

An Integrated Machine Learning Framework for Predictive Assessment of Soil Bearing Capacity and Settlement Using Multisource Geotechnical and Environmental Data

Omokaro Idama¹, Ejiro Rufus Iwemah², Afolabi Awodeyi³, Omolegho A. Ibok⁴

^{1,3,4} Department of Computer Engineering, Faculty of Engineering, Southern Delta University, Ozoro, Delta State, Nigeria

² Department of Civil Engineering, Delta State University, Oleh Campus, Delta State, Nigeria.

Article Info

Article History:

Published: 06 Sept 2026

Publication Issue:

Volume 3, Issue 9
September-2026

Page Number:

74-91

Corresponding Author:

Omolegho A. Ibok

Abstract:

Reliable foundation design requires simultaneous control of ultimate bearing capacity and serviceability settlement, yet conventional prediction methods commonly treat soil parameters as static and omit rainfall, moisture, matric suction and groundwater fluctuations. This article develops a machine-learning-driven framework that integrates geotechnical, foundation and environmental data for joint prediction of shallow-foundation bearing capacity and settlement. A transparent proof-of-concept benchmark comprising 480 physically constrained synthetic records from 12 heterogeneous sites was created from accepted strength, stiffness and load-deformation relationships. Ridge regression, support vector regression, random forest, extra trees and gradient boosting were evaluated using site-grouped cross-validation and a holdout test containing entirely unseen sites. Gradient boosting achieved the best holdout performance, with R^2 values of 0.906 for ultimate bearing capacity and 0.893 for settlement. Adding environmental variables increased bearing-capacity R^2 from 0.894 to 0.906 and settlement R^2 from 0.772 to 0.893. Permutation analysis identified friction angle, footing width, cohesion and embedment as dominant bearing-capacity predictors, whereas soil modulus, service pressure, footing width, moisture and recent rainfall governed settlement. The framework combines data-quality control, explainability, uncertainty screening and engineering checks, offering a reproducible pathway for site-calibrated decision support rather than replacement of geotechnical judgement.

Keywords: bearing capacity; settlement; machine learning; environmental data; unsaturated soil; gradient boosting; explainable AI; shallow foundation

1. Introduction

Shallow foundations remain the most economical means of transmitting structural loads where competent near-surface soils are available. Their design is governed by two related but distinct limit states: adequate resistance against shear failure and acceptable deformation under service loading. An estimate of ultimate bearing capacity that ignores settlement may produce a foundation that is stable but functionally unacceptable; conversely, a low predicted settlement does not establish an adequate factor of safety against bearing failure (Phoon & Zhang, 2023). The practical task is therefore a coupled prediction problem involving geometry, loading, soil strength, stiffness, compressibility, drainage and the environmental state of the ground.

Conventional calculations are indispensable because they express mechanics and provide interpretable design checks. However, their inputs are frequently condensed into representative values from spatially variable investigations, while empirical correction factors can be sensitive to soil type and the range of the original calibration. Recent research has consequently applied machine learning to bearing capacity, settlement, soil-property prediction and constitutive modelling. Reviews report rapid growth in artificial neural networks, tree ensembles, support-vector methods and

hybrid optimization, but they also identify recurring weaknesses: small datasets, random train-test splits that leak site information, incomplete reporting, limited external validation and insufficient explanation of predictions (P. Zhang, Yin, & Jin, 2022; Yaghoubi et al., 2024; Zhao et al., 2024).

A second limitation is environmental stationarity. Water content, matric suction and groundwater position can change through wet and dry seasons. In unsaturated soils, suction contributes to apparent strength; infiltration and a rising groundwater table can reduce that contribution, alter effective stress, lower stiffness and increase deformation. Analytical, numerical and experimental studies show that footing response depends on suction distribution, foundation geometry, drainage and transient flow (Akbari Garakani et al., 2020; Tan & Vanapalli, 2022; Ghazavi & Molavi Sarghein, 2022; Yilmazoglu & Ozocak, 2023). Climate-related changes in precipitation and temperature further strengthen the case for incorporating environmental state variables in geotechnical assessment (Kandalai et al., 2023).

Most published machine-learning studies nevertheless concentrate on one target and one data family. Bearing-capacity models often use strength and geometry, while settlement models rely on SPT, CPT, footing size or applied pressure. Strong within-sample results do not necessarily demonstrate transfer to a new site, especially when observations from the same project appear in both training and testing. A defensible framework must separate sites during validation, distinguish explanatory from merely correlated variables, compare simple and nonlinear models, and retain physics-based plausibility checks. This is consistent with calls for trustworthy, data-efficient and domain-aware machine learning in geotechnics (Phoon & Zhang, 2023; H. Liu et al., 2024).

The aim of this article is to develop and demonstrate an integrated framework for predicting ultimate bearing capacity and settlement behaviour from combined geotechnical, foundation and environmental data.

The specific objectives are to:

- (i) formulate a coupled data architecture and theoretical basis;
- (ii) compare five regression algorithms under unseen-site validation;
- (iii) quantify the incremental value of environmental predictors;
- (iv) identify the dominant controls through permutation-based explainability; and
- (v) translate the findings into a deployable decision-support workflow.

The numerical demonstration uses a clearly labelled benchmark constrained by accepted geotechnical relationships. It is not represented as field evidence and must not be used for design without local calibration.

The article contributes in four ways. First, it treats resistance and settlement as parallel outputs within one data pipeline. Second, it integrates changing environmental conditions instead of assuming immutable soil properties. Third, it uses grouped validation so that test observations originate from sites excluded from training. Fourth, it embeds model governance data checks, uncertainty flags, explainability, engineering bounds and recalibration within the proposed architecture. These features move the discussion from isolated algorithm accuracy toward reliable application in foundation planning, monitoring and climate-aware asset management.

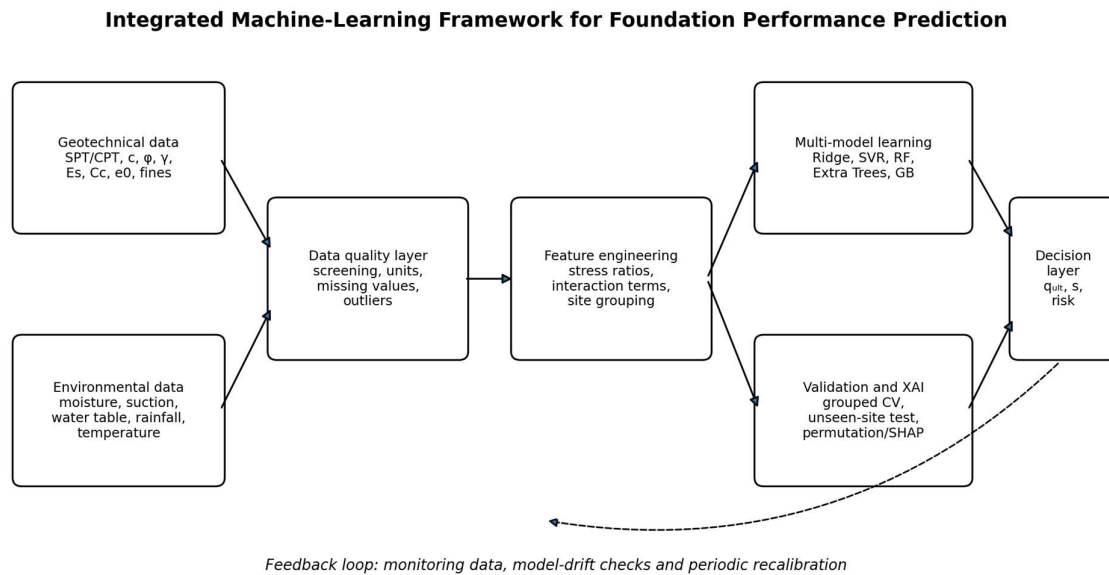


Figure 1. Integrated machine-learning framework for foundation performance prediction.

