Machine learning-driven modeling of soil plasticity and strength parameters with interpretability insights

Giovanni Spagnoli[1*], *Mohammadreza* Mahmoudi[2], *Satoru* Shimobe[3] and *Stefano* Collico[4]

[1] CDM Smith SE, Bochum, Germany
[2] Department of Civil Engineering, The Sharif University of Technology, Tehran, Iran
[3] College of Science and Technology, Nihon University, Funabashi, Japan
[4] Dipartimento di Ingegneria Civile e Ambientale, Universitá degli Studi di Firenze, Firenze, Italy

Abstract. This study proposes advanced stacking ensemble machine learning approaches to predict soil Liquidity Index (*LI*) and Undrained Shear Strength (S_u). Over 1,550 *LI* and 950 S_u measurements from global research were employed. Individual regressors—Multilayer Perceptron (MLP), Random Forest (RF), and AdaBoost—were evaluated alongside ensemble strategies, including Simple Averaging, Weighted Averaging, and stacking using a Support Vector Regression (SVR) meta-learner. Hyperparameter tuning was performed using both Bayesian Optimization (BO) and Particle Swarm Optimization (PSO). AdaBoost achieved the best results for *LI* prediction, while RF yielded the highest accuracy for S_u. Weighted Averaging ensemble methods produced outstanding predictive performances, achieving R^2 values of 0.9913 (BO) and 0.9961 (PSO) for S_u, and 0.9624 (BO) and 0.9338 (PSO) for LI. BO proved superior for *LI* models, while PSO excelled for S_u models. The results highlight the robust potential of ensemble modeling and tailored optimization in the field of geotechnical engineering.

1 Introduction

Accurate characterization of soil properties is essential for geotechnical stability. The Liquidity Index (*LI*) and Undrained Shear Strength (S_u) provide key indicators of soil behavior under varying conditions [1].

The Liquidity Index (*LI*) quantifies the natural water content relative to the plasticity range of the soil and is defined as:

$$LI = \frac{w-PL}{LL-PL} \quad (1)$$

where w is natural moisture content, PL the plastic limit, and LL the liquid limit [2]

LI expresses the soil's consistency state, with values above 1 indicating soil is wetter than its liquid limit, prone to flow, and values near zero indicating a soil near its plastic limit [2,3]. This index helps assess the susceptibility of soil to deformation and strength reduction as moisture changes.

Undrained Shear Strength (S_u) is the shear resistance of saturated fine-grained soils when drainage is prevented during loading [4]. S_u is critical for slope stability and foundation design under rapid loading. The shear strength in saturated clays under undrained conditions is nearly independent of confining effective stress because pore water pressure supports the load [4, 5]. Hence, S_u is often termed "apparent cohesion" resulting from pore pressure effects, though the true strength mechanism is frictional [6, 7].

Empirical relationships between *LI* and S_u were established by [7], showing that the remoulded shear strength decreases exponentially with increasing *LI*.

When S_u is plotted against *LI* on a logarithmic scale, the data exhibit an approximately linear trend, confirming the robustness of this empirical correlation [7]. Typically, the undrained shear strength at the liquid limit ranges around 2 kPa, while at the plastic limit it lies between 100 and 200 kPa [7, 8]. These characteristic values provide a useful benchmark for interpreting soil consistency and deformation behavior.

Traditional liquid limit test methods, such as the Casagrande cup and thread rolling methods [2], though widely adopted, are time-consuming and prone to operator-dependent variability [1]. To address these limitations, the fall cone test, standardized by [9, 10], emerged as a more objective alternative. By measuring the penetration depth of a standardized cone under controlled conditions, the fall cone method provides a consistent and reproducible determination of the liquid limit, substantially improving test precision and comparability among laboratories [1].

In the present database, both *LI* and S_u are not independently measured target variables but are directly derived from the measured cone penetration depth d through empirical fall-cone based correlations calibrated for each soil type and cone configuration. Consequently, using d simultaneously as an input feature and as the underlying source variable for the target definitions inherently induces a very strong functional dependency between predictor and response.

* Corresponding author: giovanni.spagnoli@cdmsmith.com

This tight linkage explains the unusually high R^2 values obtained in the models from a soil mechanics perspective and should not be interpreted as an indication that cone penetration depth alone can universally predict independent strength or plasticity parameters outside the context of the adopted fall-cone based formulations.

Machine learning (ML) techniques have emerged as innovative tools for handling complex geotechnical datasets [11, 12]. Recent studies show ensemble learning algorithms outperform individual models in predicting soil mechanical properties [13, 14].

The data used in this study are obtained from the work of [15], who integrated Bayesian optimization (BO) and particle swarm optimization (PSO) techniques within an ensemble machine learning framework combining random forest (RF), multilayer perceptron (MLP), and AdaBoost regressors as base learners, with support vector regression (SVR) serving as the meta-learner, to enhance prediction of soil liquidity index and shear strength.

