for the K-means analysis was determined using the Elbow Method and Silhouette Score (Fig. 21). Both criteria converged at k = 3, indicating a three-cluster solution that balances within-group cohesion and between-group separation. Cluster assignments were then projected into PC1–PC2 space (Fig. 22), where distinct grouping is visually evident, with centroids showing clear spatial separation. Convex hulls were included solely for visualization; they do not represent statistical boundaries.

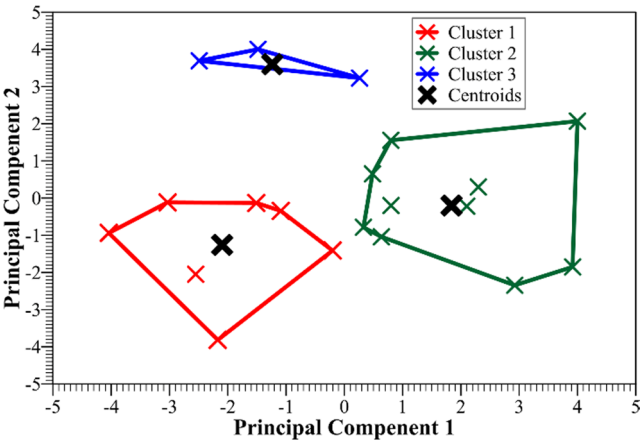


Fig. 22 PCA clustering with centroids and convex hulls

Table 5 Mean values of geotechnical and hydraulic parameters by cluster

Parameter	Cluster 1	Cluster 2	Cluster 3
PI (%)	31.52	26.14	32.67
P200 (%)	90.70	78.56	92.37
q_{dmax} (Mg/m^3)	1.49	1.57	1.53
W_{opt} (%)	23.44	21.76	22.61
Swell (kPa)	59.33	79.34	82.60
Swell (%)	1.82	2.41	3.13
Cc	0.30	0.22	0.27
Cs	0.035	0.024	0.028
m_v (m²/MN)	0.12	0.08	0.10
UCS (kPa)	173.40	243.47	261.32
Cohesion (kPa)	144.33	189.84	227.41
Friction angle (°)	18.76	21.92	26.33
AEV (kPa)	506.37	312.20	530.14
P_{break} (kPa)	351.12	401.09	480.44

Cluster characterization

The mean values of geotechnical and hydraulic properties for each cluster are provided in Table 5. These values reveal three distinct behavioral soil groups:

Cluster 1 represents moderately plastic soils (PI ≈ 31.5%) with very high fines content (P200 ≈ 90.7%) and the lowest dry density (1.49 Mg/m³). These soils exhibit the highest compressibility (Cc ≈ 0.30; m_v ≈ 0.12 m²/MN) and the lowest shear strength (UCS ≈ 173 kPa, φ ≈ 18.8°). Their AEVs (≈ 506 kPa) and mechanical breakpoints (≈ 351 kPa) fall in the mid-range of the dataset. This indicates comparatively deformable soils whose hydromechanical thresholds initiate at moderate suction and stress levels.

Cluster 2 comprises the densest and least plastic soils (MDD ≈ 1.57 Mg/m³; PI ≈ 26.1%) with the lowest fines content (78.6%). These soils exhibit the lowest compressibility and only moderate swelling. Their mechanical and hydraulic thresholds (P_{break} ≈ 401 kPa; AEV ≈ 312 kPa) are the lowest among all groups, suggesting earlier transition

to desaturation and frictional behavior. Their strength levels (UCS ≈ 243 kPa; cohesion ≈ 190 kPa) reflect a structure governed by compaction-induced density rather than suction.

Cluster 3 includes the most active, suction-controlled soils, characterized by high plasticity (PI ≈ 32.7%), high fines content (92.4%), and the greatest swelling potential (≈ 3.1%). This group displays the highest strength (UCS ≈ 261 kPa; cohesion ≈ 227 kPa) and the largest hydro-mechanical thresholds (AEV ≈ 530 kPa; P_{break} ≈ 480 kPa). These characteristics indicate microstructurally active soils in which high suction retention delays the onset of mechanical yielding.

Interpretation of multivariate structure

The joint analysis of PCA loadings and cluster means highlights a coherent hydro-mechanical framework:

- Variables associated with plasticity and compressibility load strongly on PC1 and dominate Clusters 1 and 3.
- Strength-related variables (UCS, cohesion, swell) load heavily on PC2 and primarily characterize Clusters 2 and 3.
- Friction angle, the chief component of PC3, differentiates Cluster 3 as a mechanically stronger and more structured group.

