Challenge Journal of STRUCTURAL MECHANICS

Research Article

Comparative machine learning study for allowable bearing capacity prediction of OHTL tower foundations with spatially imputed geotechnical borehole databases' interpretability

Esra Uray [a,*], Tahir Yildiz [b]

[a] *Department of Civil Engineering, KTO Karatay University, 42020 Konya, Türkiye*
[b] *Turkish Petroleum Offshore Technology Center (TP-OTC), 06530 Ankara, Türkiye*

ABSTRACT

Geotechnical site characterization for alignment-based infrastructure projects, such as overhead transmission lines (OHTLs), involves the collection of extensive borehole datasets across hundreds of kilometers and the interpretation of thousands of results from laboratory and in-situ tests. In this study, a machine learning (ML) framework is proposed for predicting the allowable bearing capacity (q_{all}, kPa) of OHTL tower spread footings using geotechnical borehole data. A dataset comprising 89 boreholes with 16 input parameters, including physical, mechanical, and seismic soil properties, was used. Missing borehole records (n=8) were spatially imputed using Inverse Distance Weighting (IDW) and K-Nearest Neighbors regression prior to model training. Ten ML algorithms, including Ridge, Lasso, ElasticNet, Support Vector Regression, K-Nearest Neighbors, Random Forest, Extra Trees, Gradient Boosting, XGBoost, and LightGBM, were trained. Model performance was assessed using coefficient of determination (R^2), root mean square error, mean absolute error, and mean absolute percentage error metrics. Random Forest was identified as the most reliable model for practical deployment, exhibiting balanced generalization behavior with training R^2 converging from 0.63 to 0.98 and a Train-CV Gap of 0.017, confirming that it learns generalizable patterns rather than memorizing individual observations. Although Gradient Boosting achieved the highest overall performance metrics, its persistent Train R^2 of 1.000 indicated memorization behavior. SHAP-based interpretability analysis identified groundwater depth, shear strength parameters, and standard penetration test blow count as the primary effective parameters of *the* q_{all} value. The proposed framework shows that ML-based approaches can significantly enhance the reliability of bearing capacity predictions along transmission line corridors.

ARTICLE INFO

Article history:
Received – April 6, 2026
Revision requested – April 27, 2026
Revision received – May 15, 2026
Accepted – May 26, 2026

Keywords:
Bearing capacity
OHTL
Borehole database
Machine learning
Geotechnical characterization
Spatial imputation

Citation: Uray E, Yildiz T (2026). Comparative machine learning study for allowable bearing capacity prediction of OHTL tower foundations with spatially imputed geotechnical borehole databases' interpretability. *Challenge Journal of Structural Mechanics*, 12(2), 125–144.

1. Introduction

Geotechnical engineering examines the influence of natural materials, such as soil and rock, on structures and their foundations. It focuses on understanding, evaluating, and using the behavior of these materials in engineering projects. Geotechnical engineering primarily aims to examine the physical and mechanical characteristics of the soil environment that will carry the structure loads. Within this scope, soil investigations and research, including bearing capacity and settlement analyses and behavior under loads, are conducted employing diverse techniques for the soils upon which structures will be constructed. The safe foundation design of overhead

* *Corresponding author*. E-mail address: esra.uray@karatay.edu.tr (E. Uray)
ISSN: 2149-8024 / DOI: https://doi.org/10.20528/cjsmec.2026.02.006

transmission line (OHTL) towers necessitates detailed investigations due to their exposure to complex multi-axial soil-structure interaction (SSI) demands on their foundations. Although steel tower systems are generally advantageous under seismic actions because of their relatively low self-weight and high ductility, their foundation performance still depends on the reliable transfer of vertical, horizontal, and overturning effects to the supporting soil (Çarbaş et al. 2025). Since OHTL systems are critical lifeline infrastructure that may extend through seismically active regions, their foundation performance should be evaluated not only for load bearing capacity but also for the potential consequences of seismic induced service disruption or structural instability (Soysal 2025). Vertical compression, horizontal shear, and overturning moments arising from conductor tension, wind loads, ice accretion, and dynamic effects necessitate a rigorous assessment of the soil's ultimate bearing capacity and its allowable bearing capacity. In this context, soil properties and layering directly govern load transfer mechanisms, foundation deformations, and overall stability under combined vertical, lateral, and moment effects, making site-specific geotechnical characterization essential for reliable foundation design (Aliyev et al. 2026). Allowable bearing capacity (q_{all}) is an important factor in design that determines the size of the footing, how deep it should be placed, and the safety measures needed to prevent failure when loads are applied together. Under the influence of multidirectional loads, the combination of vertical (N) and horizontal (H) loads leads to the formation of an overturning moment (M), which causes irregular soil pressure distributions under the foundation. Many studies in the literature indicate that the effect in cases of irregular pressure distribution is directly dependent on q_{all} and is obtained by using bearing capacity factors with appropriate shape, depth, and/or slope corrections (Braja and Sivakugan 2017). This situation represents a fundamental issue in structural mechanics and requires an interaction framework to verify that the soil has a sufficient q_{all} value when subjected to combined loads (Demirci 2023; Yilmaz et al. 2014). The interaction envelope, encompassing overturning, shear, and bearing resistance, forms the basis of SSI, a central issue in structural mechanics for the investigation of shallow foundations under non-axial loading conditions; it is directly related to moment and shear force transfer mechanisms (Panagiotidou et al. 2012; Shafan and Yi 2024). After evaluating the loads and moments generated in the structure within a mechanical framework, soil stresses are also examined separately. In this approach, where soil stress is considered the primary criterion in bearing capacity and settlement assessment analyses, the design safety of structures built on the soil is investigated. Studies indicate that soil investigations are increasingly integrating in-situ field tests, numerical modeling, and soil improvement methods to accurately determine bearing capacity and settlement for various soil conditions and foundation types (Fahad et al. 2023; Thirmanpalli et al. 2024).

Alignment-based characterization is applied to soils where highways, railways, and power transmission lines will be constructed, and determining the soil profile's characteristics along a long route through field tests and investigations is of enormous importance. On the other hand, varying site conditions and soil characteristics along the route play a significant role in achieving an idealized soil scenario. This is because polygon-based characterization can be performed with a small number of ideal soil profiles. In contrast, route characterizations are more comprehensive, requiring consideration of many different factors, such as lithological variability, seismicity, and groundwater. In such complex soil investigation studies, the route is divided into sub-units based on a geotechnical principle (Heim et al. 2012). Consequently, alignment-based characterizations must consider spatially varied subsurface characteristics, fluctuating seismicity, groundwater regimes, and geohazard potential over the whole corridor. This route-based subdivision is motivated by the spatial heterogeneity of soils and hydrogeology encountered along long corridors, which governs bearing capacity, settlement, liquefaction susceptibility, ground-improvement needs, and monitoring strategies (Lin et al. 2019). Traditional investigation methods, which rely on idealized subsurface profiles, simplified lithological continuity assumptions, and generalized engineering interpretations, are inadequate for capturing localized subsurface anomalies and identifying critical geohazard risks along extended investigation corridors (Johnson 2009). Site characterization of alignment-based infrastructure projects can be done more rigorously and computationally efficiently using reliability-based design approaches (Mozer et al. 1984; Phoon et al. 2000).

Traditional bearing capacity theories (Meyerhof 1963; Terzaghi 1943; Vesić 1973) remain foundational for designating ultimate bearing capacity with adjustments made for soil type, footing geometry, and groundwater conditions. Several studies consider these classical methods to assess which yields safer, more economical designs under varying site conditions (Cao et al. 2025; Nawaz et al. 2022). In practice, field testing results, including standard penetration test (SPT), cone penetration test (CPT), and plate load test (PLT), are employed to determine empirical ultimate bearing capacity and then allowable bearing capacity. This orientation toward in situ and non-destructive assessment is consistent with recent structural health monitoring studies, where sensor-based systems are used to track critical material or structural responses without relying solely on destructive or time-consuming test procedures (Narwade and Jadhav 2025). Several studies have shown how SPT-based methods match or differ from measurements taken in situ. This stresses the need to adapt empirical models to site-specific conditions and verify them through numerical simulation when needed (Fahad et al. 2023; Nawaz et al. 2022; Telló et al. 2021). Recent research integrated machine learning and data-driven approaches to predict bearing capacity from in situ inputs (SPT, CPT, and soil parameters) and to bridge gaps between simplified bearing models and complex soil behavior. This data driven strategy is consistent with recent applications in construction materials, where machine learning models have been used to estimate critical properties such as compressive strength from multiple interacting input parameters, thereby reducing the need for repeated,

time consuming, and labour-intensive experimental procedures (Harirchian 2024). These approaches have demonstrated promise for improving predictions of bearing capacity in heterogeneous soils and for rapid assessments in practice, though challenges remain in interpretability and validation against physical models (Cao et al. 2025; Nguyen et al. 2024).

SPT data have long been a primary in-situ input for predicting bearing capacity and settlement of foundations. Recent advances in artificial intelligence (AI), particularly machine learning (ML), have offered powerful tools to learn complex, nonlinear relationships between SPT indices and geotechnical response, potentially improving prediction accuracy, robustness, and site-specific calibration beyond classical empirical correlations (Nejad et al. 2009; Shooshpasha et al. 2013; Wu et al. 2025). This modelling strategy is consistent with recent civil engineering applications in which ML algorithms establish relationships between multiple input features and target engineering properties, thereby capturing complex nonlinear patterns that are difficult to represent through conventional empirical formulations (More and Kambekar 2025). Artificial intelligence methods, including artificial neural networks (ANN), deep learning (DL), and ensemble learning (EL), have been widely employed for many years in solving problems encountered in geotechnical engineering topics (Baghbani et al. 2022; Yaghoubi et al. 2024). Especially, the application of machine learning (ML) in geotechnical engineering has gained considerable momentum over the past decade (Harle and Wankhade 2025). Previous studies have demonstrated the effectiveness of ensemble methods for predicting bearing capacity (Zeini et al. 2025), liquefaction susceptibility (Bherde et al. 2025), and pile capacity (Kordjazi et al. 2014).