2 Data collection and pre-processing

The dataset incorporates more than 1,550 liquidity index (*LI*) measurements and 950 undrained shear strength (S_u) entries, aggregated from over 40 peer-reviewed international research studies [15]. This scope provides robust global coverage and diversity in geotechnical soil behavior. The principal features extracted for analysis include cone penetration depth (d), sample condition (disturbed or undisturbed), and cone type specifications (30°/80g, 30°/100g, 60°/60g), ensuring that soil state and laboratory setup are thoroughly considered in model development.

Descriptive statistics for all variables are presented in Table 1, summarizing distributions and central tendencies within the compiled dataset. Handling incomplete or missing data points is crucial; for this study, the K-Nearest Neighbors (KNN) imputation algorithm was utilized to resolve missing values—this method estimates unknown entries by leveraging similarities to neighboring data points, thus preserving the multivariate structure of the input set [16, 17].

Table 1. Statistical characteristics of collected data (s=soil type, c=cone type).

Test	Parameter	Max	Min	Mean	SD
S_u	*d* (mm)	100	0.8	9.83	7.23
	s	1	0	0.23	0.42
	c	2	0	1.07	0.98
	S_u (kPa)	183	0.1	16.09	26.77
LI	*d* (mm)	59.4	0.38	12.78	8.30
	s	1	0	0.09	0.28
	c	2	0	0.38	0.72
	LI	2.44	0	0.79	0.35

Additionally, categorical variables (e.g., soil type, cone type) were systematically converted to numerical encodings to facilitate seamless integration into machine learning algorithms. Subsequent to this transformation, all input variables were linearly normalized onto a [0,1] range, ensuring comparability and enhancing convergence rates during model training. The normalization is mathematically defined as:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2)$$

where x_{norm} is the normalized value, $x$is the feature to be scaled, x_{min} and x_{max} are the minimum and maximum values for that variable within the dataset.

3 Methodology

3.1 Stacking ensemble

The stacked ensemble approach enhances regression accuracy by synthesizing predictions from diverse base learners to mitigate individual model bias and variance [18].

As illustrated in Fig.1, the preprocessing workflow encompasses data cleaning, normalization, feature selection, and partitioning to prepare the dataset for predictive modeling.

The system employs three primary-base regressors—Multilayer Perceptron (MLP), Random Forest (RF), and AdaBoost regressor—whose individual outputs are combined by a meta-learner based on Support Vector Regression (SVR). This meta-learner, trained on the predictions of the base learners, aims to deliver optimally blended outputs that outperform any single constituent model.

To guard against overfitting, a K-fold cross-validation strategy with K=5 is adopted. In this setup, the full training set is divided into five subsets, iteratively using each as a validation set while training on the remaining portions to ensure robust and reliable performance estimation. Two ensemble prediction fusing strategies are employed. In the Simple Averaging (SA), the ensemble prediction P_{SA} is the arithmetic mean of all base model predictions:

$$P_{SA} = \frac{1}{N}\sum_{i=1}^{N} P_i \quad (3)$$

where P_i denotes the prediction from the i-th base learner and N is the total number of learners.

Weighted averaging (WA) enhances ensemble accuracy by assigning a specific weight to each base learner's prediction:

$$P_t = \sum_{i=1}^{N} \omega_i P_i(t) \quad (4)$$

where each weight ω_i reflects the base learner's contribution and predictive reliability.

In this formulation, each $P_i(t)$ is the prediction made by the i-th base learner at time t, and the weights ω_i reflect the base learner's individual contribution and

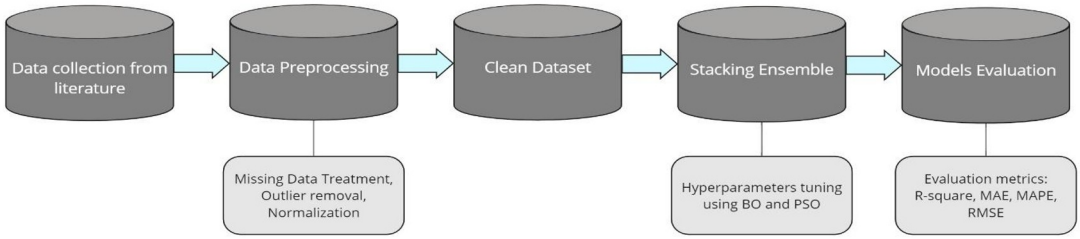

Fig. 1. Preprocess flow diagram.

predictive reliability. The weights are often calculated based on past performance, accuracy, or a reliability measure of each learner. Generally, these weights can be obtained through various schemes such as validation performance or optimization of prediction error.

The weights are calculated as:

$$\omega_i = \frac{DC_i}{\sum_{i=1}^{N} DC_i} \tag{5}$$

where DC_i is a data-driven performance coefficient (such as inverse error or accuracy score) for the i-th learner, ensuring that more effective models exert greater influence on the final output [11, 19]. This method dynamically emphasizes high-performing predictors within the ensemble.