This structure confirms that the clusters arise from a combination of plasticity-driven suction behavior and stress-dependent strength mechanisms.

Regression evaluation and hydro-mechanical coupling

Regression analysis was initially explored to predict AEV from geotechnical variables; however, because Perera-AEV is derived directly from PI and P200 in the model formulation, such regression produced circular dependencies and was not statistically meaningful. A weak association was observed between AEV and mechanical breakpoint stress (P_{break}), but formal regression was also unsuitable due to overlapping parameter derivations.

Nevertheless, a conceptual relationship emerges: samples with higher AEV values also show higher mechanical breakpoints, particularly in Cluster 3. This suggests that the suction threshold controlling desaturation may align with the stress level at which triaxial behavior transitions between mechanical regimes—a pattern consistent with the hydro-mechanical coupling interpretations presented in Sect. 4.5.

The PCA–K-means framework provided a statistically sound and physically interpretable classification of the soils into three behavioral groups. The clusters reflect distinct combinations of plasticity, suction retention, compressibility, and strength, offering a holistic view of soil behavior under compacted conditions. The integration of AEV, P_{break}, and strength parameters supports the broader interpretation that suction-related hydraulic thresholds and mechanical yielding transitions share a common microstructural foundation.

Conclusions

This study investigated the hydraulic, mechanical, and multivariate behavioral characteristics of 20 compacted high-plasticity soils. The main objective was to determine whether the hydraulic transition threshold, represented by the air entry value (AEV), corresponds to the mechanical transition threshold, represented by the breakpoint stress (P_{break}) obtained from connected bilinear p–q shear strength envelopes. A comprehensive experimental program incorporating swelling, consolidation, unconfined compression, and UU triaxial tests, together with SWCC modeling and multivariate statistical classification, provided an integrated dataset from which hydro-mechanical linkages could be evaluated.

The laboratory test results demonstrated substantial variability in swelling behavior, compressibility, and shear strength among the compacted soils, largely governed by fines content, plasticity index, and the compaction-derived microfabric. Swell percentages ranged from 1.1% to 4.9%, and corresponding swell pressures exceeded 45 kPa in the most active clays, reflecting the influence of microstructure and matric suction. Consolidation parameters (Cc, Cs, m_v) showed consistent alignment with swelling trends, indicating that soils exhibiting greater wetting-induced expansion also possessed more deformable and suction-sensitive fabrics. Unconfined compression strengths ranged from 107 to 250 kPa, with suction-controlled stiffening contributing to strength differences among samples. Together, these results confirmed that the hydro-mechanical behavior of compacted fine-grained soils is strongly conditioned by plasticity and pore-size distribution.

SWCCs modeled using the Perera et al. (2005) and Zapata et al. (2000) approaches revealed distinctive differences in suction retention behavior. The Perera model produced AEVs ranging from 263 to 633 kPa, consistent with the high plasticity of the soils, while the Zapata model yielded significantly lower values (52–206 kPa). Comparison of AEV with mechanical breakpoints from the UU triaxial envelopes showed that many samples displayed close correspondence between the two thresholds. In particular, high-plasticity soils tended to exhibit AEV and P_{break} values of similar magnitude, suggesting that the onset of desaturation and the transition in mechanical behavior may arise from closely related microstructural mechanisms. This interpretation is supported by recent findings in the literature indicating

that suction governs both hydraulic transitions and apparent cohesion in unsaturated fine-grained soils.

The connected bilinear shear strength envelopes successfully captured the dual-regime behavior observed under UU loading, with most samples exhibiting a clear breakpoint at which the controlling failure mechanism changed. This bilinear approach provides a practical, nonparametric approximation of the continuous nonlinear strength–suction relationship characteristic of unsaturated soils. The comparison of AEV and P_{break} further confirmed that mechanical yielding in compacted clays is strongly influenced by suction dissipation and pore-pressure evolution, although the two thresholds do not coincide perfectly in all soils.

Multivariate evaluation using PCA and K-means clustering classified the soils into three distinct behavioral groups. These clusters reflected coherent combinations of plasticity, compressibility, suction retention, and strength characteristics, reinforcing the integrated hydro-mechanical interpretation. Soils with high fines content, high plasticity, and large AEV and P_{break} values formed a cluster governed predominantly by suction-controlled behavior. In contrast, denser and less plastic soils formed a lower-suction, strength-moderate cluster primarily controlled by compaction-derived structure. The clustering results highlight the value of multivariate statistical tools in classifying compacted soils according to their coupled hydraulic and mechanical responses.

The findings of this study demonstrate that the AEV derived from SWCC modeling and the mechanical transition stress P_{break} from shear strength envelopes are interrelated indicators of suction-controlled behavior in compacted plastic soils.