2. Literature Review

2.1 Conceptual Review

2.1.1 Soil Bearing Capacity

Bearing capacity is the maximum contact pressure that a soil-foundation system can sustain before a defined failure mechanism or excessive deformation occurs. For a shallow strip footing, the conventional form combines contributions from cohesion, surcharge and soil self-weight. Shape, depth, load inclination, groundwater and stratification modify these terms. In unsaturated ground, suction-dependent effective stress can increase mobilized strength, although the benefit is not permanent when infiltration reduces suction. Akbari Garakani et al. (2020) demonstrated the relevance of suction-dependent effective stress through analytical, numerical and experimental comparison, while Ghazavi and Molavi Sarghein (2022) and Yilmazoglu and Ozocak (2023) proposed or validated approaches that explicitly account for unsaturated response.

For modelling purposes, ultimate bearing capacity is not controlled by one variable. Footing width and embedment affect the stress field and failure zone; cohesion and friction angle define shear resistance; unit weight controls the self-weight term; groundwater and moisture influence effective stress; and stratification changes the governing mechanism. Liu and Liang (2024) showed that integrated neural models can capture nonlinear footing-soil interactions, but their study also illustrates the need for careful attention to data scarcity and generalizability. A machine-learning model should therefore learn interactions without discarding known monotonic expectations, such as increasing capacity with greater friction angle under otherwise comparable conditions.

2.1.2 Settlement Behaviour

Settlement is the downward movement of a foundation caused by deformation of the supporting soil. It may include immediate or elastic settlement, primary consolidation and time-dependent secondary compression. The relative importance of these components depends on soil type, permeability, stress history, drainage path and loading duration. In granular soils, immediate settlement commonly dominates; in saturated fine-grained deposits, consolidation may

govern. Serviceability is often more restrictive than ultimate resistance because differential or total movements can damage finishes, utilities and structural alignment at loads well below failure.

Recent machine-learning studies have modelled settlement using footing geometry, applied load, SPT or CPT indices and soil descriptors. Aouadj and Bouafia (2022) combined CPT information with an artificial neural network and mathematical formulation to reproduce load-settlement behaviour. Ngo and Tran (2024) compared gradient boosting, random forest, support-vector and nearest-neighbour methods and reported strong performance for gradient boosting. Khajehzadeh et al. (2024) further demonstrated the value of optimized nonlinear models. These studies support data-driven prediction, but most datasets remain focused on cohesionless soils and do not systematically integrate changing environmental conditions.

2.1.3 Integrated Geotechnical and Environmental Data

Integrated data refer to synchronized information from site investigation, laboratory testing, foundation design and environmental observation. Geotechnical variables include classification, density or unit weight, cohesion, friction angle, stiffness, compressibility, SPT resistance, CPT cone resistance, void ratio and fines content. Foundation variables describe width, embedment and service pressure. Environmental variables characterize moisture, matric suction, groundwater depth, antecedent rainfall and temperature. Their integration is conceptually important because environmental changes act through the mechanical state rather than as unrelated covariates.

A useful integrated record must be traceable to a site, depth, test method, date and quality grade. Units should be standardized, and laboratory values should not be combined with in-situ indices without metadata. Rainfall needs a meaningful accumulation window; groundwater measurements require a consistent datum; and suction measurements should identify method and range. Data integration is therefore an engineering information-management task before it becomes a machine-learning task. Poor alignment between a soil test and an environmental observation can create spurious patterns that even sophisticated algorithms will learn.

2.1.4 Machine-Learning Predictive Modelling

Machine learning estimates a mapping from predictors to a target using patterns in data. Linear regularized models provide transparent baselines; support-vector regression can represent smooth nonlinear relationships; random forest and extra trees capture interactions and are relatively robust to scaling; and gradient boosting sequentially corrects residual errors. Broad geotechnical reviews find that no algorithm is universally best, and performance is strongly conditioned by sample size, feature selection, validation design and parameter tuning (W. Zhang et al., 2022; Gao, 2024; Yaghoubi et al., 2024; Zhao et al., 2024).

The principal attraction of tree ensembles is their ability to represent thresholds and interactions without requiring a single assumed equation. Their limitations include extrapolation risk, sensitivity to unrepresentative training data and possible overconfidence outside the observed domain. Explainability methods can rank influential variables, but importance is not proof of causality. The engineering interpretation must test whether the learned direction and magnitude agree with mechanics. Physics-informed and hybrid methods offer a complementary path by embedding governing equations or constitutive constraints within learning (Bekele, 2021; Guan et al., 2023; H. Liu et al., 2024).

2.2 Theoretical Framework

The proposed theoretical framework combines effective-stress reasoning, load-deformation theory and statistical learning. Effective stress explains how total stress and pore-fluid conditions govern soil skeleton behaviour. In unsaturated soil, matric suction contributes to strength and stiffness through a suction-dependent stress state. Rainfall infiltration or groundwater rise can reduce suction, increase degree of saturation and modify both resistance and compressibility. This pathway justifies treating moisture, suction, rainfall and groundwater as state variables rather than background descriptors (Akbari Garakani et al., 2020; Tan & Vanapalli, 2022).

Bearing resistance is represented by the generic mechanics structure in Equation 1, where c is cohesion, q is surcharge at foundation level, γ is unit weight, B is footing width, N terms are bearing factors and F terms collect shape, depth, inclination and environmental corrections. The equation is not used as a substitute label generator for real projects; it provides the directional constraints used to create the proof-of-concept benchmark and the plausibility rules used after prediction.

$$q_{ult} = cNcFc + qNqFq + 0.5\gamma BN\gamma F\gamma \quad (1)$$

Settlement theory is represented by immediate and consolidation components. A simplified immediate-settlement expression relates service pressure and footing width to soil modulus, Poisson effects and an influence factor, while consolidation depends on compressibility, initial void ratio, layer thickness and the change in effective stress. Equations 2 and 3 show the conceptual structure. Their value in the framework is not that one closed form describes every soil, but that they establish expected dependencies for feature selection, scenario testing and prediction screening.

$$s_i = (qB/_s)(1 - \nu^2)I_s \quad (2)$$

$$s_c = [CcH/(1 + e_0)] \log_{10}[(\sigma'_0 + \Delta\sigma')/\sigma'_0] \quad (3)$$

Statistical learning supplies the flexible function that links the integrated predictor vector to the two outputs. For target y and predictor vector x , the fitted model minimizes a regularized empirical loss. Site-grouped cross-validation estimates performance when entire sites are withheld, which is more demanding and more relevant than a random record split. The framework therefore accepts a machine-learning model only when it improves on a transparent baseline, remains stable across site folds, produces plausible sensitivity, and operates within a documented applicability domain.