Fig. 2 presents the complete computational pipeline—from raw data ingestion and preprocessing to sequential model training, prediction stacking, and final output generation.

3.2 Machine learning models

MLP is a type of feed-forward artificial neural network composed of an input layer, one or more hidden layers, and an output layer. Each neuron applies a nonlinear activation function, enabling the network to capture complex, non-linear relationships between features. MLPs are foundational to deep learning and provide the structural basis for more sophisticated neural architectures. Their capacity for universal function approximation makes them particularly effective in multivariate regression and pattern recognition tasks [20].

The Random Forest algorithm aggregates the predictions from an ensemble of individual decision trees, each constructed from random bootstrap samples of the training data. Furthermore, at each node, only a random subset of predictor variables is considered for splitting. This built-in randomness decorrelates individual trees, substantially reducing variance and the risk of overfitting commonly encountered with single tree models. The RF's robustness, scalability, and ability to estimate variable importance make it a preferred choice for regression and classification problems in complex, high-dimensional datasets [21, 22].

AdaBoost, or Adaptive Boosting, combines the outputs of a sequence of weak learners—typically shallow trees—by iteratively adjusting their influence based on previous performance. In each iteration, higher weights are assigned to observations that have been mis-predicted, compelling the ensemble to focus on difficult cases. The final prediction is derived from a weighted sum of all base learners, yielding significantly improved overall accuracy without substantially increasing bias. The AdaBoost method is especially valued for its resilience to noise and its strong theoretical guarantees for convergence under weak learning assumptions [23, 24].

SVR extends the principles of Support Vector Machines to regression, aiming to find a function that deviates from the true targets by a value no greater than a defined precision margin (epsilon), while maintaining model flatness. Utilizing a radial basis function (RBF) kernel enables SVR to efficiently handle complex, non-linear relationships by implicitly mapping input data into a higher-dimensional space. This makes SVR highly adept at minimizing the overall deviation of sample points from the regression hyperplane, particularly in cases with non-linear target dependencies [25].

3.3 Hyperparameter optimization

Optimizing hyperparameters is a critical step for ensuring robust and generalizable machine learning models. Toward this goal, two state-of-the-art optimization algorithms are utilized: Bayesian Optimization (BO) and Particle Swarm Optimization (PSO) [26].

Bayesian Optimization leverages Bayes' theorem to estimate the posterior distribution of the unknown objective function based on historical evaluations. This approach employs Gaussian Processes (GP) to probabilistically model the hyperparameter search space, allowing BO to balance exploration and exploitation effectively when selecting new candidate solutions [15]. The conditional mean $\mu[x^*]$ and variance $\sigma^2[x^*]$ at a test candidate x^*, which guide the acquisition function in BO, are given by:

$$\mu[x^*] = K[x^*, X] \cdot K[X, X]^{-1} \cdot f \tag{6}$$

$$\sigma^2[x^*] = K[x^*, x^*] - K[x^*, X] \cdot K[X, X]^{-1} \cdot K[X, x^*] \tag{7}$$

Here, $K[\cdot,\cdot]$ denotes the kernel function encoding similarity between points, X is the set of sampled hyperparameters, and f are the corresponding validation scores. This modeling guides the search towards promising regions, efficiently discovering optimal or

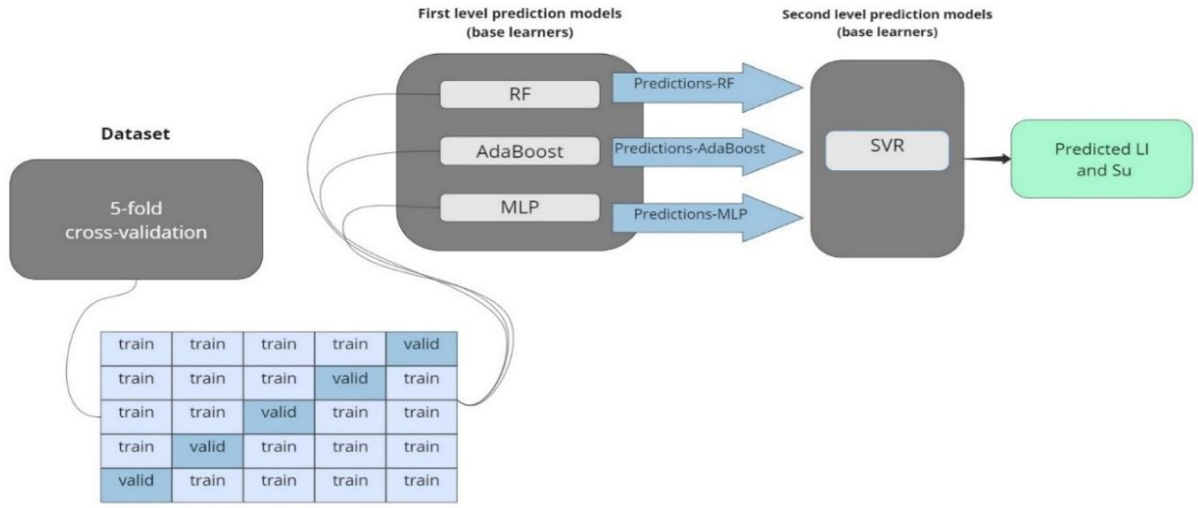

Fig. 2. Pipeline of the proposed method.

near-optimal hyperparameters even in high-dimensional, non-convex spaces.