Despite the growing body of literature on ML-based bearing capacity prediction, a critical gap remains at the intersection of three domains: OHTL tower foundation design, corridor-based geotechnical decision-making, and data-driven predictive modeling. Existing ML studies on bearing capacity prediction predominantly focus on single-site or point-scale datasets, where subsurface conditions are relatively uniform and borehole data are complete (Cao et al. 2025; Nguyen et al. 2024; Zeini et al. 2025). Corridor-scale infrastructure projects are complicated by geotechnical variability, spatial data sparsity, and the need for consistent foundation design decisions across hundreds of tower locations, which site-specific empirical methods cannot solve alone effectively. For OHTL transmission corridor studies, the alignment-based characterization of subsurface conditions must take into account lithological variability, seismicity zonation, groundwater fluctuations, and geohazard potential over large areas (Heim et al. 2012; Lin et al. 2019). This means that traditional investigation methods, which rely on idealized subsurface profiles and simplified lithological continuity assumptions, are not adequate enough to identify localized problems and make reliable foundation design decisions at each tower location (Johnson 2009). Additionally, SPT-based empirical correlations are still the most common way to estimate geotechnical parameters in corridor investigations. However, utilizing them to make predictions across spatially heterogeneous datasets may raise systematic uncertainty, which ML-based frameworks can address via nonlinear pattern recognition and cross-validation-based generalization assessment. The present study therefore bridges these three domains by developing an explicitly tailored ML pipeline to the spatial and geotechnical characteristics of corridor-scale OHTL foundation design, integrating spatial imputation, feature engineering, and model interpretability within a unified and practically deployable framework.

While machine learning-based bearing capacity estimation and SPT-based geotechnical regression are well-known research areas, this study focuses on a problem that has not been extensively explored. The development of a prediction model for the q_{all} value of a tower foundation within a wide-area corridor, considering incomplete borehole records and varying subsurface conditions, highlights unique aspects of this study: a dataset context and infrastructure application that have not been extensively addressed in the geotechnical machine learning literature. First, an ML-based q_{all} prediction framework is developed and benchmarked across ten ML algorithms using SPT-derived geotechnical data from an OHTL. Second, a novel spatial imputation decision matrix integrating Moran's I spatial autocorrelation index and a composite Spatial Score is proposed to guide the selection between Inverse Distance Weighting (IDW) and K-Nearest Neighbors (KNN) imputation for missing borehole records, directly addressing the data sparsity challenge inherent to corridor-scale geotechnical investigations. Third, SHAP-based interpretability analysis is conducted to verify that model predictions are driven by physically meaningful geotechnical parameters consistent with established bearing capacity theory. This study aims to offer a realistic, scalable, and interpretable tool for fundamental design assistance throughout extensive transmission line corridors, where conventional site-specific empirical methods are prohibitively expensive in terms of computing.

2. Data and Methodology

2.1. Theoretical background for bearing capacity predictions

During site investigations, groundwater levels were partially observed throughout the project area. Where fine-grained deposits were dominant along the investigation depth, undrained conditions were considered for regions with an observed groundwater level. The allowable bearing capacity (q_{all}) was obtained using the general bearing capacity formula given in Eq. (1), where F_s corresponds to a safety factor and q_{ult} is the ultimate bearing capacity.

$$q_{all} = \frac{q_{ult}}{F_s} \tag{1}$$

The bearing capacity of tower footings placed above fine-grained deposits with groundwater was calculated for undrained conditions per Terzaghi's (1943) bearing capacity equation for square footings, as given in Eq. (2).

$$q_{ult} = 1.3\times c\times N_c\times + q\times N_q + 0.4\times\gamma\times B\times N_\gamma \quad (2)$$

Bearing capacity factors N_c, N_q, and N_γ are dimensionless coefficients that govern the contributions of soil cohesion, overburden pressure, and self-weight to the ultimate bearing capacity, respectively. These factors are functions of the soil's internal friction angle ($\emptyset'$) and are determined based on established theoretical relationships. As the internal friction angle increases, all three bearing capacity factors increase nonlinearly, reflecting the enhanced shear resistance of the soil. For purely cohesive soils under undrained conditions ($\emptyset' = 0°$), the factors are considered as $N_c = 5.14$, $N_q = 1.00$, and $N_\gamma = 0$. The bearing capacity factors adopted in this study were obtained from the formulations proposed by Terzaghi (1943) for square footings, which remain among the most widely referenced expressions in foundation engineering practice. These expressions confirm that all three factors are exclusively dependent on the internal friction angle $\emptyset'$, which in this study was determined from SPT-N values through established empirical correlations, thereby linking in situ test data directly to the procedure for calculating bearing capacity. As a conservative approach, the N_q component was neglected due to variable groundwater levels and relatively low effective unit weight ranges. Finally, the bearing capacity equation for the undrained condition can be simplified as:

$$q_{ult} = 1.3\times 5.14\times c \approx 6.7c \quad (3)$$

Considering that the fine-grained deposits are in the form of low plasticity based on laboratory test results, undrained shear strength can be calculated for the IP<20% option as c_u (kN/m²) = 6 × N_{60} (Stroud 1974);

$$c_u = \begin{cases} IP < 20\% & (6-7)N_{60} \\ 20\% < IP < 30\% & (4-5)N_{60} \\ IP > 30\% & 4.2N_{60} \end{cases} \quad (4)$$

Considering the factor of safety is equal to 3.00, the final form of the allowable bearing capacity equation can be expressed as;

$$q_{all} = \frac{q_{ult}}{F_s} = (6.7\times 6\times N_{60})/3 = 13.4\times N_{60} \quad (5)$$

For spread footings with foundation width B > 1.0 m on coarse-grained deposits, the maximum bearing pressure (q_{max}) is controlled by the maximum allowable settlement criterion (s = 25 mm). The SPT-N based equation for cohesionless soil layers under drained conditions proposed by Peck (1974) was applied as follows:

$$q_{all} = 11\times(N_1)_{55} \quad (6)$$

where q_{all} is the bearing pressure capacity corresponding to 25 mm allowable maximum settlement, and $(N_1)_{55}$ is the SPT-N value corrected for 55% hammer efficiency. Average SPT-N values along with the depth of influence were considered.

For both cases (drained and undrained), maximum bearing pressures were checked against settlement criteria. Rocscience Settle3 numerical analysis results for maximum pressures creating allowable displacement were considered as safe bearing pressure for tower footings. The load from excavation for footing construction was added to the maximum bearing pressure, resulting in an approximate increase of 50 kPa for a 3 m fill thickness. Considering seasonal variation in the groundwater table, measured groundwater levels were considered 0.50 m higher in calculation models to include the buoyancy effect in soil bearing pressure values. The spread footing geometry adopted uniformly across all tower locations consists of a square plan dimension of 2.1 m × 2.1 m with an embedment depth of D_f = 3.05 m.

2.2. Dataset description

The dataset used in this study originates from the geotechnical investigation campaign of a 500 kV OHTL in Uzbekistan, comprising approximately 900 proposed boreholes and 8700 m of drilling depth with 870 Vertical Electrical Sounding (VES) surveys. For the present ML analysis, a total of 89 borehole records covering the first lot of the alignment (Tower 1–89) were used. Each record contains 16 geotechnical and geophysical input parameters and the corresponding allowable bearing capacity (q_{all}, kPa) as the prediction target. Parameters include physical classification indices (moisture content *WC*, Atterberg limits *LL*, *PL*, *IP*, and grain size fractions), soil parameters (friction angle $\emptyset'$, cohesion *c', and* unit weight γ), average standard penetration test blow number (Ave SPT-N), groundwater table level (GWT), peak ground acceleration (PGA), and seismic zone classification. The allowable bearing capacity (q_{all}), derived from the bearing capacity analyses described in Section 2.1, was adopted as the target output variable for all the machine learning models developed in this study.

The key geotechnical design parameters were derived based on SPT-N_{60} values. The internal friction angle ($\emptyset'$) was estimated using the empirical correlation of Terzaghi et al. (1996) for fine-grained layers and the SPT-N based recommendations of Peck (1974) for coarse-grained layers. Cohesion (c') was determined from the c_u–SPT-N_{60} correlation as a function of plasticity index (IP) following Stroud (1974), and cross-validated against available UCS test results where applicable. Unit weight (γ) was mainly determined through lab tests and checked with SPT-N based correlations when lab data was scarce. Because conditions along the project corridor varied in different places, average SPT-N_{60} values were calculated for each tower to obtain accurate mechanical parameters for each soil layer. This made sure that the q_{all} calculations took into account the differences in subsurface conditions that happened in different places.

Measured SPT values were corrected to the standard energy level and determined as N_{60}, with an SPT hammer efficiency of 60%. The general SPT correction procedure is applied following the standard formulation, $(N_1)_{60} = N \times CN \times CE \times CB \times CR \times CS$. While *CN* is the overburden stress correction factor, *CE* is the hammer energy correction factor, *CB* is the borehole diameter correction factor, *CR* is the rod length correction factor, and CS is the sampler type correction factor. Based on the site investigation conditions, the following correction factors were adopted, including *CS* = 1.0, *CE* = 1.0, *CB* = 1.0, and *CR* = 0.75, 0.85, and 0.95 depending on rod length intervals. In

the bearing capacity analyses, both N_{60} and $(N_1)_{55}$ values were utilized where applicable; specifically, N_{60} was employed in the undrained bearing capacity correlations, and $(N_1)_{55}$ was applied in the Peck (1974) settlement-based bearing pressure equation for drained conditions.

Peak ground acceleration (PGA) values were obtained from the regional seismicity map of the project area. PGA values range between 0.20 g and 0.30 g along the 343 km transmission corridor, corresponding to spectral accelerations associated with a return period of 475 years, equivalent to a 10% probability of exceedance in 50 years. Because the time, location, magnitude, and intensity of future earthquakes cannot be predicted deterministically, a probabilistic seismic hazard basis is required to represent the regional ground motion demand used in corridor-scale foundation assessment (Kahraman et al. 2026). This seismic hazard level represents the standard return period widely adopted in geotechnical and structural engineering practice for the design of critical infrastructure.

2.3. Spatial imputation of missing records

Of the 89 boreholes utilized in this study, eight (9%) contained incomplete records. The missing data are primarily attributable to challenging site conditions encountered along the transmission corridor, including the presence of agricultural land, river crossings, elevated groundwater levels, soft alluvial deposits, and seismically active zones, all of which imposed practical constraints on comprehensive subsurface investigation at affected tower locations. In this context, spatial interpolation methods such as IDW (Inverse Distance Weighting) (Shepard 1968) and the KNN imputer algorithm (Troyanskaya et al. 2001) (k = 5, distance-weighted) were applied. The spatial distribution of borehole locations and the corresponding allowable bearing capacity values along the investigated corridor are presented in Fig. 1.