The synthesis is a physics-guided rather than physics-replacing position. Mechanics determines variables, constraints and rejection rules; data-driven learning represents residual nonlinearities and interactions; and monitoring provides updating evidence. Bekele (2021) illustrated how governing consolidation physics can be incorporated into deep learning, while Guan et al. (2023) integrated a learned soil model into finite-element analysis. The present framework is simpler and deployable with routine data, but it follows the same principle that predictive accuracy should be reconciled with physical meaning.

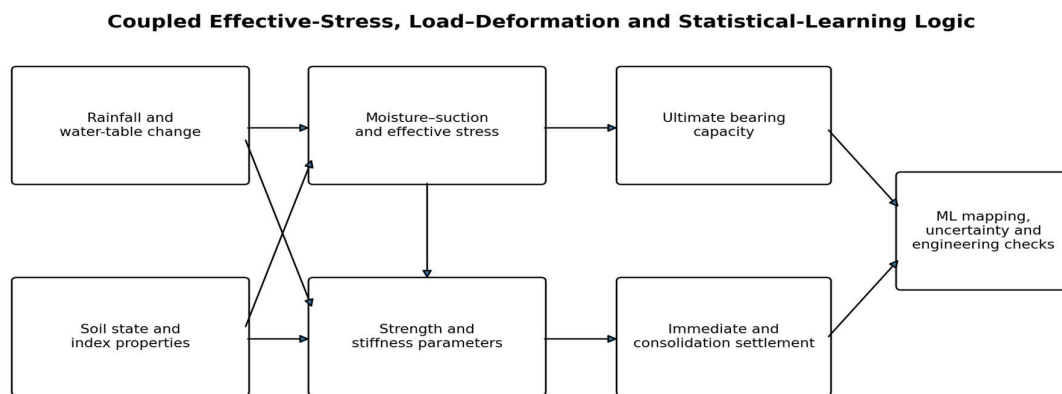


Figure 2. Coupled theoretical framework linking environmental forcing, effective stress, strength, stiffness and machine learning.

2.3 Empirical Review

Table 1 synthesizes recent empirical, analytical and review studies that define the current evidence base. The studies demonstrate strong progress in nonlinear prediction and unsaturated-soil mechanics, but they also reveal a separation between bearing-capacity models, settlement models and environmental analysis.

Table 1. Empirical Review of Recent Studies (2020–2024)

Study	Data/context	Target	Method	Main finding	Limitation relevant to this study
Akbari Garakani et al. (2020)	Analytical, 3D numerical and experimental footing response	Bearing capacity on unsaturated soil	Suction-dependent effective-stress approach	Capacity depends on soil properties, geometry and embedment; numerical response agreed with the effective-stress concept.	Not a data-driven multi-target framework.
Bekele (2021)	One-dimensional consolidation cases	Consolidation response	Physics-informed deep learning	Governing consolidation physics can constrain learning for forward and inverse analysis.	Focused on idealized 1D consolidation rather than routine foundation datasets.
Aouadj & Bouafia (2022)	CPT-based shallow-footing observations in sand	Load-settlement curve	Hybrid ANN-mathematical model	Hybridization improved prediction of load-settlement behaviour from CPT-related inputs.	Environmental state and cross-site transfer were not central.
Tan & Vanapalli (2022)	Transient-flow unsaturated slope-foundation analysis	Bearing capacity	Slip-line analytical method	Transient hydraulic conditions and suction profiles materially affect foundation capacity.	Analytical study without machine-learning integration.
P. Zhang et al. (2022)	Published geotechnical property datasets	Soil properties	Comparative machine learning review/tool	Algorithm suitability depends on dataset, feature structure and evaluation design.	Broad property focus; not a unified footing-performance model.
Guan et al. (2023)	Stress-strain data and finite-element boundary-value problems	Constitutive response	Deep-learning soil model within FEM	Learned constitutive behaviour can be embedded in numerical geotechnical analysis.	Requires richer stress-strain data and computational integration.

Yilmazoglu & Ozocak (2023)	Direct shear and model-footing tests on unsaturated silt	Ultimate bearing capacity	Experimental and analytical equation	Matric suction improved capacity and the proposed relation agreed with model tests.	Single soil domain; no settlement or ML validation.
Liu & Liang (2024)	200 bearing-capacity records for layered footing systems	Stability/bearing capacity	ANN optimized with BSA and EO	Integrated optimization improved neural prediction under nonlinear footing-soil interaction.	Small dataset and limited evidence of external-site generalization.
Ngo & Tran (2024)	Shallow-foundation settlement database	Settlement	GB, RF, SVM, KNN with tuning	Gradient boosting achieved the strongest reported test accuracy and useful feature ranking.	Environmental variables and bearing capacity were outside scope.
Khajehzadeh et al. (2024)	Cohesionless-soil shallow-foundation cases	Settlement	Optimized LSTM, SVR and MLP	Optimization improved nonlinear settlement prediction, with optimized LSTM strongest.	Algorithmically intensive; narrow soil and target coverage.
Zhao et al. (2024)	Multiple published geotechnical-property datasets	Geotechnical properties	ML comparison with feature selection	Feature selection and validation strongly influence the optimal model.	Review scope does not provide a deployable dual-target footing workflow.
Yaghoubi et al. (2024)	Systematic review/meta-analysis across geotechnics	Multiple geotechnical targets	ANN, ML, DL and ensembles	Ensembles and nonlinear models are widely effective, but data and reporting quality constrain conclusions.	Heterogeneous evidence; limited common external validation.

2.4 Gap in Literature

Recent studies demonstrate that machine learning can predict footing resistance or settlement, but the evidence remains fragmented. Bearing-capacity research emphasizes geometry and strength parameters, whereas settlement studies rely on load, SPT/CPT indices and stiffness; few models estimate both outcomes through one governed pipeline. Environmental state is underrepresented, despite evidence that rainfall, suction and groundwater transitions alter effective stress and footing response (Tan & Vanapalli, 2022; Yilmazoglu & Ozocak, 2023). Many studies use random record splits, which can place observations from one site in both training and testing and inflate generalization. Reviews consequently call for external validation, feature governance, uncertainty analysis and interpretable predictions (Phoon & Zhang, 2023; Zhao et al., 2024). Although Liu and Liang (2024) and Ngo and Tran (2024) report strong nonlinear models, neither supplies a unified environmental, bearing-capacity and settlement workflow. This study addresses that gap through integrated features, unseen-site validation, ablation testing and engineeringly constrained interpretation for practice.

3. Methodology

3.1 Research Design and Data Architecture

The study adopted a design-science and computational benchmarking approach. The output is an operational framework, while the numerical experiment tests whether the proposed data architecture can support accurate and interpretable dual-target prediction. The workflow comprised data definition, physically constrained benchmark generation, preprocessing, model development, site-grouped validation, holdout testing, feature-ablation analysis and explainability. All numerical results are demonstrative rather than claims about a named field site.

The benchmark contained 480 records allocated to 12 synthetic sites, with 40 records per site. Four broad soil classes sand, silty sand, sandy clay and clay were represented. Site-level parameter centers were first sampled to create realistic between-site heterogeneity. Record-level variation was then introduced for foundation geometry, load and soil state. Ultimate bearing capacity was generated from a conventional cohesion-surcharge-self-weight structure modified by geometry, groundwater and moisture-suction factors. Settlement combined an elastic term with a compressibility term and an environmental softening multiplier. Random noise and site effects were added to prevent an unrealistically deterministic dataset.