PSO is a population-based metaheuristic inspired by the collective behavior of social organisms. Each particle represents a candidate solution, which updates its search position in the hyperparameter space based on both its prior experience and the information shared by its neighbors [15]. The update rules for position and velocity at iteration t are:

$$x^p(t+1) = x^p(t) + v^p(t+1) \tag{8}$$

$$v_j^p(t+1) = k_1 v_j^p(t) + k_2 \omega_{1j}(t)(y_j^p(t - x_j^p(t)) + k_3 \omega_{2j}(t)(\tilde{y}_j(t) - x_j^p(t)) \tag{9}$$

In this algorithm, the velocity update incorporates both cognitive components (the tendency to return to a particle's best-known position) and social components (the influence of the swarm's global best position). The random coefficients $\omega_{1j}(t)$ and $\omega_{2j}(t)$ instill stochasticity, helping the search avoid local optima [27].

Table 2 provides a comprehensive overview of the specific hyperparameters optimized for each model in this study, reflecting the tailored search for optimal performance across different learning architectures

3.4 Performance metrics

Model performance was rigorously evaluated using multiple statistical metrics to ensure a comprehensive assessment of predictive accuracy and reliability. The combination of these metrics allows for nuanced measurement of model error, robustness, and practical utility in regression scenarios.

R^2 (Coefficient of Determination) measures the proportion of variance in the observed values explained by the model predictions:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2} \tag{10}$$

Where y_i and $\hat{y}_i$ are the actual and predicted values, respectively, and $\bar{y}$ is the mean of actual values. An R^2 value close to 1 indicates an excellent model fit.

RMSE (Root Mean Square Error) quantifies average magnitude of prediction errors (in the same units as the target variable), penalizing larger deviations more heavily:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}} \tag{11}$$

Lower RMSE values reflect higher prediction accuracy.

Table 2. Hyperparameter ranges for machine learning models.

Hyperparameter	Model	Range
hidden_layer_size	MLP	0-10
alpha (L2)	MLP	0.001-0.1
momentum	MLP	0.2-0.9
n_estimators	RF	5-100
random_state	RF	1-100
n_estimators	AdaBoost	5-100
random_state	AdaBoost	1-100
C	SVR	1-300
gamma	SVR	0.01-30
epsilon	SVR	0.01-0.9

MAE (Mean Absolute Error) is the average of absolute differences between actual and predicted values, offering a straightforward interpretation of typical model error:

$$MAE = \frac{\sum_{i=1}^{N}|y_i - \hat{y}_i|}{N} \tag{12}$$

As it does not square the errors, MAE is less sensitive to outliers compared to RMSE.

MAPE (Mean Absolute Percentage Error) expresses prediction accuracy as a percentage, facilitating intuitive comparison across different datasets:

$$MAPE = \frac{\sum_{i=1}^{N} \left| \frac{y_i - \hat{y}_i}{y_i} \right|}{N} \times 100\% \quad (13)$$

Here y_i and $\hat{y}_i$ represent actual and predicted values for each data point, respectively.

4 Results and discussion

4.1 Individual model performance

Assessment of regression models for the prediction of *LI* and S_u reveals distinct strengths across algorithms. For LI prediction, the AdaBoost regressor consistently delivered the highest R^2, with values of 0.8871 during Bayesian Optimization (BO) training and 0.8787 for BO testing. This notable performance highlights AdaBoost's effectiveness in leveraging weak learners to capture complex, nonlinearities in *LI* data—surpassing both Multilayer Perceptron (MLP) and Random Forest (RF) models in this context.

In contrast, for S_u prediction, Random Forest demonstrated clear superiority among individual learners, returning R^2 values exceeding 0.97 in both training and testing phases, irrespective of whether BO or PSO was employed for hyperparameter tuning. The ensemble of de-correlated trees in RF enables robust modeling of the heterogeneous and high-dimensional geotechnical S_u dataset, minimizing variance and adapting well to different optimization strategies.