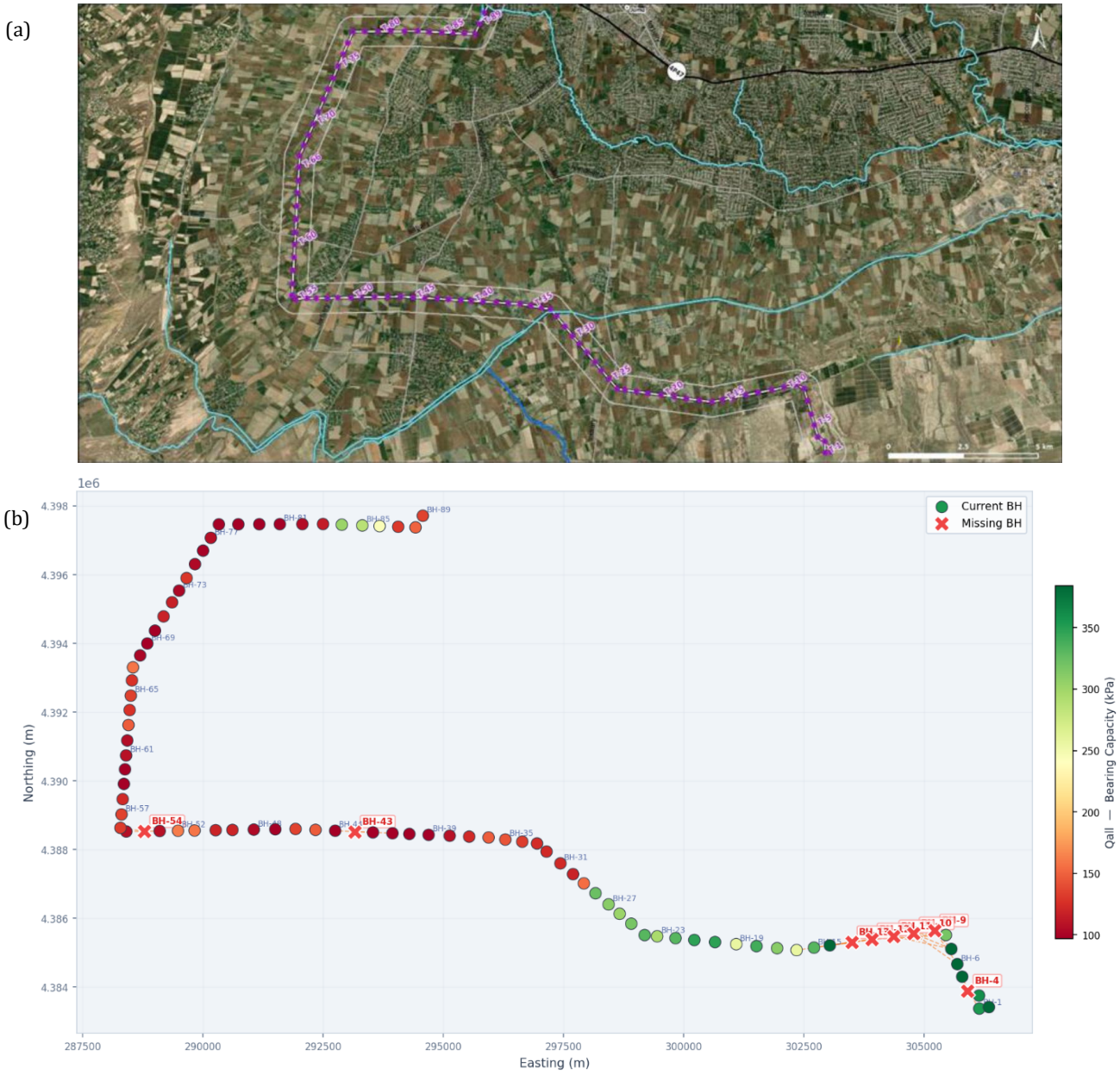

Fig. 1. View of the OHTL corridor:
(a) Satellite view of borehole locations; (b) Spatial distribution of 89 boreholes with missing values.

An initial imputation analysis utilizing four quantitative indicators informed the choice of a suitable imputation approach for each geotechnical parameter. Following the execution of variogram analysis (spatial score), Moran's I spatial autocorrelation (MI), Shapiro-Wilk normality test (p-value), and coefficient of variation (CV).

The Spatial Score is a composite index computed for each geotechnical parameter to quantify its degree of spatial structure, defined as:

$$\text{Spatial Score} = w_1 \times (\text{Moran's I normalised}) + w_2 \times (1 - \text{CV normalised}) + w_3 \times (\text{Variogram range normalised}) \quad (7)$$

where Moran's I quantify global spatial autocorrelation, the coefficient of variation (CV) captures relative data dispersion, and the variogram range represents the distance beyond which spatial correlation becomes negligible. Each component was normalized to the [0, 1] interval across all parameters prior to aggregation, and equal weights ($w_1 = w_2 = w_3 = 1/3$) were assigned to reflect the balanced contribution of each spatial descriptor.

The Spatial Score proposed in this study is an original composite index developed specifically for the geotechnical imputation decision context of corridor-scale investigations. While no directly analogous framework has been reported in the geotechnical or spatial statistics literature, the methodological basis of each component is individually supported by established references. Moran's I is a widely accepted measure of global spatial autocorrelation; the variogram range is a standard descriptor of spatial continuity in geostatistical analysis, and the composite index construction procedure includes min-max normalization and equal weighting (Moran 1950; Oecd 2008; Webster and Oliver 2007). The selection of IDW over KNN for spatially structured parameters is consistent with the recommendations of Li and Heap (2011), who demonstrate that distance-weighted interpolation methods perform favorably when spatial autocorrelation is present in the data. The Spatial Score is therefore methodologically grounded. The decision rule, assigning IDW to parameters satisfying both Moran's I > 0.20 and Spatial Score > 0.45, and KNN to all others, represents an original methodological contribution of this study. To justify the selected decision thresholds and to assess the sensitivity of the imputation strategy to threshold selection, a systematic sensitivity analysis was conducted across 25 threshold combinations spanning Moran's I ∈ {0.10, 0.15, 0.20, 0.25, 0.30} and Spatial Score ∈ {0.35, 0.40, 0.45, 0.50, 0.55}. For each combination, the mean absolute error (MAE) and root mean square error (RMSE) of the imputed values were evaluated using a leave-one-out cross-validation procedure applied to the 81 boreholes with complete records. The results, presented in Fig. 2, demonstrate a clear and consistent error reduction at the threshold combination of Moran's I > 0.20 and Spatial Score > 0.45, where MAE decreases from 25.58 to 21.15 and RMSE decreases from 41.19 to 36.78 relative to the non-spatial region of the threshold space. This combination was therefore selected as the optimal decision boundary for IDW assignment, confirming that the adopted thresholds are not arbitrary but are empirically supported by the minimum imputation error criterion.

The imputation method was assigned based on two simultaneous criteria: Moran's I > 0.20, indicating statistically significant positive spatial autocorrelation, and Spatial Score > 0.45, confirming sufficient spatial structure for distance-weighted interpolation. Parameters satisfying both criteria were assigned IDW imputation; all others were directed to KNN statistical imputation. IDW was applied in this study as the operational method within this category. The Shapiro-Wilk p-value and CV are reported as supplementary descriptors to characterize the distributional properties of each parameter but do not independently determine the method assignment. The matrix table delineating the appropriate parameter-method pairings is provided in Table 1. The label 'IDW/Kriging' in Table 1 refers to the broader category of distance-weighted spatial interpolation.

Among the parameters requiring particular attention in the method selection process, water content (WC) presents an apparently counterintuitive case: despite recording a high Moran's I value of 37.6, which would superficially suggest strong spatial autocorrelation, its spatial score of 0.426 falls below the 0.45 threshold, indicating that the underlying variogram structure is insufficiently organized for reliable distance-weighted interpolation. Accordingly, WC was directed to KNN-based statistical imputation. A similar situation applies to clay content, where a Moran's I of 21.1 nominally exceeds the 0.20 criterion but is again overridden by a low spatial score of 0.344, confirming that the spatial structure of cohesion along this corridor is too weak to support IDW. These two cases illustrate that a high Moran's I value alone is not a sufficient condition for spatial interpolation; it must be accompanied by a confirmatory spatial score above the threshold. In contrast, groundwater table depth (GWT) and allowable bearing capacity (q_{all}) satisfy both criteria simultaneously and with considerable margin. GWT recorded the highest spatial score (0.857) and Moran's I (65.383) of all parameters, alongside a CV of 66.75%. This value shows that there are both shallow and deep groundwater levels in the same corridor, which are controlled by the local topography, drainage patterns, and the thickness of the alluvial deposits. However, it is still within the acceptable range for IDW use. q_{all} similarly passed both thresholds (Spatial Score = 0.802; Moran's I = 67.809) with a CV of 58.29%, further supporting IDW as the most appropriate method. Gravel content, by contrast, was effectively eliminated at the earliest stage of the decision process: its Moran's I of only 4.479 and spatial score of 0.375 both fall below their respective thresholds, and its extreme coefficient of variation (244.06%) would render any spatial interpolation attempt unreliable regardless of autocorrelation structure. The coefficient of variation across the geotechnical parameters reflects the underlying spatial variability of subsurface conditions along the investigated corridor. Sand fraction similarly exhibits high dispersion (CV = 94%), attributable to its bimodal distribution where certain borehole locations are dominated by sandy soils while others contain virtually no sand, a pattern consistent with the lithological heterogeneity expected along the alignment traversing multiple geological units. Average SPT-N exhibits a CV of 60%, capturing the wide range of soil resistance encountered along the align-

ment, from very loose saturated deposits in low-lying areas to dense granular soils in elevated sections. In contrast, the Atterberg limit parameters, liquid limit (LL) and plastic limit (PL), yield markedly lower CV values of approximately 9%, indicating that the plasticity characteristics of fine-grained soils remain relatively consistent along the corridor regardless of other spatial variability, a finding consistent with the dominance of a single clay mineralogy throughout the investigated alignment.