3.2 Variables and Benchmark Construction

The predictor set comprised footing width B , embedment depth D_f , service pressure, unit weight, cohesion, friction angle, deformation modulus E_s , compression index C_c , initial void ratio e_0 , SPT N , CPT q_c , clay and fines percentages, moisture content, matric suction, groundwater depth, 30-day antecedent rainfall and temperature, plus soil class. The targets were ultimate bearing capacity in kPa and settlement in millimetres. Values were bounded to plausible study ranges, but those ranges are not universal design limits.

Table 2. Variables Used in the Integrated Benchmark Dataset

Variable	Unit/type	Illustrative range	Role in framework
Soil class	Categorical	Sand, silty sand, sandy clay, clay	Stratification proxy and categorical context
Footing width, B	M	0.8–4.0	Failure-zone scale and settlement influence
Embedment depth, D_f	M	0.5–2.5	Surcharge and depth effect
Service pressure, q	KPa	70–450	Primary settlement demand
Unit weight, γ	kN/m ³	15–21	Self-weight bearing component
Cohesion, c	KPa	0–55	Shear-strength contribution
Friction angle, ϕ	Degrees	18–40	Shear resistance and bearing factors
Deformation modulus, E_s	MPa	6–90	Immediate settlement stiffness
Compression index, C_c	–	0.08–0.55	Consolidation compressibility
Initial void ratio, e_0	–	0.40–1.20	Compressibility state

SPT resistance, N	blows/0.3 m	3–45	In-situ density/consistency proxy
CPT cone resistance, q_c	MPa	0.8–14	In-situ resistance proxy
Clay content	%	3–65	Texture and compressibility
Fines content	%	5–85	Drainage and soil behaviour
Moisture content	%	7–38	Hydraulic/mechanical state
Matric suction	KPa	2–120	Unsaturated effective-stress contribution
Groundwater depth	M	0.2–6.0	Saturation and effective stress
30-day rainfall	Mm	0–420	Antecedent infiltration forcing
Temperature	°C	18–36	Seasonal environmental context
Ultimate bearing capacity	KPa	387–3,200	Target 1
Settlement	Mm	2–150	Target 2

3.3 Preprocessing and Candidate Models

Preprocessing was performed within each training fold to avoid information leakage. Numerical variables were standardized for ridge and support-vector regression; categorical soil class was one-hot encoded; tree models received the same pipeline for consistency. Missing-data imputation is included in the proposed operational framework, although the synthetic benchmark contained complete records. For field implementation, imputation parameters must be learned only from the training set, and a missingness indicator should be retained when absence may carry operational meaning.

Five candidate regressors were compared: ridge regression, radial-basis support-vector regression, random forest, extra trees and gradient boosting. Ridge provided a low-variance linear baseline. Support-vector regression represented smooth nonlinear relationships. Random forest and extra trees provided bagged tree ensembles, while gradient boosting constructed an additive sequence of shallow trees using a robust loss. Hyperparameters were selected conservatively to reduce overfitting rather than maximized through an extensive search on the holdout set.

3.4 Validation Strategy and Performance Metrics

Settlement was log-transformed during training because its distribution was positively skewed and extreme movements should not dominate the loss. Predictions were returned to millimetres before calculating engineering metrics. The main test divided the data by site: nine sites (360 records) formed the development set and three sites (120 records) formed the unseen-site holdout. A six-fold GroupKFold procedure, again separating sites, assessed stability across alternative site combinations.

Performance was measured with the coefficient of determination, root mean squared error and mean absolute error. R^2 measures explained variation, RMSE penalizes large errors and MAE expresses typical absolute deviation in target units. No single statistic was treated as sufficient. Model selection considered holdout accuracy, fold stability, physical plausibility and interpretability. Predictions outside the training domain or violating configured design bounds would be flagged rather than silently accepted.

$$R^2 = 1 - \Sigma(y_i - \hat{y}_i)^2 / \Sigma(y_i - \bar{y})^2 \quad (4)$$

$$RMSE = \sqrt{[(1/n)\Sigma(y_i - \hat{y}_i)^2]}; MAE = (1/n)\Sigma|y_i - \hat{y}_i| \quad (5)$$

3.5 Ablation, Explainability and Governance

Environmental value was tested through ablation. The best-performing algorithm was retrained using only geotechnical and foundation variables, then compared with the full integrated feature set on the same unseen sites. Permutation importance quantified how much holdout R^2 decreased when one variable was randomly disrupted. This method preserves the fitted model and directly measures predictive dependence, although correlated variables may share importance and rankings must not be interpreted as causal proof.

The proposed deployment layer adds uncertainty and governance not fully represented by point metrics. Each prediction should be accompanied by an applicability-domain score based on distance from the training distribution, an ensemble or conformal interval, a data-quality grade and a list of the most influential case-level features. A geotechnical engineer then compares the prediction with conventional calculations, allowable settlement criteria, factor-of-safety requirements and site observations. Model output is advisory until those checks are satisfied.

4. Results

4.1 Descriptive Characteristics

The benchmark spans a wide response range, with ultimate bearing capacity from approximately 387 to 3,200 kPa and settlement from 2 to 150 mm. The broad range was intentional because narrow synthetic data would make the learning task unrealistically easy. Bearing-capacity values were less skewed than settlement values, supporting the log transformation adopted for settlement. The three-site holdout also differed in combinations of soil class, environmental condition and foundation geometry, providing a meaningful transfer test.

Table 3. Selected Descriptive Statistics for the Benchmark Dataset (n = 480)

Variable	Mean	SD	Minimum	Maximum
Footing width (m)	2.37	0.94	0.80	3.99
Service pressure (kPa)	250.42	101.30	70.16	449.74
Cohesion (kPa)	19.54	18.65	0.00	54.95
Friction angle (°)	29.19	6.88	18.01	39.99
Deformation modulus (MPa)	38.70	24.08	6.08	89.95
Moisture content (%)	21.05	8.51	7.02	37.99
Matric suction (kPa)	52.67	34.29	2.03	119.95
30-day rainfall (mm)	207.84	121.94	0.76	419.43
Ultimate bearing capacity (kPa)	1,464.46	734.99	387.02	3,200.00
Settlement (mm)	25.21	30.61	2.00	150.00

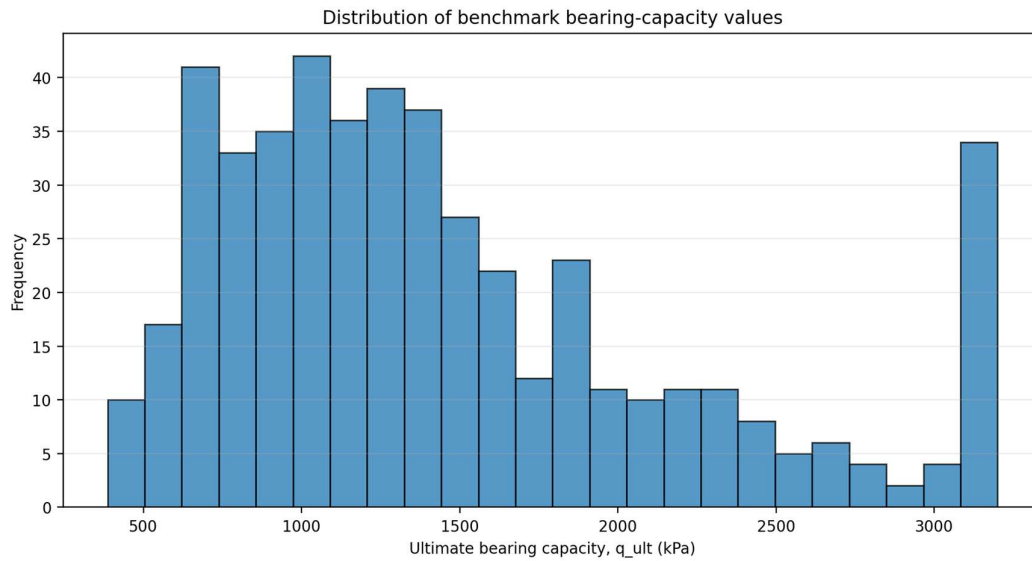


Figure 3. Distribution of synthetic benchmark ultimate bearing-capacity values.