4.2 Ensemble model performance

Ensemble methods substantially enhanced predictive accuracy across all tested tasks. For *LI* prediction, the weighted averaging ensemble (WA) with Bayesian Optimization (WA-BO) achieved an outstanding R^2 score of 0.9624 on the test set—a pronounced improvement over the best single-model result. The PSO-optimized WA ensemble (WA-PSO) also yielded a strong performance, reaching a test R^2 of 0.9338. These results underscore the benefit of aggregating diverse base models, as ensemble strategies leverage complementary strengths while reducing model-specific weaknesses.

For S_u, ensemble averaging produced remarkable accuracy: the WA-BO configuration attained test $R^2 = 0.9913$, while WA-PSO pushed this even further to $R^2 = 0.9961$ on testing data. Such high predictive power demonstrates the ensemble method's reliability for critical engineering estimations of undrained shear strength.

Notably, stacking ensembles with a Support Vector Regression (SVR) meta-learner yielded additional improvements. The RF-AdaBoost-SVR combination achieved exceptional Su predictions, with BO and PSO test R^2 values of 0.9943 and 0.9946, respectively demonstrating effective synergy between decision tree, boosting, and kernel-based learning. For *LI*, the MLP-AdaBoost-SVR stack delivered the highest BO testing performance (R^2 = 0.9251), indicating the value of blending neural and tree-based regressors before SVR meta-learning.

4.3 Optimization comparison

A comparative evaluation of Bayesian Optimization (BO) and Particle Swarm Optimization (PSO) reveals that each algorithm exhibits distinctive optimization strengths depending on the specific prediction target and model architecture. For *LI* prediction tasks, BO routinely achieves superior R^2 values and reduced error metrics across a broad spectrum of models. This advantage stems from BO's probabilistic, model-based search process, which excels at navigating complex, high-dimensional hyperparameter spaces where subtle interactions between features significantly influence model outcomes [28].

In contrast, PSO demonstrates notable superiority in S_u prediction, particularly when utilized within ensemble modeling frameworks. The adaptive, population-based search dynamics of PSO are especially effective at capturing intricate relationships in S_u datasets. PSO's global search capabilities and resilience to local optima make it adept at uncovering robust hyperparameter configurations in scenarios characterized by strongly aggregated or ensemble architectures [29].

These results highlight the importance of aligning the optimization strategy with the underlying prediction task:

- BO is preferable for *LI* modeling due to its consistency, efficient convergence, and ability to quantify uncertainty during the search, supporting the identification of near-optimal solutions in non-linear and high-dimensional landscapes [28].
- PSO is better suited for S_u prediction, leveraging swarm intelligence to efficiently explore diverse solution spaces and capitalize on collaborative effects in ensemble models, which is reflected in both higher accuracy rates and improved model generalization [30].

The feature importance analysis highlights cone penetration depth *d* as the dominant predictor for both *LI* and S_u models, which is consistent with the fact that both target variables are derived from *d* in the underlying fall-cone formulations. For *LI*, *d* shows the highest contribution, followed by soil type *s*, indicating that grain-size distribution and plasticity modulate the penetration–*LI* relationship, while cone type *c* (angle and mass) plays a smaller but non-negligible role.

In the S_u models, *d* again explains most of the variance, but the relative importance of *c* is higher than for *LI*, reflecting the sensitivity of undrained shear strength estimates to cone configuration and test setup. Soil type *s* still contributes to the predictions by capturing the influence of mineralogy and structure on the conversion from penetration depth to shear strength,

while overall the models mainly exploit the strong physical link between d and the derived target variables.

Radar diagrams such as those presented in Fig.3 visually substantiate these trends, providing clear evidence that BO outperforms PSO for LI while PSO consistently leads in S_u predictions across tested model configurations. As such, selecting an optimization algorithm tailored to the specific predictive task and data structure is crucial for maximizing machine learning model performance.

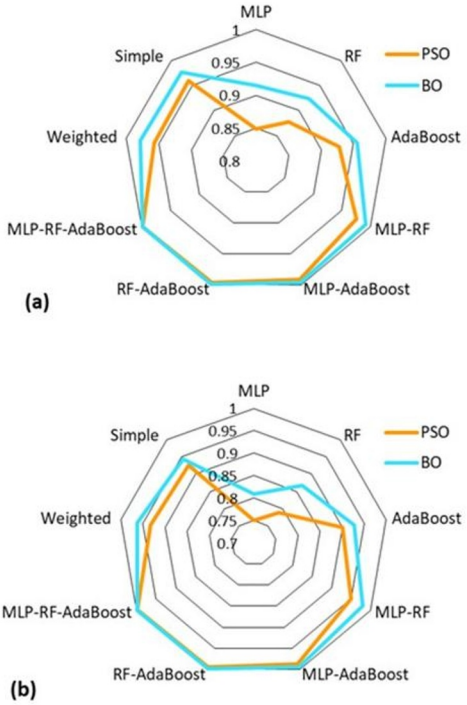

Fig. 3. The radar diagrams based on R^2, (a) training data set, (b) testing data set.