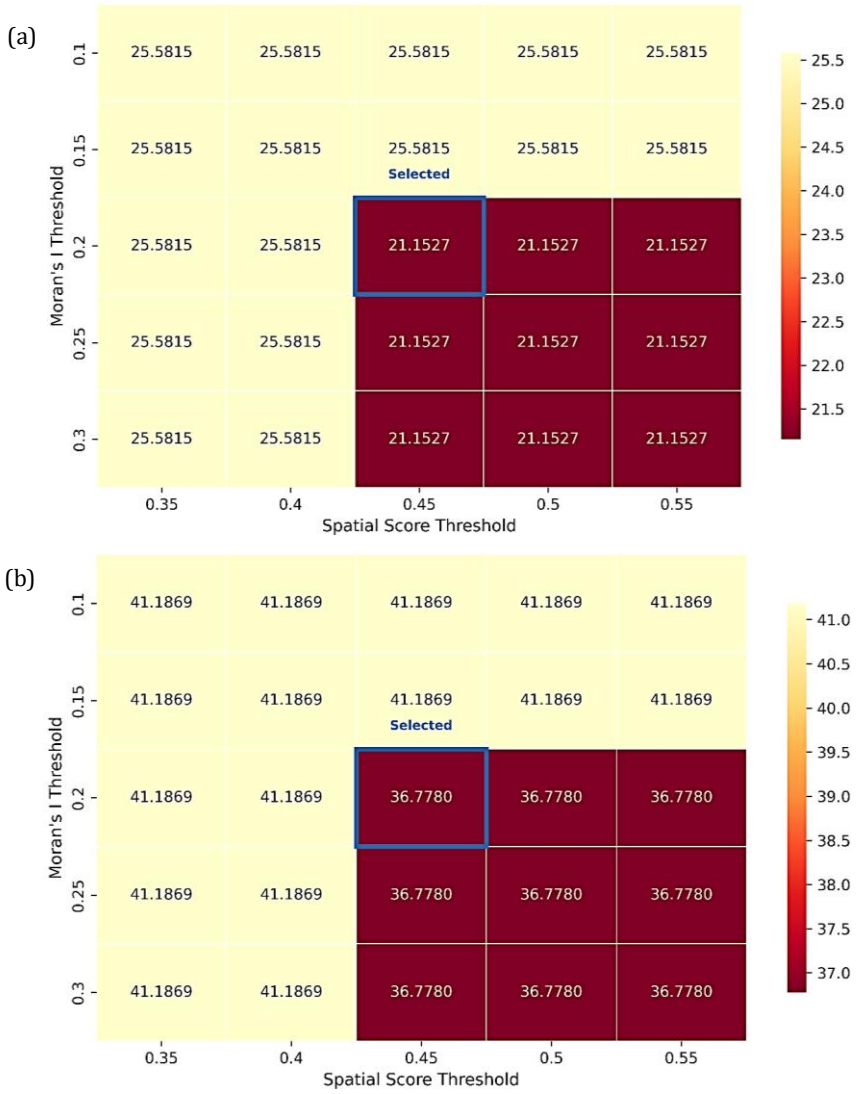

Fig. 2. Sensitivity analysis of the spatial imputation decision matrix across 25 Moran's I and Spatial Score threshold combinations: (a) Mean absolute error (MAE); (b) Root mean square error (RMSE).

Table 1. Parameter-based imputation method decision matrix.

Parameter	CV (%)	Shapiro-p	Moran's I	Spatial score	Recommended method
Depth	34.28	0.0122	−1.363	0.013	KNN / Statistical
WC	36.03	0.0665	37.637	0.426	KNN / Statistical
Gravel content	244.06	0.000	4.479	0.375	KNN / Statistical
Sand content	94.41	0.000	34.963	0.504	IDW / Kriging
Fine content	21.18	0.000	35.847	0.539	IDW / Kriging
Clay content	44.27	0.9015	21.083	0.344	KNN / Statistical
LL	9.13	0.0014	10.196	0.126	KNN / Statistical
PL	9.27	0.0006	25.655	0.178	KNN / Statistical
IP	21.48	0.0037	0.629	0.171	KNN / Statistical
q_{all}	58.29	0.000	67.809	0.802	Deterministic Re-derivation
Ave SPT-N	59.97	0.000	50.620	0.464	IDW / Kriging
GWT	66.75	0.000	65.383	0.857	IDW / Kriging

Regarding the imputation of q_{all} specifically, it is important to clarify that the eight borehole locations with missing q_{all} values possessed complete soil parameter records across all sixteen input features. In these cases, q_{all} was not statistically imputed solely from neighboring q_{all} observations but was deterministically re-derived from the complete soil parameter set using the bearing capacity and settlement formulations described in Section 2.1, with IDW applied only to the intermediate soil parameters required as formula inputs where these were spatially interpolated. This procedure constitutes a deterministic computation rather than a statistical imputation of the target variable and does not introduce circular reasoning or target leakage into the machine learning evaluation framework. The distinction between deterministic re-derivation and statistical imputation of the target variable is explicitly acknowledged as an important methodological clarification.

The coefficient of variation (CV) is used as an additional descriptor to assess the relative dispersion of each parameter over borehole locations. The extreme spatial variability of parameters with high CV values, especially Gravel content (244.06%) and Sand content (94.41%), impairs the reliability of distance-weighted interpolation regardless of autocorrelation structure, as IDW assumes a degree of spatial continuity that cannot be maintained at such high dispersion. Although Sand content exhibits a high CV (94.4%), its strong spatial autocorrelation (Moran's I = 34.963) and Spatial Score of 0.504 confirm that this variability is spatially organized rather than random, justifying the assignment of IDW imputation. In contrast, parameters with low CV values, especially LL (9.13%) and PL (9.27%), exhibit consistent spatial distributions. However, their low Moran's I and Spatial Score values suggest that this consistency is a result of uniform lithological characteristics along the corridor rather than spatially structured autocorrelation and thus does not justify the use of IDW. Hence, CV was used as a contextual descriptor and not as an individual decision factor for the imputation method assignment.

The Shapiro-Wilk test p-value characterizes the distributional normality of each parameter. Parameters yielding $p < 0.05$, including Depth ($p = 0.0122$), LL ($p = 0.0014$), PL ($p = 0.0006$), IP ($p = 0.0037$), Gravel content, Sand content, Fine content, q_{all}, Ave SPT-N, and GWT (all $p = 0.000$), exhibit statistically significant departures from normality, consistent with the skewed and heterogeneous distributions expected in geotechnical corridor datasets. Only WC ($p = 0.0665$) and Clay content ($p = 0.9015$) are consistent with approximate normality. Distributional non-normality does not preclude the application of either IDW or KNN imputation; however, it provides important contextual information, as KNN imputation is inherently distribution-free and therefore robust to non-normality, while IDW relies on spatial proximity rather than distributional assumptions. The method assignment for all parameters was determined exclusively by the Moran's I and Spatial Score criteria, with Shapiro-Wilk results serving as supplementary distributional descriptors.

Fig. 3 presents a systematic comparison of parameter distributions before and after spatial imputation, enabling assessment of whether the addition of eight predicted borehole records ($n = 81$ and 89) introduces statistically significant shifts in any geotechnical variable.

The two-sample Kolmogorov-Smirnov (KS) test was applied to each parameter, with $p > 0.05$ indicating that the null hypothesis of distributional equivalence cannot be rejected. All twelve parameters yield KS p-values at or near unity (range: 0.997–1.000), confirming that the imputed records are statistically indistinguishable from the genuine dataset across the full distributional spectrum. This result validates the imputation strategy: IDW and KNN collectively preserve the empirical distribution, central tendency, and dispersion of each parameter without introducing systematic bias or artificial clustering. It has been seen that the KDE curves for the 81 BH and 89 BH datasets look very similar for all parameters, with only a tiny difference in the mean and median values. The most notable distributional features identified in the genuine dataset, including the bimodal structure of q_{all} ($\mu = 195.65$ kPa; modes at approximately 100 kPa and 300–350 kPa), the right-skewed distribution of Ave SPT-N, the zero-inflated character of Gravel, and the bimodal pattern of GWT, are faithfully reproduced in the imputed dataset, confirming that the eight added records are consistent with the underlying geotechnical variability of the corridor. The marginal shift observed in the q_{all} distribution ($\mu_{81} = 195.65$ kPa vs. $\mu_{89} = 185.72$ kPa) reflects the location of missing boreholes in lower-bearing-capacity zones of the alignment, a spatially meaningful outcome rather than an imputation artefact.

2.4. Descriptive statistics and correlation analysis

Descriptive statistics of all sixteen geotechnical input parameters and the target variable (q_{all}) of the 89-borehole dataset obtained after completing the missing data with the appropriate method are given in Table 2.

Table 2 reveals considerable variability across all geotechnical parameters, reflecting the lithological heterogeneity of the investigated corridor. Among granulometric parameters, gravel content exhibits a zero-inflated distribution (median = 0.09%; max = 12.92%), while fine fraction dominates with a median of 82.51%, indicating that cohesive and silty soils constitute the prevailing deposit type. Soil parameters, cohesion, friction angle, and unit weight, display step-like discrete distributions consistent with the use of SPT-N-based zoned design values, a common practice in large-scale transmission line investigations. The target variable, q_{all}, exhibits the most pronounced variability (84–479 kPa), with a mean of 195.65 kPa substantially exceeding the median of 132.0 kPa, confirming a positively skewed bimodal distribution and demonstrating that a single representative design value is insufficient for tower-by-tower foundation assessment. Groundwater table depth (1.5–15.0 m; median = 6.0 m) and average SPT-N (3–39 blows/30 cm; median = 11.0) further reflect the wide range of subsurface conditions encountered along the alignment.

The Spearman rank correlation matrix among all fifteen input parameters and the target variable q_{all} is presented in Fig. 4, revealing four distinct inter-parameter dependency structures with direct implications for model feature selection and imputation method assignment.