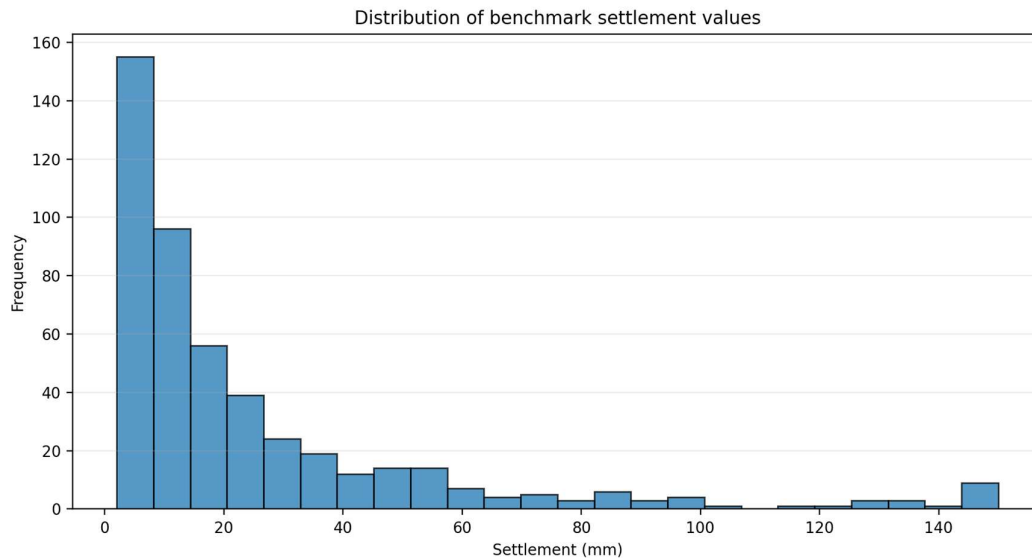


Figure 4. Distribution of synthetic benchmark settlement values.

4.2 Comparative Model Performance

Gradient boosting produced the strongest holdout result for both targets. For bearing capacity it achieved $R^2 = 0.906$, RMSE = 244.03 kPa and MAE = 188.86 kPa. Ridge, extra trees and random forest were competitive, whereas support-vector regression performed materially worse. For settlement, gradient boosting achieved $R^2 = 0.893$, RMSE = 7.95 mm and MAE = 4.48 mm. Ridge and support-vector regression were the next strongest. The relative ordering confirms that algorithm performance is target-dependent even within one integrated dataset.

Grouped cross-validation supported the holdout result. The selected bearing-capacity model achieved mean $R^2 = 0.896$ across six site folds, with a standard deviation of 0.039; mean RMSE and MAE were 213.82 and 165.61 kPa. The settlement model achieved mean $R^2 = 0.914$ with a standard deviation of 0.022; mean RMSE and MAE were 7.74 and 4.20 mm. The variability in RMSE reflects differences in target scale among site folds, which reinforces the need to examine both normalized and unit-based metrics.

Table 4. Unseen-Site Holdout Performance for Ultimate Bearing Capacity

Model	R ²	RMSE (kPa)	MAE (kPa)
Gradient boosting	0.906	244.03	188.86
Ridge regression	0.897	255.68	203.65
Extra trees	0.893	260.92	209.11
Random forest	0.887	267.96	213.21
SVR-RBF	0.672	456.57	313.72

Table 5. Unseen-Site Holdout Performance for Settlement

Model	R ²	RMSE (mm)	MAE (mm)
Gradient boosting	0.893	7.95	4.48
Ridge regression	0.876	8.57	4.90
SVR-RBF	0.848	9.46	4.64
Random forest	0.806	10.70	6.27
Extra trees	0.700	13.30	8.00

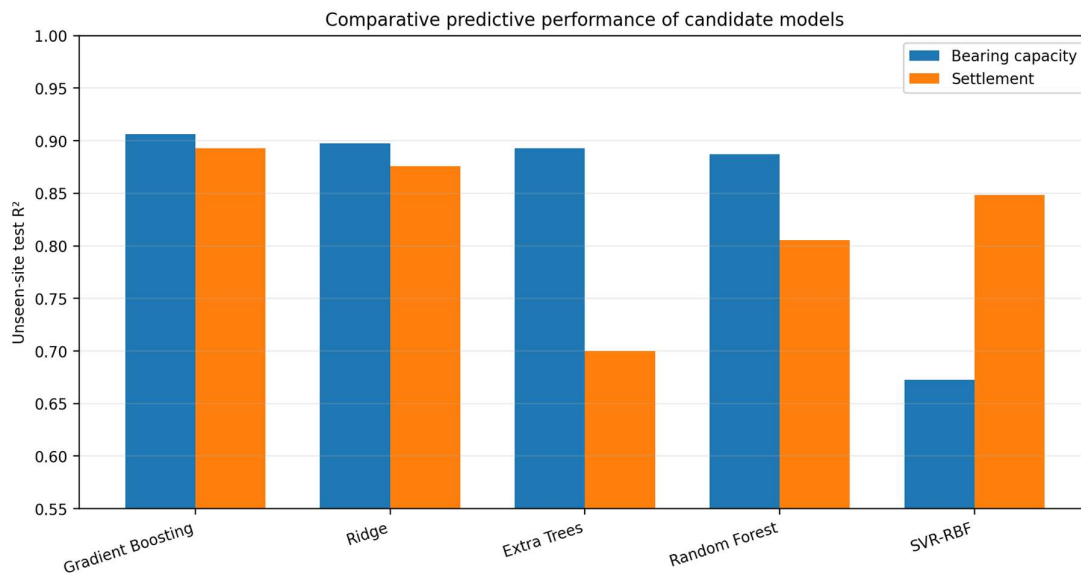


Figure 5. Comparative unseen-site R² values for the candidate models.