Future work should incorporate suction-controlled triaxial testing, microscopic imaging of fabric evolution, and alternative SWCC calibration techniques to validate and refine the hydro-mechanical coupling relationships identified here. Application of machine learning models to integrate SWCC, strength, and microstructural parameters also represents a promising direction for extending predictive capabilities in unsaturated soil mechanics.

Funding Open access funding provided by the Scientific and Technological Research Council of Türkiye (TÜBİTAK).

Data availability Data are available on reasonable request.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Abd IA, Fattah MY, Mekkiyah H (2020) Relationship between the matric suction and the shear strength in unsaturated soil. Case Stud Constr Mater 13:e00441. https://doi.org/10.1016/j.cscm.2020.e00441

ASTM D2166/D2166M-16 (2016) Standard Test Method for Unconfined Compressive Strength of Cohesive Soil. ASTM International, West Conshohocken, PA

ASTM D2435/D2435M-11 (2020) Standard Test Methods for One-Dimensional Consolidation Properties of Soils Using Incremental Loading. ASTM International, West Conshohocken, PA

ASTM D2850-23 (2024) Standard Test Method for Unconsolidated-Undrained Triaxial Compression Test on Cohesive Soils. ASTM International, West Conshohocken, PA

ASTM D698-12e2 (2012) Standard Test Methods for Laboratory Compaction Characteristics of Soil Using Standard Effort (12 400 ft-lbf/ft3 (600 kN-m/m3)). ASTM International, West Conshohocken, PA

Bekdaş G, Aydın Y, Nigdeli SM, Ünver İS, Kim WW, Geem ZW (2025) Modeling soil behavior with machine learning: static and cyclic properties of high plasticity clays treated with lime and fly ash. Buildings 15(2):288. https://doi.org/10.3390/buildings15020288

Bello N, Satyanaga A, Irawan S et al (2025) An innovative modelling technique for bimodal soil water characteristic curve under wetting process. Sci Rep 15:11837. https://doi.org/10.1038/s41598-025-93987-2

Bozorgzadeh N, Feng Y (2024) Evaluation structures for machine learning models in geotechnical engineering. Georisk Assess Manage Risk Eng Syst Geohazards 18(1):52–59. https://doi.org/10.1080/17499518.2024.2313485

Chae J, Kim B, Park SW, Kato S (2010) Effect of suction on unconfined compressive strength in partly saturated soils. KSCE J Civ Eng 14(3):281–290. https://doi.org/10.1007/s12205-010-0281-7

Chiu CF, Ng CW (2012) Coupled water retention and shrinkage properties of a compacted silt under isotropic and deviatoric stress paths. Can Geotech J 49(8):928–938. https://doi.org/10.1139/t2012-055

Fondjo AA, Theron E, Ray RP (2021) Investigation of the ınfluencing soil parameters on the air entry values in soil-water characteristic curve of compacted heaving soils. Civ Eng Archit 9(1):91–114. https://doi.org/10.13189/cea.2021.090108

Fredlund DG, Rahardjo H (1993) Soil Mechanics for Unsaturated Soils. John Wiley & Sons

Fredlund DG, Xing A (1994) Equations for the soil-water characteristic curve. Can Geotech J 31(4):521–532. https://doi.org/10.1139/t94-061

Gao W (2024) The application of machine learning in geotechnical engineering. Appl Sci 14(11):4712. https://doi.org/10.3390/app14114712

Goktepe AB, Altun S, Sezer A (2005) Soil clustering by fuzzy c-means algorithm. Adv Eng Softw 36(10):691–698. https://doi.org/10.1016/j.advengsoft.2005.01.008

Harle SM, Wankhade RL (2025) Machine learning techniques for predictive modelling in geotechnical engineering: a succinct review. Discover Civil Eng 2:86. https://doi.org/10.1007/s44290-025-00224-w

Houlsby GT (1997) The work input to an unsaturated granular material. Géotechnique 47(1):193–196. https://doi.org/10.1680/geot.1997.47.1.193

Hsiao CH, Kumar K, Rathje EM (2024) Explainable AI models for predicting liquefaction-induced lateral spreading. Front Built Environ. https://doi.org/10.3389/fbuil.2024.1387953

Johari A, Habibagahi G, Ghahramani A (2006) Prediction of soil–water characteristic curve using genetic programming. J Geotechn Geoenviron Eng 132(5):661–665. https://doi.org/10.1061/(ASCE)1090-0241(2006)132:5(661)