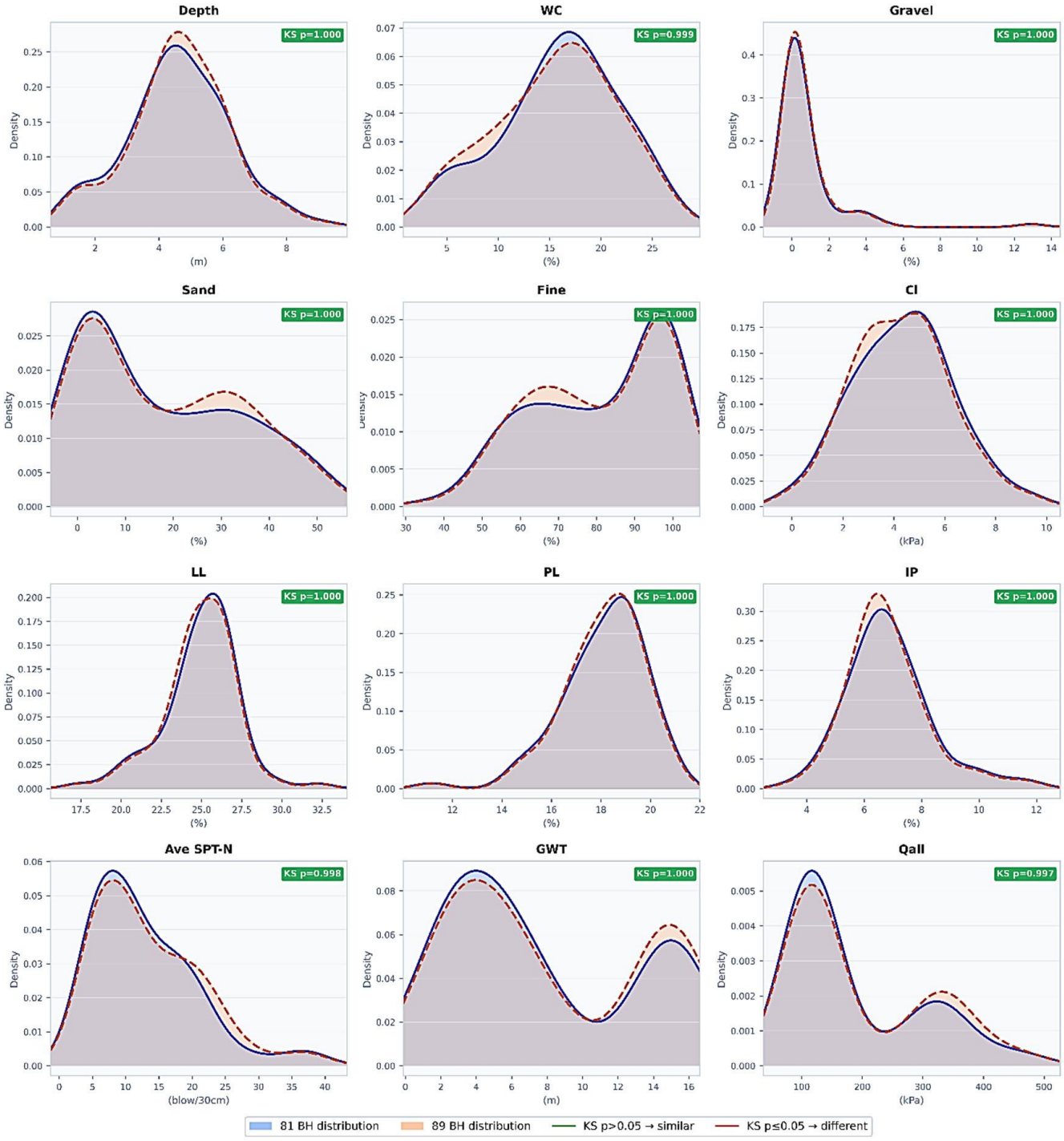

Fig. 3. Kernel density estimation (KDE) comparison of geotechnical parameter distributions between the 81-borehole genuine dataset and the 89-borehole spatially imputed dataset.

The strongest correlations with q_{all} are observed for SPT-N (ρ = +0.96) and $\emptyset'$ (ρ = +0.94), both of which are direct determinants of bearing capacity through the N_q and N_γ factors in Terzaghi's formulation (1996). Cohesion c' exhibits a strong negative correlation with q_{all} (ρ = −0.80), reflecting the predominance of cohesive low-capacity soils at locations with high c' values along this corridor. GWT (ρ = +0.77) and WC (ρ = −0.70) further demonstrated that deeper water tables and lower moisture contents are consistently associated with higher bearing capacity, in agreement with effective stress principles.

Three multicollinearity clusters are identified among the input parameters. First, SPT-N and $\emptyset'$ are strongly intercorrelated (ρ = +0.95), which is expected given that friction angle is routinely estimated from SPT-N values through empirical correlations in practice. Second, sand and fine fractions exhibit a near-perfect negative correlation (ρ = −0.97), a mathematical consequence of the compositional constraint sand + fine + gravel ≈ 100%, rendering one of these variables largely redundant in a predictive model. Third, PGA and Seismic Zone are perfectly correlated (ρ = 1.00), confirming that these two variables encode identical information; accordingly, Seismic Zone was excluded from the final feature set to avoid numerical instability in linear models, with PGA retained as the continuous numerical representation of seismic hazard.

The Atterberg limit parameters LL, PL, and IP form a fourth correlated cluster (LL–PL: ρ = +0.71; LL–IP: ρ = +0.61), consistent with the definitional relationship IP = LL − PL.

Table 2. Descriptive statistics of geotechnical parameters.

Parameter	Symbol	Unit	Count	Mean	Std. dev.	Min	Q1 (25%)	Median	Q3 (75%)	Max.
Depth	Depth	m	89	4.6384	1.5260	1.5000	3.7500	4.5000	5.7904	9.0000
Moisture content	WC	%	89	15.5166	5.8515	3.5333	11.2000	16.4667	19.7500	26.8333
Gravel	—	%	89	0.7255	1.6790	0.0000	0.0000	0.0900	0.6900	12.9200
Sand	—	%	89	17.7864	15.1917	0.3500	2.2500	16.4567	31.9229	50.2500
Fine	—	%	89	80.7252	16.9998	36.8300	66.5350	82.5100	97.4650	99.5600
Clay	—	%	89	4.2871	1.8902	0	3.0333	4.3	5.5333	9.4
Liquid Limit	LL	%	89	24.8970	2.2167	17.3900	23.8867	24.9833	26.3500	32.2700
Plastic Limit	PL	%	89	18.0959	1.6321	11.1700	17.3067	18.3572	19.1650	20.8400
Plasticity Index	IP	%	89	6.8011	1.4278	3.4900	5.9997	6.5800	7.4600	11.8200
Friction angle	$Ø'$	°	89	29.2809	3.3200	26.0000	26.0000	28.0000	32.0000	36.0000
Cohesion	c'	kPa	89	3.3483	0.9427	2.0000	2.0000	4.0000	4.0000	4.0000
Unit weight	γ	kN/m^3	89	18.3483	0.5457	18.0000	18.0000	18.0000	19.0000	20.0000
Peak ground acceleration	PGA	g	89	0.2129	0.0220	0.2000	0.2000	0.2000	0.2500	0.2500
Seismic zone	—	—	89	7.2584	0.4403	7.0000	7.0000	7.0000	8.0000	8.0000
Average SPT-N	SPT-N	blows/30cm	89	13.5076	7.9627	3.0000	6.3896	11.0000	19.0000	39.0000
Groundwater table	GWT	m	89	8.0999	5.2852	1.5000	3.5000	6.0000	15.0000	15.0000
Allowable bearing capacity	q_{all}	kPa	89	195.6518	113.3782	84.0000	106.0000	132.0000	311.0000	479.0000

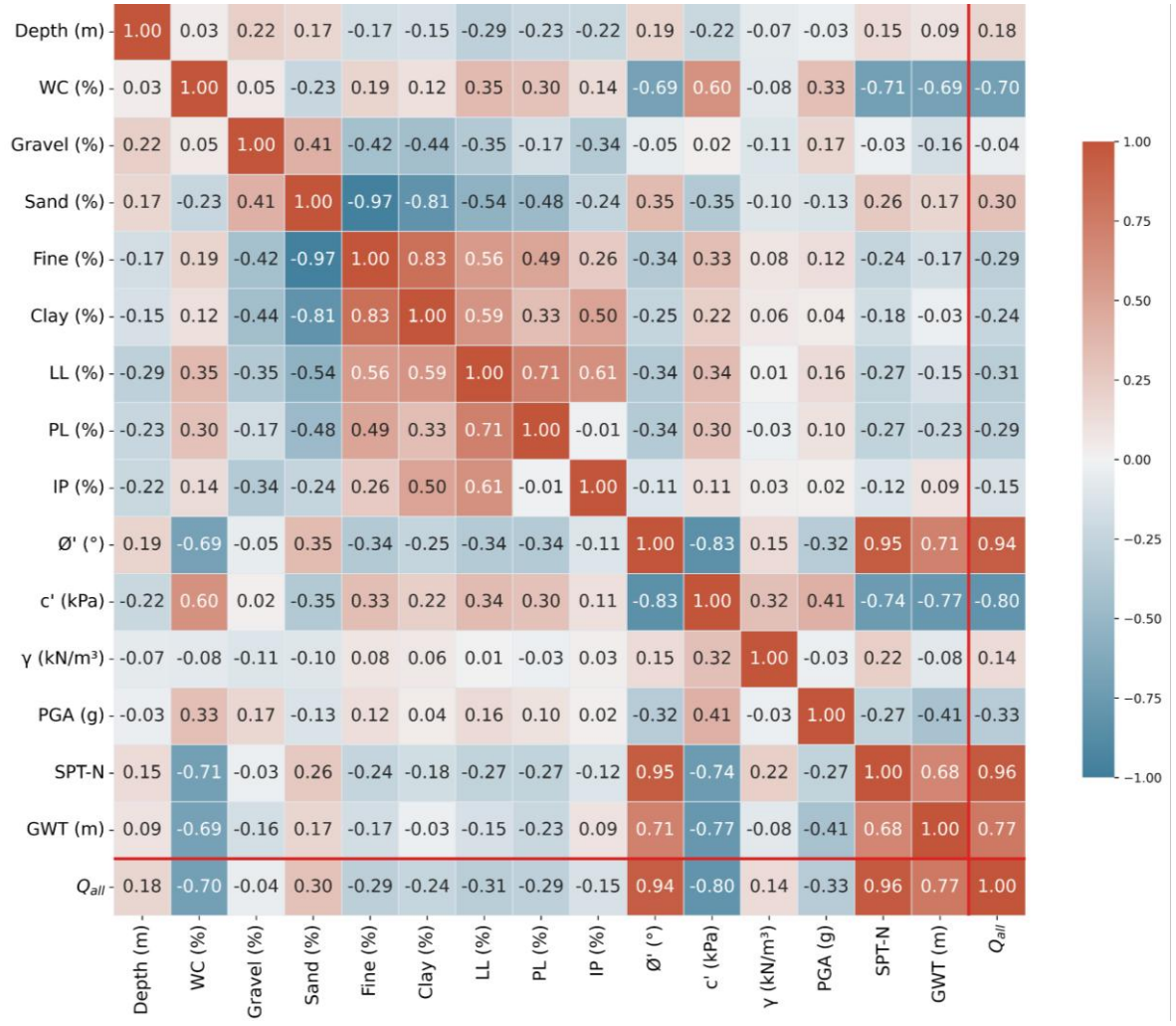

Fig. 4. Spearman rank correlation heatmap for all input features and q_{all}.