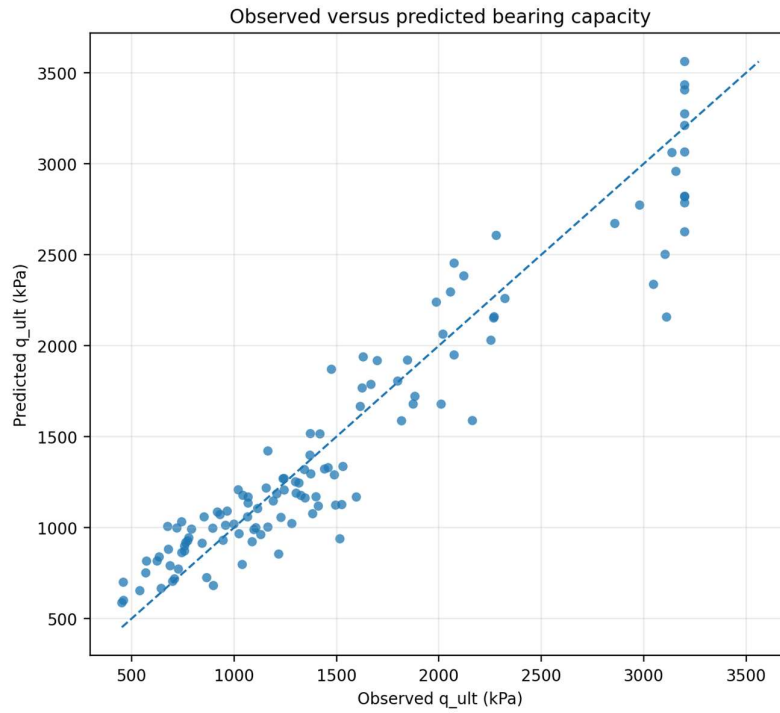


Figure 6. Observed versus predicted ultimate bearing capacity for unseen sites.

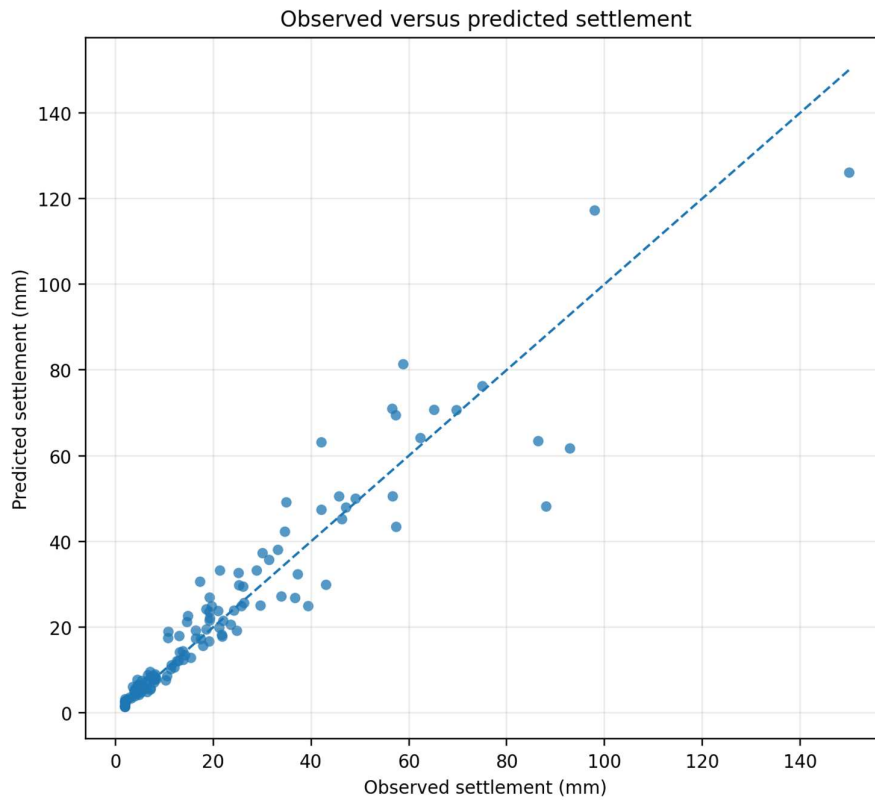


Figure 7. Observed versus predicted settlement for unseen sites.

4.3 Environmental Ablation and Explainability

Environmental ablation produced the clearest substantive result. Removing moisture, suction, groundwater depth, rainfall and temperature reduced bearing-capacity R^2 from 0.906 to 0.894, a modest but consistent decline. The same removal reduced settlement R^2 from 0.893 to 0.772 and increased RMSE from 7.95 to 11.60 mm. Environmental information therefore supplied substantially more predictive value for deformation than for ultimate resistance in this benchmark.

Permutation analysis also matched the governing mechanics. Friction angle was the dominant bearing-capacity feature, followed by footing width, cohesion and embedment depth. Rainfall and moisture had lower direct importance but improved integrated performance. For settlement, deformation modulus had the largest importance, followed by service pressure and footing width. Moisture content and antecedent rainfall were the leading environmental predictors. The observed-versus-predicted plots show close concentration around the equality line, with larger deviations concentrated at the highest responses where fewer training cases were available.

Table 6. Environmental-Feature Ablation Results

Target	Feature set	R^2	RMSE	MAE
Bearing capacity	Geotechnical only	0.894	260.07 kPa	204.62 kPa
Bearing capacity	Integrated	0.906	244.03 kPa	188.86 kPa
Settlement	Geotechnical only	0.772	11.60 mm	7.08 mm
Settlement	Integrated	0.893	7.95 mm	4.48 mm

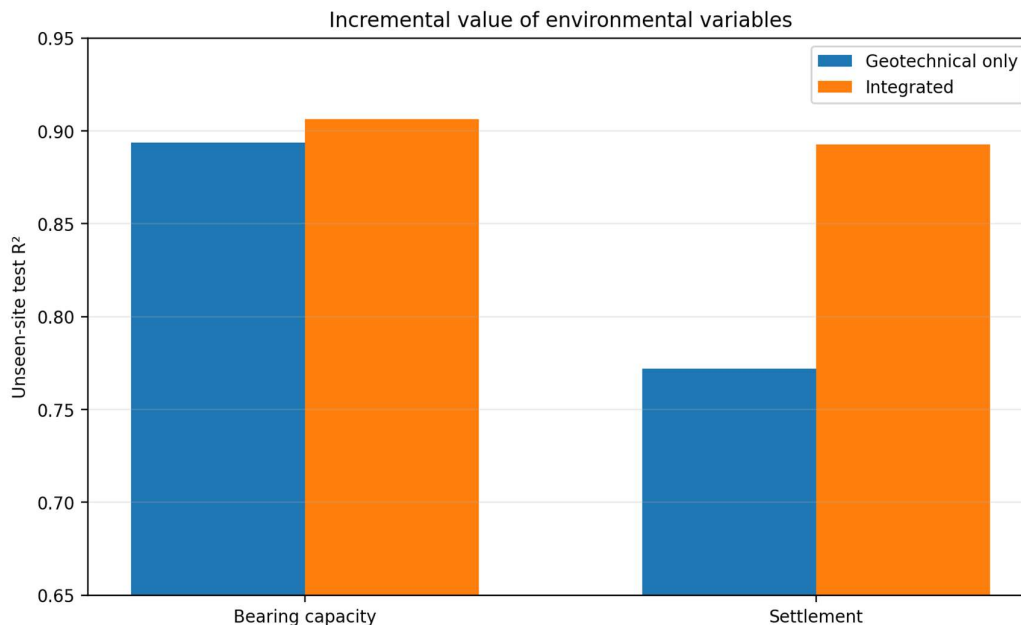


Figure 8. Incremental predictive value of environmental variables.