4.4 Feature importance and sensitivity

Feature importance analysis revealed that cone penetration depth (d) is the most influential factor driving both LI and S_u predictions. This finding is consistent with theoretical understanding of the fall cone test, where penetration depth directly reflects a soil's consistency and shear behavior under load [31]. Cone type ranked as the second most significant predictor, followed by soil type, highlighting how variations in test apparatus and material conditions also contribute notably to predictive accuracy.

Further validation of the models demonstrated no evidence of systematic bias: residuals were symmetrically distributed about the one-to-one reference line across all scatter plots, confirming the reliability and generalization capability of the ensemble prediction framework. Notably, MAPE for the best-performing ensemble models remained below 15%, underlining their practical value for engineering applications where prediction errors must be strictly controlled.

4.5 Feature importance and sensitivity

Feature importance analysis revealed that cone penetration depth (d) is the most influential factor driving both LI and S_u predictions. This finding is consistent with theoretical understanding of the fall cone test, where penetration depth directly reflects a soil's consistency and shear behavior under load [31]. Cone type ranked as the second most significant predictor, followed by soil type, highlighting how variations in test apparatus and material conditions also contribute notably to predictive accuracy.

Further validation of the models demonstrated no evidence of systematic bias: residuals were symmetrically distributed about the one-to-one reference line across all scatter plots, confirming the reliability and generalization capability of the ensemble prediction framework. Notably, MAPE for the best-performing ensemble models remained below 15%, underlining their practical value for engineering applications where prediction errors must be strictly controlled.

Fig. 4 provides a visual comparison of predicted and measured LI in the testing dataset using PSO-optimized models. The close clustering of points along the identity line, together with regression equation fitting and narrow scatter, reflects both model fidelity and practical deployment potential in real-world geotechnical engineering design scenarios.

4.6 Feature importance and sensitivity

Feature importance analysis revealed that cone penetration depth (d) is the most influential factor driving both LI and S_u predictions. This finding is consistent with theoretical understanding of the fall cone test, where penetration depth directly reflects a soil's consistency and shear behavior under load [31]. Cone type ranked as the second most significant predictor, followed by soil type, highlighting how variations in test apparatus and material conditions also contribute notably to predictive accuracy.

Further validation of the models demonstrated no evidence of systematic bias: residuals were symmetrically distributed about the one-to-one reference line across all scatter plots, confirming the reliability and generalization capability of the ensemble prediction framework. Notably, MAPE for the best-performing ensemble models remained below 15%, underlining their practical value for engineering applications where prediction errors must be strictly controlled.

Fig. 4 provides a visual comparison of predicted and measured LI in the testing dataset using PSO-optimized models. The close clustering of points along the identity line, together with regression equation fitting and narrow scatter, reflects both model fidelity and practical deployment potential in real-world geotechnical engineering design scenarios.