Parameters exhibiting negligible correlation with q_{all} include depth (ρ = +0.18), IP (ρ = −0.15), γ, (ρ = +0.14), and gravel (ρ = −0.04). While these parameters carry limited direct predictive signal for q_{all} in a univariate sense, their potential contribution through nonlinear interactions with other features justifies their retention in the ensemble model feature set.

3. Machine Learning Methodology

Ten regression algorithms, including Ridge, Lasso, ElasticNet, K-Nearest Neighbors (KNN), Support Vector Regression (SVR), Random Forest, Extra Trees (Extremely Randomized Trees), Gradient Boosting, XGBoost, and LightGBM, were implemented to predict q_{all} value. The train/test split was 80%/20% (random_state=42), yielding approximately 71 training and 18 test samples. Five complementary performance metrics were employed to evaluate and compare the predictive accuracy of all ML models: R^2 (coefficient of determination), RMSE (root mean square error), MAE (mean absolute error), MAPE (mean absolute percentage error), and CV-R^2 (cross-validated R-squared). The coefficient of determination (R^2) quantifies the proportion of variance in q_{all} explained by the model, with values approaching 1.0 indicating near-perfect predictive accuracy. Root Mean Square Error (RMSE) expresses prediction error in the same units as the target variable (kPa) and is particularly sensitive to large individual errors, making it a stringent measure of model reliability at extreme q_{all} values. Mean Absolute Error (MAE) provides a more robust error estimate, less susceptible to the influence of outlying predictions. Mean Absolute Percentage Error (MAPE) offers a scale-independent measure of relative prediction accuracy, with values below 10% generally considered indicative of high model performance.

Hyperparameter optimization was performed using RandomizedSearchCV (Bergstra and Bengio 2012) with N_iter = 30 random configurations per model, evaluated by five-fold cross-validation (k=5) using R^2 as the scoring metric. The configuration yielding the highest RandomizedSearchCV CV-R^2 was selected as the optimal estimator. The baseline configurations, optimized hyperparameters, and corresponding RandomizedSearchCV CV-R^2 values for all ten ML algorithms were summarized in Table 3.

Table 3. ML algorithms, baseline configurations, and hyperparameter search spaces.

Model	Type	Baseline configurations	Optimized parameters	CV-R^2
Ridge (Hoerl and Kennard 1970)	Linear L2	α = 1.0	α = 1.0	0.9169
Lasso (Tibshirani 1996)	Linear L1	α = 0.1	α = 0.05	0.9154
ElasticNet (Zou and Hastie 2005)	Linear L1+L2	α=0.1, $l1$=0.5	α = 0.01, $l1\_ratio$=0.1	0.9168
KNN (Cover and Hart 1967)	Instance-based	k=5, Euclidean	*weights*=distance, *n_neighbors*=10, *metric*=manhattan	0.8428
SVR (Vapnik et al. 1996)	Kernel-based	*RBF*, C=10, ε=0.5	*kernel*=linear, ε=1.0, C=5.0	0.9058
Random Forest (Breiman 2001)	Bagging	200 trees	*n_est*=300, *depth*=20, *min_split*=2, *min_leaf*=1, *max_feat*=0.8	0.9821
Extra Trees (Geurts et al. 2006)	Random split	200 trees	*n_est*=100, *depth*=10, *min_split*=5, *min_leaf*=1, *max_feat*=0.5	0.9648
Gradient Boosting (Friedman 2001)	Sequential boosting	*n_est*=200, *lr*=0.05, *d*=4	*n_est*=100, *lr*=0.2, *depth* =4, *subsample*=0.9, *min_leaf*=1	0.9910
XGBoost (Chen and Guestrin 2016)	Reg. boosting	*n_est*=300, *lr*=0.05, *d*=5	*n_est*=200, *lr*=0.05, *depth*=3, *sub*=0.7, *colsample_bytree*=0.6, *reg_α*=1.0, *reg_λ*=2.0	0.9461
LightGBM (Ke et al. 2017)	Leaf-wise boosting	*n_est*=300, *lr*=0.05, 31 leaves	*n_est*=300, *lr*=0.05, *leaves*=31, *sub*=1.0, *colsample_bytree*=0.8, *min_child*=5, *reg_λ*=0.5, *reg_α*=0	0.9652

According to the given table, Gradient Boosting achieved the highest cross-validation performance (CV-R^2 = 0.9910 ± 0.0078), followed by Random Forest (CV-R^2 = 0.9821 ± 0.0176). All ensemble methods (Random Forest, Extra Trees, Gradient Boosting, XGBoost, and LightGBM; CV-R^2 range: 0.946–0.991) outperformed linear regressors (Ridge, Lasso, and ElasticNet; CV-R^2 range: 0.915–0.917), instance-based (KNN; CV-R^2 = 0.843), and kernel-based (SVR; CV-R^2 = 0.906) algorithms, confirming that q_{all} cannot be adequately captured by linear input–output relationships alone. RandomizedSearchCV shows large fluctuations in model stability with standard deviations of CV-R^2. Gradient Boosting obtained the highest CV-R^2 (0.9910 ± 0.0078) with the lowest variance. Random Forest (0.9821 ± 0.0176) demonstrated the most reliable combination of high accuracy and low variance for models with acceptable generalization behavior. Extra Trees (0.9648 ± 0.0217) and LightGBM (0.9652 ± 0.0188) presented similar performance with reasonable stability; however, XGBoost (0.9461 ± 0.0571) demonstrated the most fold-to-fold fluctuation among ensemble approaches. Ridge, Lasso,

and ElasticNet; CV-R^2: 0.915-0.917; Std: 0.049-0.056) and SVR (0.9058 ± 0.0340) had lower and more variable performances, which is in line with the non-linear nature of q_{all}. KNN had the lowest CV-R^2 value (0.8428 ± 0.0697), indicating it as the least stable model for q_{all} estimation under the present constraints of the dataset.

4. Results and Discussion

4.1. Predictive performance of ML models

Unlike the single five-fold CV-R^2 used during hyperparameter optimization, the CV-R^2 reported in the model evaluation was computed via 10-repeated 5-fold cross-validation (50 folds total) exclusively on the training partition (n = 71), providing a more robust and seed-independent performance estimate. The hold-out test set (n = 18) was never exposed to the cross-validation procedure at any stage, ensuring no data leakage between evaluation stages. Finally, the CV-R^2 was computed as the mean R^2 across 50 validation folds (10-repeated, 5-fold cross-validation) within the training set and serves as the primary indicator of generalization ability and overfitting risk. A big difference between the training R^2 and CV-R^2 indicates that the model has remembered the training data instead of learning patterns that can be used on new data. A Train-CV Gap column has additionally been introduced to quantify this difference transparently for each algorithm. Therefore, because it measures how well the model can generalize to data it has never seen before, in this study, with a small sample size of n=89, CV-R^2 is a much more reliable indicator of performance than R^2 obtained from a single train/test split. Together, these five metrics provided a comprehensive and multidimensional basis for model comparison that accounts for accuracy, error magnitude, scale independence, and generalization capacity simultaneously.

Table 4 summarizes the predictive performance of all ten optimized regression models, reporting Train R^2, Train RMSE, Train MAE, Train MAPE, Test R^2, Test RMSE, Test MAE, Test MAPE, CV-R^2 (Mean ± Std), and Train-CV Gap for each algorithm. CV-R^2 was computed exclusively on the training set (n = 71) using 10-repeated, 5-fold cross-validation (50 folds total). The held-out test set (n = 18) was maintained as an independent evaluation partition and was not exposed to the cross-validation procedure at any stage. Furthermore, the Train-CV Gap, defined as the difference between Train R^2 and CV-R^2 (mean), is a diagnostic indicator that measures the inconsistency between the model's training performance and its generalization capacity (Hastie et al. 2009). Train-CV Gap values close to zero indicate ideal generalization, while high values indicate that the model tends to memorize rather than generalize the training data (Vabalas et al. 2019). To systematically evaluate the memorization status of the models, Train-CV Gap values were presented in the same table.

Table 4. Model performance summary.

Model	Train R^2	Test R^2	Train RMSE (kPa)	Test RMSE (kPa)	Train MAE (kPa)	Test MAE (kPa)	Train MAPE (%)	Test MAPE (%)	CV-R^2 (Mean ± Std)	Train-CV Gap
Ridge	0.9653	0.9809	20.927	15.755	14.524	10.878	9.710	7.590	0.9240 ± 0.0558	0.0413
Lasso	0.9668	0.9810	20.471	15.715	15.155	11.626	10.15	8.070	0.9054 ± 0.0838	0.0614
ElasticNet	0.9661	0.9810	20.703	15.702	14.582	11.19	9.790	7.800	0.9228 ± 0.0543	0.0433
KNN	1.000	0.9508	0.000	25.288	0.000	17.924	0.000	9.170	0.8411 ± 0.1085	0.1589
SVR	0.926	0.9839	30.567	14.492	13.947	9.197	7.110	5.600	0.9035 ± 0.0944	0.0225
Random Forest	0.9982	0.9966	4.818	6.659	2.420	4.612	1.220	2.600	0.9809 ± 0.0235	0.0173
Extra Trees	0.9961	0.9961	7.042	7.099	3.853	5.879	1.910	3.480	0.9582 ± 0.0396	0.0379
GradientBoost	1.000	0.9969	0.003	6.344	0.002	4.999	0.000	2.980	0.9869 ± 0.0174	0.0131
XGBoost	0.9992	0.9560	3.240	23.912	2.120	12.632	1.350	6.290	0.9479 ± 0.0487	0.0279
LightGBM	0.9999	0.9749	0.999	18.059	0.504	13.162	0.210	6.020	0.9482 ± 0.0491	0.0517