Khalili N, Khabbaz MH (1998) A unique relationship for χ for the determination of the shear strength of unsaturated soils. Geotechnique 48(5):681–687

Lade PV, Duncan JM (1973) Cubical triaxial tests on cohesionless soil. J Soil Mech Found Div 99(10):793–812

Lin Z, Qian J, Shi Z, Zhai Q (2022) Bi-linear strength envelope of coarse- and fine-grained unsaturated soils with bimodal water-retention curve. Eng Geol 304:106694. https://doi.org/10.1016/j.enggeo.2022.106694

Lu N, Likos WJ (2004) Unsaturated soil mechanics. John Wiley & Sons Inc, New York

Niu G, Sun D, Kong L et al (2024) Investigation into the shear strength of a weakly expansive soil over a wide suction range. Acta Geotech 19:3059–3073. https://doi.org/10.1007/s11440-024-02246-7

Ozcoban MS, Isenkul ME, Durak SG, Acar TO, Övez S, Tüfekci N (2018) Predicting permeability of compacted clay filtrated with landfill leachate by k-Nearest Neighbors modelling method. Water Sci Technol 77(8):2155–2164. https://doi.org/10.2166/wst.2018.139

Pande P, Giri J, Ali MS, Mohammad F, Raut J, Raut S, Sathish T, Giri P (2024) Comparative analysis of saturated–unsaturated shear strength under undrained loading: Experimental validation and ANN prediction of clayey soils. AIP Adv 14:075118. https://doi.org/10.1063/5.0206783

Perera YY, Zapata CE, Houston WN, Houston SL (2005) Prediction of the soil-water characteristic curve based on grain-size-distribution and index properties. Adv Pavement Eng. https://doi.org/10.1061/40776(155)4

Quevedo R, Lopez M, Roehl D (2024) Hydromechanical modeling of unsaturated soils considering compaction effects on soil physical and hydraulic properties. Acta Geotech 19:5051–5064. https://doi.org/10.1007/s11440-024-02305-z

Shafqat K, Khalid U, Rehman Z (2025) Coupling effect of cyclic wet dry environment and compaction state on desiccation cracking and mechanical behavior of low and high plastic clays. Bull Eng Geol Environ 84:66. https://doi.org/10.1007/s10064-024-04049-2

Sharma S, Rathor APS, Sharma JK (2025) Prediction of soil water characteristic curve of unsaturated soil using machine learning. Multiscale Multidiscip Model Exp Des 8:72. https://doi.org/10.1007/s41939-024-00664-4

Tripathy S, Tadza MYM, Thomas HR (2014) Soil-water characteristic curves of clays. Can Geotech J 51(8):869–883. https://doi.org/10.1139/cgj-2013-0089

Türköz M (2019) Computer-controlled equipment for the direct measurement of the swell potential of expansive soils. Arab J Geosci 12:743. https://doi.org/10.1007/s12517-019-4987-7

Uzundurukan S (2023) Prediction of soil–water characteristic curve for plastic soils using PSO algorithm. Environ Earth Sci 82:37. https://doi.org/10.1007/s12665-022-10687-0

Xiao Y, Liu S, Wu H, Shi J, Liu H (2023) Strengths and ifinite slope stability of unsaturated soils. Int J Geomech 23(12):1–9. https://doi.org/10.1061/IJGNAI.GMENG-9021

Zapata CE, Houston WN, Houston SL, Walsh KD (2000) Soil–water characteristic curve variability. Adv Unsaturat Geotech. https://doi.org/10.1061/40510(287)7

Zhai Q, Rahardjo H, Satyanaga A, Dai G (2019) Estimation of unsaturated shear strength from soil–water characteristic curve. Acta Geotech 14:1977–1990. https://doi.org/10.1007/s11440-019-00785-y

Zhai Q, Zhu Y, Rahardjo H, Satyanaga A, Dai G, Gong W, Zhao X, Ou Y (2023) Prediction of the soil–water characteristic curves for the fine-grained soils with different initial void ratios. Acta Geotech 18:5359–5368. https://doi.org/10.1007/s11440-023-01833-4

Zhang F, Cui Y, Zeng L, Robinet JC, Conil N, Talandier J (2018) Effect of degree of saturation on the unconfined compressive strength of natural stiff clays with consideration of air entry value. Eng Geol 237:140–148. https://doi.org/10.1016/j.enggeo.2018.02

Zhang C, Li J, Zhang Y, Hu T (2022) Experimental and discrete element modeling study on suction stress characteristic curve and soil–water characteristic curve of unsaturated reticulated red clay. Bull Eng Geol Environ 81:363. https://doi.org/10.1007/s10064-022-02834-5

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.