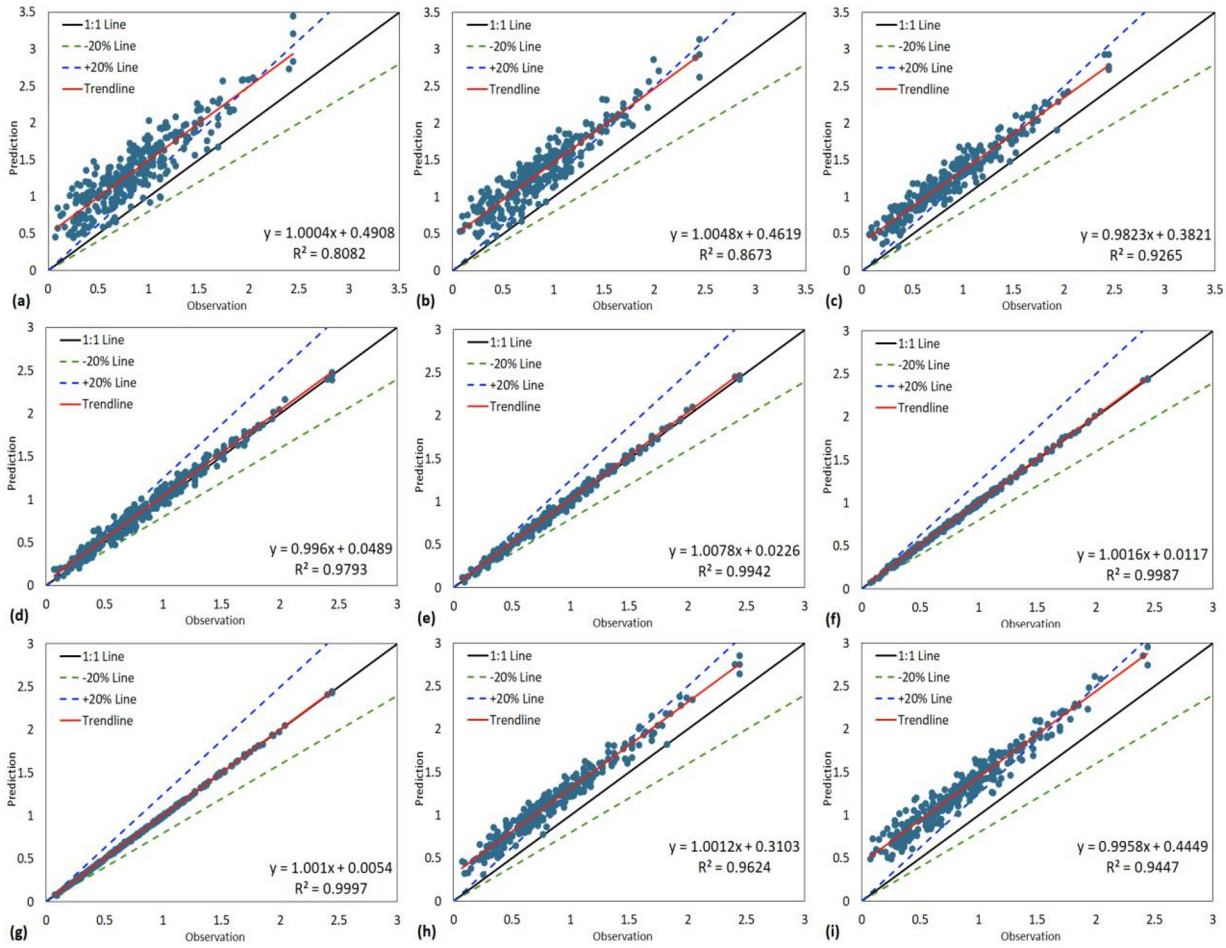

Fig. 4. Scatter plots from the test datasets for LI predictions optimized using PSO, corresponding to the following model setups. Y-axis shows the prediction, X-axis shows the observation: (a) multilayer perceptron (MLP), (b) random forest (RF), (c) AdaBoost, (d) combined MLP and RF, (e) combined MLP and AdaBoost, (f) combined RF and AdaBoost, (g) integrated MLP-RF-AdaBoost ensemble, (h) weighted averaging approach, and (i) simple averaging method.

5.Conclusions

This study demonstrated the effective development and validation of ensemble machine learning models for predicting soil Liquidity Index (*LI*) and Undrained Shear Strength (S_u), offering substantial improvements over traditional approaches. The principal findings are as follows:

- AdaBoost regressor emerged as the top-performing individual model for *LI* prediction, while Random Forest (RF) consistently achieved the highest accuracy for S_u prediction, underscoring the distinctive strengths of boosting and tree-based methods for different geotechnical properties.
- The application of ensemble techniques—especially Weighted Averaging—markedly boosted predictive performance, achieving R^2 values above 0.96 for *LI* and exceeding 0.99 for S_u in validation tests. These results highlight the ensemble's ability to capture complex interactions and reduce generalization error across datasets.
- Advanced stacking architectures incorporating a Support Vector Regression (SVR) meta-learner further improved results compared to traditional simple averaging, leveraging diverse model strengths for superior accuracy.
- Bayesian Optimization (BO) proved most effective for optimizing *LI* prediction models, while Particle Swarm Optimization (PSO) delivered the best performance in S_u prediction tasks, reflecting the value of tailoring optimization strategies to the target parameter and model ensemble structure.
- The ensemble modeling framework established in this research provides a robust, data-driven, and computationally efficient alternative for preliminary geotechnical assessments, complementing or substituting conventional laboratory testing in early-stage engineering evaluations.

Future research should explore integration of additional soil parameters, application to diverse geological conditions, and development of real-time prediction systems for field applications. The methodology can be extended to other geotechnical parameters, contributing to safer and more efficient civil engineering practices.