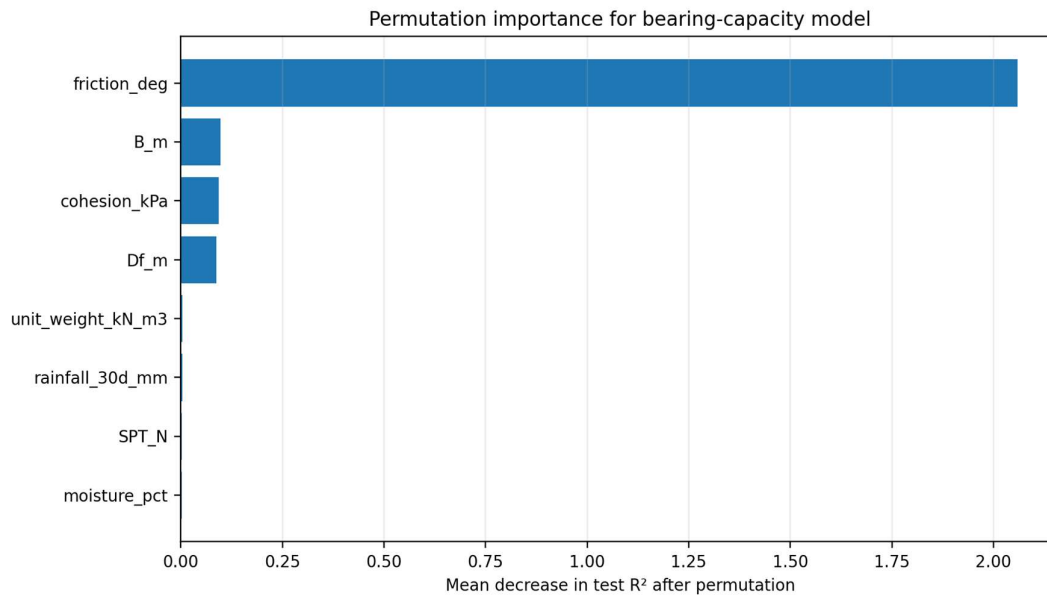


Figure 9. Permutation importance for ultimate bearing-capacity prediction.

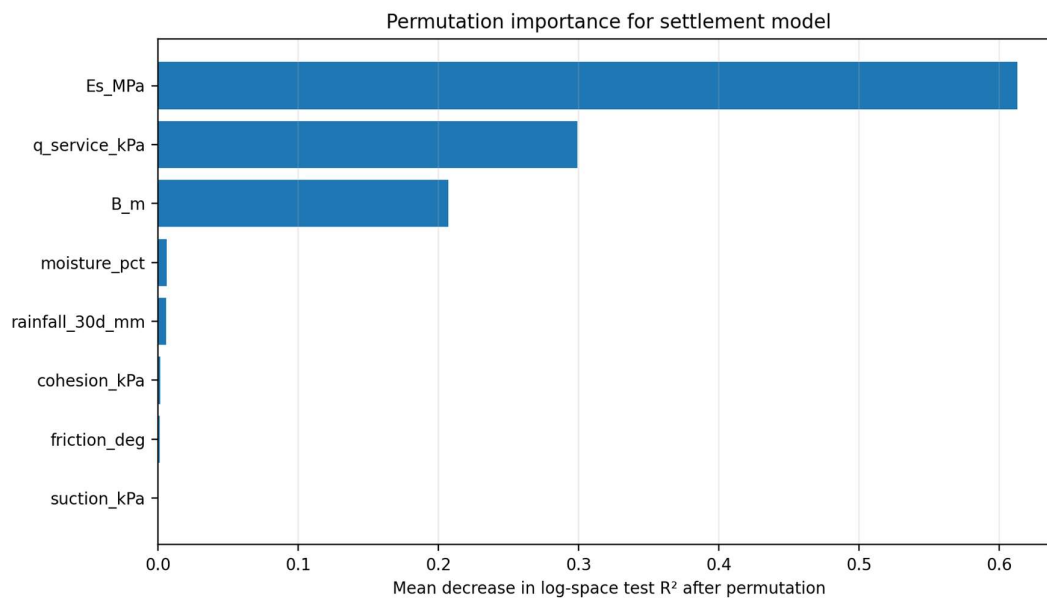


Figure 10. Permutation importance for settlement prediction.

5. Discussion

The findings support the central proposition that combined data improve prediction, but they also show that the benefit is not uniform across engineering outputs. Ultimate bearing capacity is primarily governed by strength and geometry, so adding environmental state yielded an incremental improvement. Settlement is more directly sensitive to stiffness and compressibility changes associated with wetting, suction loss and groundwater variation; the environmental contribution was consequently much larger. This distinction agrees with unsaturated-soil studies showing that transient flow modifies footing response and with climate-geotechnics research emphasizing changes in hydrological boundary conditions (Tan & Vanapalli, 2022; Kandalai et al., 2023).

Gradient boosting performed well because shallow trees can represent interactions such as friction angle by footing width, pressure by modulus and rainfall by moisture without assuming one global linear form. Similar strength has

been reported in recent settlement modelling (Ngo & Tran, 2024). Nevertheless, the competitive ridge result is important. It indicates that a well-constructed feature set may capture much of the signal with a simpler model. A complex model should therefore be selected for demonstrated generalization, not novelty.

The unseen-site protocol is a more credible test than a random record split. When multiple observations share site geology, investigation method or construction context, random splitting can allow the model to recognize site signatures rather than transferable relationships. The lower and more variable grouped scores are not a weakness of the study; they provide a less optimistic estimate of deployment performance. This directly responds to the generalization concerns emphasized by Phoon and Zhang (2023) and Zhao et al. (2024).

The synthetic benchmark is also a limitation. It is valuable for testing workflow integrity because the generating relationships and environmental pathways are known, but it cannot reproduce all features of natural deposits, including anisotropy, cementation, fissuring, stress history, sampling disturbance and complex stratification. Reported errors should not be interpreted as expected errors on Nigerian, tropical or any other field site. A publication-quality empirical application would replace the benchmark with multi-season investigation and monitoring data and preserve the same validation logic.

From a safety perspective, machine learning must not directly issue a final allowable pressure. A safe decision combines predicted ultimate capacity with an appropriate factor of safety, settlement criteria, conventional calculations, uncertainty bounds and professional review. Cases near the boundary of the training domain should trigger additional investigation rather than automated confidence. This human-in-the-loop position is consistent with the broader literature on trustworthy AI-enhanced geotechnical mechanics (H. Liu et al., 2024).

6. Proposed Operational and Deployment Framework

Operational implementation begins with a site data model. Every borehole, sounding, laboratory specimen, foundation element and sensor reading receives a unique identifier and coordinates. Depth intervals and timestamps are explicit. A controlled dictionary defines units, methods and quality codes. This prevents the common error of combining values that appear numerically compatible but represent different stress conditions or test procedures.

The modelling layer trains separate target models within a shared pipeline. A baseline model is always retained. Site-grouped cross-validation, temporal back-testing and an external project holdout are preferred when data volume permits. Hyperparameter search is nested inside the training process, and the final test set remains untouched until model selection is complete. Model cards document data coverage, intended use, exclusions, metrics and known failure modes.