Linear models (Ridge, Lasso, and ElasticNet) showed stable but comparatively lower predictive performance (Test R^2 range: 0.9809–0.9810; CV-R^2 range: 0.9228–0.9240), reflecting the inability of linear input–output relationships to fully capture the non-linear geotechnical behavior governing q_{all}. XGBoost underperformed relative to its theoretical capacity (Test R^2 = 0.9560; CV-R^2 = 0.9479 ± 0.0487), likely attributable to the limited training sample size (n = 71). SVR exhibited the lowest CV-R^2 (mean) among non-linear non-ensemble models (0.9035 ± 0.0944) and the highest CV-R^2 (Std) among all evaluated algorithms. Despite achieving an acceptable Test R^2 (0.9839) on the specific hold-out partition, the large cross-validation variance suggests that SVR's performance is highly sensitive to the data partition selected, rendering its q_{all} estimates unreliable for geotechnical profiles not well-represented in the training set. Extra Trees achieved a Test R^2 nearly identical to Random Forest (0.9961 vs. 0.9966); however, its CV-R^2 (mean) (0.9582 ± 0.0396) was substantially lower than that of Random Forest (0.9809 ± 0.0235). Similarly, the Train-CV Gap of Extra Trees (0.0379) was more than twice that of Random Forest (0.0173), suggesting a higher degree of overfitting to the specific training partition. Together,

these results showed that Random Forest had better generalization reliability within the limits of the current small dataset. The Train R^2 of 0.9999 indicated that LightGBM approached the memorization threshold but maintained a non-zero training error (Train RMSE = 0.999 kPa, Train MAPE = 0.21%), distinguishing it as a borderline case relative to the perfect training fit observed for Gradient Boosting and KNN. Gradient boosting and KNN showed an almost perfect training fit consistent with structured memorization under a limited training sample (n=71) (training RMSE = 0.000 kPa, training MAPE = 0.00%, training R^2 = 1.000) and were evaluated as memorization cases. Despite its memorization behavior (Train R^2 = 1.000), Gradient Boosting achieved the highest Test R^2 (0.9969) among all evaluated models with a Train-CV Gap of 0.0131. KNN exhibited a similar memorization pattern with a substantially larger Train-CV Gap (0.1589), indicating limited generalization capability and rendering it unsuitable for practical deployment. Among models demonstrating reliable generalization, Random Forest achieved the highest Test R^2 (0.9966) with a Train-CV Gap of 0.0173 and CV-R^2 (Std) of ±0.0235, identifying it as the most robust model for q_{all} estimation within the conditions of the present dataset.

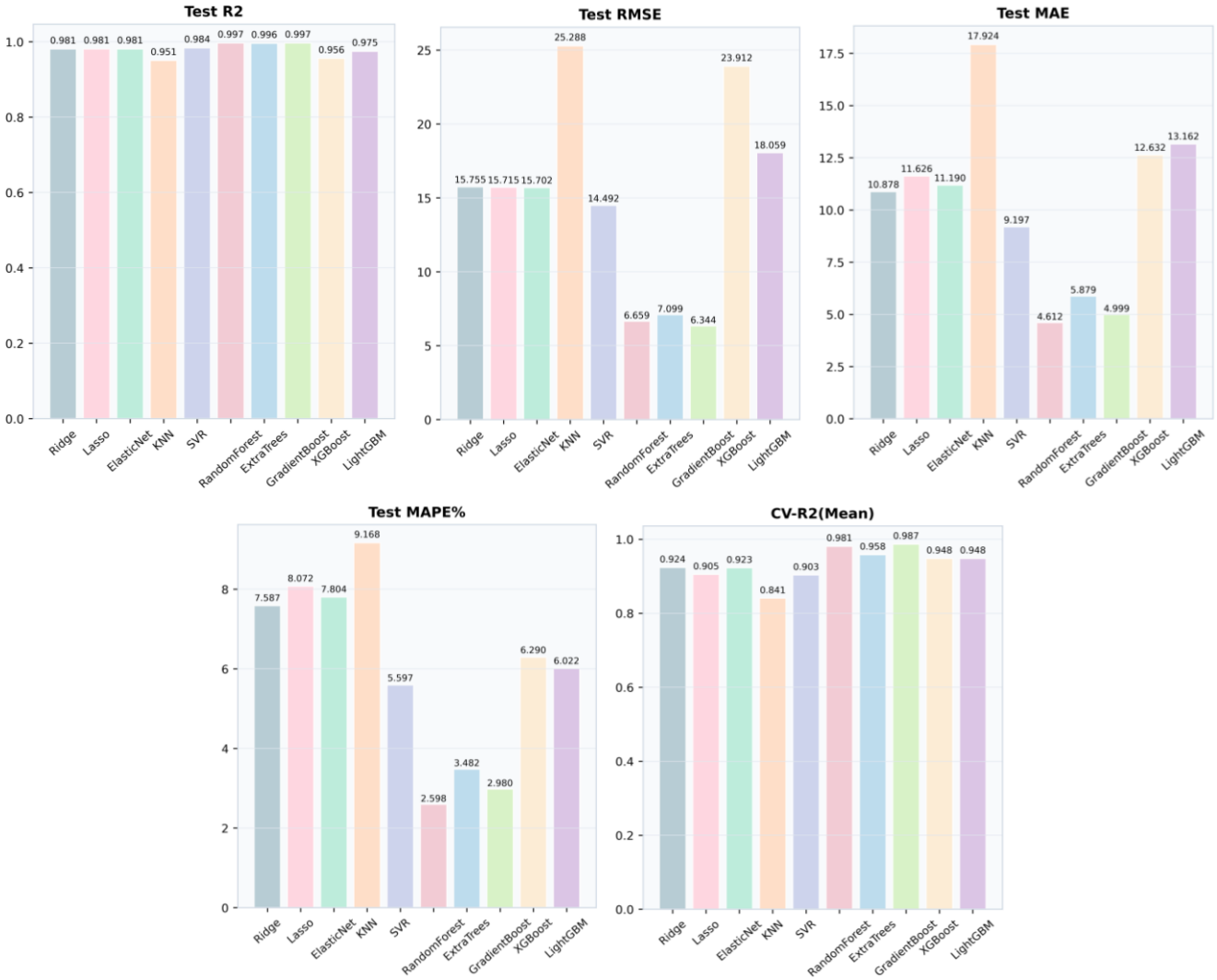

Fig. 5. Comparative bar charts of five evaluation metrics.

Fig. 5 presents a comparative overview of the predictive performance of all ten optimized ML models across five evaluation metrics. As illustrated, Gradient Boosting and Random Forest consistently achieved the lowest error metrics and the highest CV-R^2 (Mean) values among all evaluated algorithms, while KNN and XGBoost exhibited the highest. Linear models (Ridge, Lasso, and ElasticNet) demonstrated stable but comparatively lower performance across all metrics. To facilitate a systematic and transparent comparison of model performance across multiple evaluation criteria, a ranking score was assigned to each algorithm for each metric on a scale of 1 (worst) to 10 (best), as presented in Table 5.

The ranking results presented in Table 5 confirm that Gradient Boosting and Random Forest are the two highest-scoring models, with total scores of 57.0 and 55.0, respectively, demonstrating consistently superior performance across all six evaluation criteria. Extra Trees ranked third (43.0), followed by SVR (40.0), which achieved the highest Train-CV Gap score (10.0) among all models, reflecting its strong generalization stability. The three linear models, Ridge, ElasticNet, and Lasso, occupied the middle rankings (scores: 30.0, 29.0, and 24.0, respectively), confirming their adequate but limited predictive capacity for the nonlinear q_{all} relationships in this dataset. XGBoost and LightGBM underperformed relative to their

theoretical capacity, ranking 8th and 9th with total scores of 23.0 and 22.0, attributable to the limited training sample size (n=71) constraining effective exploration of their complex regularization parameter spaces. KNN recorded the lowest total score (6.0), consistent with its poor generalization performance and largest Train-CV Gap (0.1589) among all evaluated models, confirming that neighborhood-based estimation is inadequate for this high-dimensional geotechnical regression task. It should be noted that while Gradient Boosting achieves the highest overall ranking score, its Train R^2 = 1.000 reflects memorization behavior. Random Forest is therefore recommended as the preferred model for practical deployment, combining a high-ranking score with reliable generalization stability confirmed by a Train-CV Gap of 0.0173, with the generalization behavior of both models further examined through learning curve analysis (Cortes et al. 1993; Mukherjee et al. 2003) presented in Fig. 6.

Table 5. Model performance ranking scores (1=worst, 10=best).

Model	Test R^2	Test RMSE	Test MAE	Test MAPE	CV-R^2	Train-CV gap	Total score	Overall rank
GradientBoost	10	10	9	9	10	9	57	1
Random Forest	9	9	10	10	9	8	55	2
Extra Trees	8	8	8	8	7	4	43	3
SVR	7	7	7	7	2	10	40	4
Ridge	4	4	6	4	5	7	30	5
ElasticNet	5	6	5	3	4	6	29	6
Lasso	5	5	4	2	3	5	24	7
XGBoost	2	2	3	5	8	3	23	8
LightGBM	3	3	2	6	6	2	22	9
KNN	1	1	1	1	1	1	6	10