References

1. M. Carter, S.P. Bentley. Soil properties and their correlations. John Wiley & Sons, Ltd, 2nd Edition (2016)
2. A. Casagrande. The liquid limit of soils. ASTM Proceedings **48**, 1033-1049 (1948)
3. R.D. Holtz, W.D. Kovacs. An introduction to geotechnical engineering. Prentice Hall, New Jersey (1981)
4. K. Terzaghi, R.B. Peck. Soil mechanics in engineering practice. John Wiley, London (1967)
5. A.W. Skempton. The pore pressure coefficients A and B. Geotechnique **4**(4), 143-147 (1954)
6. R. Singh, D.J. Henkel, D.A. Sangay. Shear and K0 swelling of overconsolidated clay. Proc. 8th ICSMFE, vol. 1, 367-376 (1973)
7. A.W. Skempton, R.D. Northey. The sensitivity of clays. Géotechnique **3**, 305-353 (1952)
8. L. Bjerrum, N.E. Simons. Comparison of shear strength characteristics of normally consolidated clays. Proc. ASCE Research Conf. on Shear Strength of Cohesive Soils, Boulder, 711-726 (1960)
9. British Standards Institution. Methods of test for soils for civil engineering purposes. BS 1377 (1990)
10. ASTM International. Standard test methods for liquid limit, plastic limit, and plasticity index of soils. ASTM D4318
11. S.F.F. Mojtahedi, I. Ebtehaj, M. Hasanipanah, H. Bonakdari, H.B. Amnieh. Proposing a novel hybrid intelligent model for the simulation of particle size distribution resulting from blasting. Eng. Comput. **35**, 47-56 (2019)
12. N. Yousefpour, F.F. Mojtahedi. Early detection of internal erosion in earth dams: combining seismic monitoring and convolutional AutoEncoders. Georisk (2023): 1-21
13. W. Chen, P. Xie, J. Gu, H. Wang. A comparative study of ensemble learning methods for predicting the uniaxial tensile strength of expansive soil. Eng. Comput. **39**, 2163-2178 (2022)
14. W. Chen, J. Zhang, H. Wang. Prediction of the shear strength of fly ash-amended soils with lignin reinforcement using ensemble learning algorithms. Constr. Build. Mater. **369**, 130583 (2023)
15. M. Mahmoudi, G. Spagnoli, S. Shimobe. Integrating Bayesian and Particle Swarm Optimization in Ensemble Machine Learning for Enhanced Prediction of Soil Liquidity Index and Shear Strength. Int. J. Geosynth. Ground Eng. **11**, 50 (2025). https://doi.org/10.1007/s40891-025-00656-5
16. G.E.A.P.A. Batista, M.C. Monard. An analysis of four missing data treatment methods for supervised learning. Appl. Artif. Intell. **17**, 519-533 (2003)
17. L. Beretta, A. Santaniello. Nearest neighbor imputation algorithms: a critical evaluation. BMC Med. Inform. Decis. Mak. **16**, 74 (2016)
18. F. Farooq, A. Czarnecki, M.U. Akbar Nizami, K. Javed, et al. A comparative study for the prediction of the compressive strength of self-compacting concrete modified with fly ash. Materials **14**, 4934 (2021)
19. V. Nourani, G. Elkiran, S.I. Abba. Wastewater treatment plant performance analysis using artificial intelligence--an ensemble approach. Water Sci. Technol. **78**, 2064-2076 (2018)
20. M.C. Popescu, V.E. Balas, L. Perescu-Popescu, N. Mastorakis. Multilayer perceptron and neural networks. WSEAS Trans. Circuits Syst. **8**, 579-588 (2009)
21. L. Breiman. Random forests. Mach. Learn. **45**, 5-32 (2001)
22. R. Genuer, J.M. Poggi, C. Tuleau-Malot, N. Villa-Vialaneix. Random forests for big data. Big Data Res. **9**, 28-46 (2020)
23. D.P. Solomatine, D.L. Shrestha. AdaBoost.RT: a boosting algorithm for regression problems. 2004 IEEE Int. Joint Conf. Neural Netw., vol. 2, 1163-1168 (2004)
24. D.L. Shrestha, D.P. Solomatine. Experiments with AdaBoost.RT, an improved boosting scheme for regression. Neural Comput. **18**, 1678-1710 (2006)
25. M. Awad, R. Khanna, M. Awad, R. Khanna. Support vector regression. In: Efficient learning machines (Springer, Berkeley, 2015), pp. 67-80
26. P.R. Lorenzo, J. Nalepa, L.S. Ramos, J.R. Pastor. Hyper-parameter selection in deep neural networks using parallel particle swarm optimization. Proc. Genetic Evol. Comput. Conf. Companion, 1864-1871 (2017)
27. R. Eberhart, J. Kennedy. A new optimizer using particle swarm theory. MHS'95 Proc. Sixth Int. Symp. Micro Mach. Hum. Sci., 39-43 (1995)
28. L. Tani, C. Veelken. Comparison of Bayesian and particle swarm algorithms for hyperparameter optimisation in machine learning applications in high energy physics. Comput. Phys. Commun. 294, 108955 (2024). https://doi.org/10.1016/j.cpc.2023.108955
29. M. Fang, C. Pan, X. Yu, W. Li, B. Wang, H. Zhou, Z. Xu, G. Yang. Risk prediction of hyperuricemia based on particle swarm fusion machine learning solely dependent on routine blood tests. BMC Med. Inform. Decis. Mak. 25, 131 (2025). https://doi.org/10.1186/s12911-025-02956-2
30. A. Chauhan, Shivaprakash S.J., Sabireen H., Abdul Quadir Md., Neelanarayanan Venkataraman. Stock price forecasting using PSO hypertuned neural nets and ensembling. Appl. Soft Comput. **147**, 110835 (2023). https://doi.org/10.1016/j.asoc.2023.110835
31. S. Hansbo. A new approach to the determination of the shear strength of clay by the fall-cone test. Proc. Royal Swedish Geotech. Inst. **14**, 1-47 (1957)