Table 7. Recommended Model-Governance Controls

Control stage	Minimum requirement	Decision consequence
Data acceptance	Traceable site/depth/time, standard units, quality flag	Reject or quarantine records that fail critical checks
Model validation	Grouped site/project split and untouched external holdout	No deployment based only on random record split
Applicability domain	Range and multivariate distance screening	Flag extrapolative cases for additional investigation
Uncertainty	Prediction interval or calibrated error band	Do not present point estimate as certainty

Engineering review	Conventional calculation, code criteria and plausibility check	Engineer approves, modifies or rejects model advice
Monitoring	Bias and drift checks by site, soil class and season	Recalibrate or withdraw model when performance degrades

7.0 Conclusion

This article developed a machine-learning-driven framework for joint prediction of soil bearing capacity and settlement using integrated geotechnical, foundation and environmental data. The proof-of-concept benchmark showed that gradient boosting generalized strongly to unseen synthetic sites and that environmental variables provided a particularly large improvement for settlement. The resulting feature patterns were consistent with mechanics: strength and geometry governed bearing capacity, while modulus, pressure, width and wetting-related variables governed settlement.

The principal implication is that foundation intelligence should not be organized as isolated algorithms. It should be organized as a governed system in which data lineage, mechanics, site-wise validation, uncertainty, explainability and professional review are connected. The framework is therefore suitable as a research and implementation blueprint, not as a substitute for site investigation or code-compliant design.

Future empirical work should assemble multi-site and multi-season datasets containing matched SPT/CPT, laboratory strength and compressibility tests, groundwater, rainfall, moisture or suction, foundation geometry, applied loads and measured settlement. It should compare point predictions with probabilistic or conformal intervals, test transfer across geological regions and investigate physics-constrained boosting or neural models. Most importantly, model success should be judged by safe decisions and calibrated uncertainty, not R^2 alone.

7.1 Recommendations

- I. Geotechnical organizations should establish standardized, site-linked databases before investing in complex algorithms.
- II. Environmental measurements should be aligned with investigation and performance records, using explicit time windows and quality codes.
- III. Validation should withhold entire sites or projects; random record splitting should not be the sole evidence of generalization.
- IV. Every deployed model should retain a conventional baseline, uncertainty interval, applicability-domain warning and engineering approval step.
- V. Field studies should replace the illustrative benchmark and report failures, outliers and seasonal drift as carefully as headline accuracy.

References

1. Akbari Garakani, A., Sadeghi, H., Saheb, S., & Lamei, A. (2020). Bearing capacity of shallow foundations on unsaturated soils: Analytical approach with 3D numerical simulations and experimental validations. *International Journal of Geomechanics*, 20(3), 04019181. [https://doi.org/10.1061/\(ASCE\)GM.1943-5622.0001589](https://doi.org/10.1061/(ASCE)GM.1943-5622.0001589)
2. Aouadj, A., & Bouafia, A. (2022). CPT-based method using hybrid artificial neural network and mathematical model to predict the load-settlement behaviour of shallow foundations. *Geomechanics and Geoengineering*, 17(1), 321–333. <https://doi.org/10.1080/17486025.2020.1755459>
3. Bekele, Y. W. (2021). Physics-informed deep learning for one-dimensional consolidation. *Journal of Rock Mechanics and Geotechnical Engineering*, 13(2), 420–430. <https://doi.org/10.1016/j.jrmge.2020.09.005>

4. Gao, W. (2024). The application of machine learning in geotechnical engineering. *Applied Sciences*, 14(11), 4712. <https://doi.org/10.3390/app14114712>
5. Ghazavi, M., & Molavi Sarghein, A. (2022). Analytical method for determining bearing capacity of shallow foundations on unsaturated soils. *International Journal of Geomechanics*, 22(11), 04022195. [https://doi.org/10.1061/\(ASCE\)GM.1943-5622.0002530](https://doi.org/10.1061/(ASCE)GM.1943-5622.0002530)
6. Guan, Q., Yang, Z., Guo, N., & Hu, Z. (2023). Finite element geotechnical analysis incorporating deep learning-based soil model. *Computers and Geotechnics*, 154, 105120. <https://doi.org/10.1016/j.compgeo.2022.105120>
7. Kandalai, S., John, N. J., & Patel, A. (2023). Effects of climate change on geotechnical infrastructures—state of the art. *Environmental Science and Pollution Research*, 30, 16878–16904. <https://doi.org/10.1007/s11356-022-24788-7>
8. Khajehzadeh, M., Keawsawasvong, S., Kamchoom, V., Shi, C., & Khajehzadeh, A. (2024). Developing effective optimized machine learning approaches for settlement prediction of shallow foundation. *Heliyon*, 10(17), e36714. <https://doi.org/10.1016/j.heliyon.2024.e36714>
9. Liu, H., Su, H., Sun, L., & Dias-da-Costa, D. (2024). State-of-the-art review on AI-enhanced computational mechanics in geotechnical engineering. *Artificial Intelligence Review*, 57, 196. <https://doi.org/10.1007/s10462-024-10836-w>
10. Liu, Y., & Liang, Y. (2024). Integrated machine learning for modeling bearing capacity of shallow foundations. *Scientific Reports*, 14, 8319. <https://doi.org/10.1038/s41598-024-58534-5>
11. Ngo, T. T. H., & Tran, V. Q. (2024). Predicting and evaluating settlement of shallow foundation using machine learning approach. *Science Progress*, 107(4), 00368504241302972. <https://doi.org/10.1177/00368504241302972>
12. Phoon, K.-K., & Zhang, W. (2023). Future of machine learning in geotechnics. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 17(1), 7–22. <https://doi.org/10.1080/17499518.2022.2087884>
13. Tan, M., & Vanapalli, S. K. (2022). Foundation bearing capacity estimation on unsaturated soil slope under transient flow condition using slip line method. *Computers and Geotechnics*, 148, 104804. <https://doi.org/10.1016/j.compgeo.2022.104804>
14. Yaghoubi, E., Yaghoubi, E., Khamees, A., & Vakili, A. H. (2024). A systematic review and meta-analysis of artificial neural network, machine learning, deep learning, and ensemble learning approaches in field of geotechnical engineering. *Neural Computing and Applications*, 36, 12655–12699. <https://doi.org/10.1007/s00521-024-09893-7>
15. Yilmazoglu, M. U., & Ozocak, A. (2023). Bearing capacity of shallow foundations on unsaturated silty soils. *Applied Sciences*, 13(3), 1308. <https://doi.org/10.3390/app13031308>
16. Zhang, P., Yin, Z.-Y., & Jin, Y.-F. (2022). Machine learning-based modelling of soil properties for geotechnical design: Review, tool development and comparison. *Archives of Computational Methods in Engineering*, 29(2), 1229–1245. <https://doi.org/10.1007/s11831-021-09615-5>
17. Zhang, W., Gu, X., Tang, L., Yin, Y., Liu, D., & Zhang, Y. (2022). Application of machine learning, deep learning and optimization algorithms in geoengineering and geoscience: Comprehensive review and future challenge. *Gondwana Research*, 109, 1–17. <https://doi.org/10.1016/j.gr.2022.03.015>
18. Zhao, T., Shen, F., & Xu, L. (2024). Review and comparison of machine learning methods in developing optimal models for predicting geotechnical properties with consideration of feature selection. *Soils and Foundations*, 64(6), 101523. <https://doi.org/10.1016/j.sandf.2024.101523>