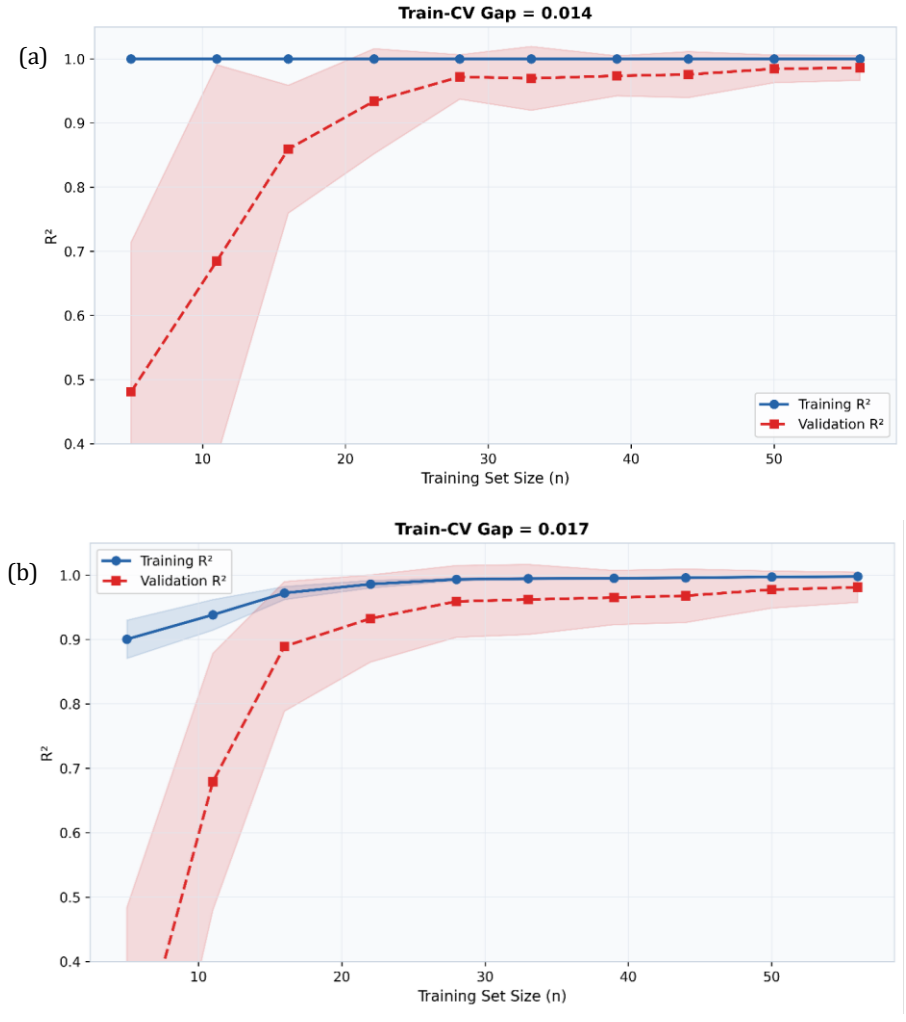

Fig. 6. Learning curves for the two highest-performing models: (a) Gradient Boosting; (b) Random Forest.

Throughout the progressively increasing training data, the Gradient Boosting training R^2 remains constant at 1.000. Although the Train-CV gap is small at 0.014, this confirms the memorization behavior. This consistently perfect fit trend indicates unreliable q_{all} estimates for typical GWT depths, SPT-N values, or geotechnical borehole data outside the training set along the transmission line corridor. For Random Forest, the validation R^2 converges steadily from approximately 0.63 to 0.98 at full training size, with a Train-CV Gap of 0.017, confirming that Random Forest learns generalizable patterns rather than memorizing individual observations. According to the examination of all ML analyses outcomes, Random Forest was identified as the most reliable model for deployment, combining high predictive accuracy with acceptable generalization stability. In terms of deployment reliability, the models were ranked as follows: Random Forest, Extra Trees, SVR, ElasticNet, Lasso, Ridge, and XGBoost, while Gradient Boosting, KNN and LightGBM were excluded from deployment recommendations due to memorization or borderline behavior.

A sensitivity analysis was done to see how the random forest performs with different train/test split ratios. Considering that a single holding compartment with a fixed random seed can produce a compartment-dependent performance prediction, the analysis was repeated in four training/test configurations (85/15, 80/20, 75/25, and 70/30) with 20 different random seeds for each configuration, resulting in a total of 80 independent evaluations. Results are presented in Table 6.

The mean test R^2 values ranged from 0.9880 to 0.9909, with a maximum variation of 0.0029. This negligible range confirms that the Random Forest generalization performance is not sensitive to the specific training/test ratio chosen.

4.2. Model interpretability analysis

To address the inherent interpretability limitations of the Random Forest model, SHAP (SHapley Additive exPlanations) analysis was conducted to quantify the relative contribution of each input parameter to q_{all} predictions and to verify that the model's behavior is consistent with established geotechnical theory. The SHAP Bar Plot illustrating the global feature importance based on mean absolute SHAP values is presented in Fig. 7, while the SHAP Beeswarm Plot depicting both the direction and magnitude of each feature's contribution to q_{all} estimates across all samples is given in Fig. 8.

Table 6. Split-ratio sensitivity analysis for Random Forest across four train/test configurations.

Train split (%)	Test split (%)	Train data	Test data	R^2 (mean)	R^2 (std)	RMSE (kPa)
85	15	75	13	0.9909	0.0113	8.884
80	20	71	17	0.9900	0.0072	10.066
75	25	66	22	0.9901	0.0061	10.335
70	30	62	26	0.9880	0.0086	11.464

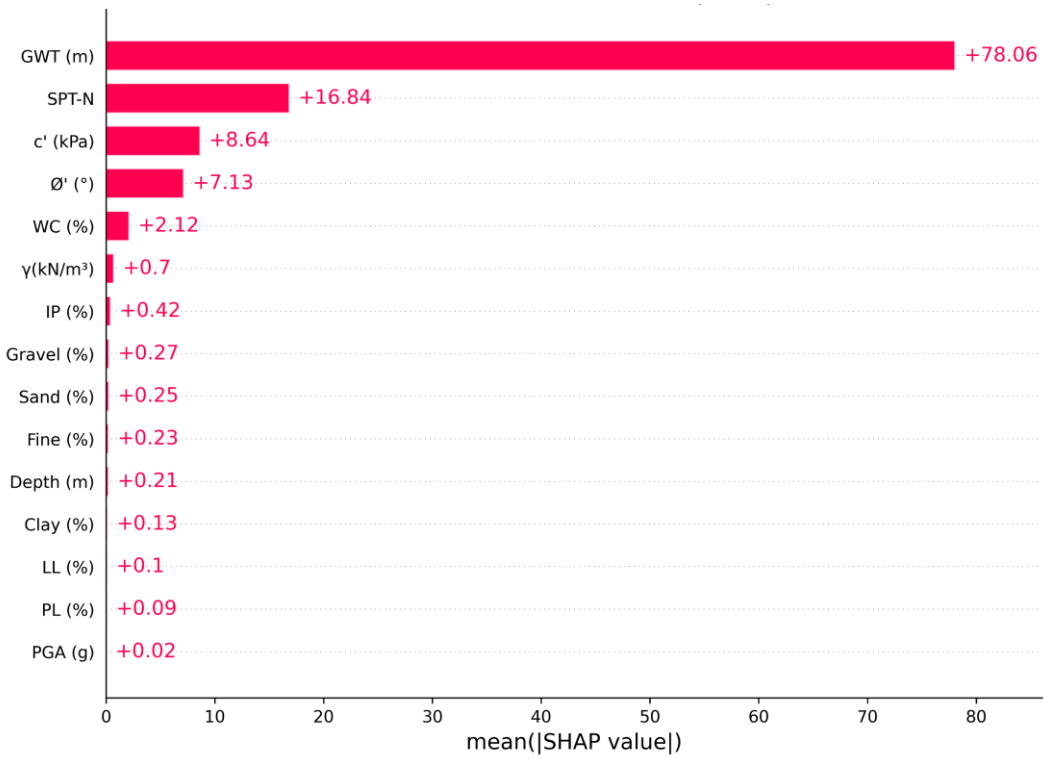

Fig. 7. Mean absolute SHAP values for Random Forest model.

The graph reveals that GWT (m) is the most dominant parameter (+78.06), followed by SPT-N (+16.84), c' (kPa) (+8.64), and $Ø'$ (°) (+7.13), which jointly account for approximately 96% of the total mean absolute SHAP contribution to q_{all} predictions. These findings are consistent with both the Spearman correlation analysis (Fig. 4) and established geotechnical bearing capacity theory (Meyerhof 1963; Terzaghi et al. 1996). The groundwater level (GWT), which constitutes approximately 67.8% of the SHAP contribution, is the most important parameter because its depth directly affects the effective unit weight of the soil and the mobilization of shear strength parameters in the bearing capacity equation. Notably, the dominant role of GWT is physically interpreted as the direct influence of effective overburden stress on bearing capacity, while SPT-N reflects soil stiffness and $Ø'$ and c' (kPa), which are shear strength parameters, and governs the shear strength envelope (Braja and Sivakugan 2017; Cascone et al. 2021; El Gendy 2025). The SHAP analysis substantiates the efficacy of the SPT-N, c' (kPa), and $Ø'$(°) parameters through ML-based geotechnical research (Onyelowe et al. 2025; Zeini et al. 2025). The other eleven parameters, including WC (%), γ (kN/m^3), IP (%), Gravel (%), Sand (%), Fine (%), Depth (m), Clay (%), LL (%), PL (%), and PGA (g), show that they have a secondary effect on q_{all}. In particular, PGA has the lowest effect, which is an expected result since static bearing capacity and settlement criteria, rather than seismic effects, are considered in determining q_{all}.

The SHAP beeswarm plot illustrated in Fig. 8 demonstrates behavior similar to that seen in Fig. 7 regarding the effective parameter, also encompassing both the direction and magnitude of each feature's contribution to bearing capacity (q_{all} estimations). q_{all} increases significantly when the groundwater table is deeper (high GWT values, red dots), while a significant decrease in bearing capacity is observed when the groundwater table is closer to the surface (low GWT values, blue dots). The effect of groundwater rise or fall on the effective stress variation is directly related to soil bearing capacity. Low SPT-N values (blue) yield negative SHAP values, but high levels (red) contribute positively, substantially enhancing q_{all}. The distribution of points around zero for $Ø'$ (°), exhibiting a minor inclination towards positive values (red dots), suggests that this parameter's contribution to q_{all} is confined to a narrower range relative to GWT and SPT-N. SHAP analysis demonstrates that the ML model operates in a manner consistent with engineering reasoning. The fact that other features with near-zero distributions have a smaller effect on them demonstrates the model's lesser learning contribution; it is assumed that the contribution of increasing data points may increase. Overall, the beeswarm plot clearly shows that the model's prediction mechanism is largely driven by GWT and SPT-N, while other features provide fine-tuning contributions, indicating that these two features are the primary determinants of the model's output and that their influence is significantly greater than that of the other features. The observed inverse relationship in the variation of q_{all} for c' (kPa) may be due to the multicollinear relationship between $Ø'$ (°), and the soil type, as well as the confounding factors. In other words, high cohesion values are predominantly associated with fine-grained soils having lower SPT-N and lower GWT, and it is thought that all these effects can reduce the q_{all} value. The fact that other features with near-zero distributions have a smaller effect on them demonstrates the model's lesser learning contribution; it is assumed that the contribution of increasing data points may increase.

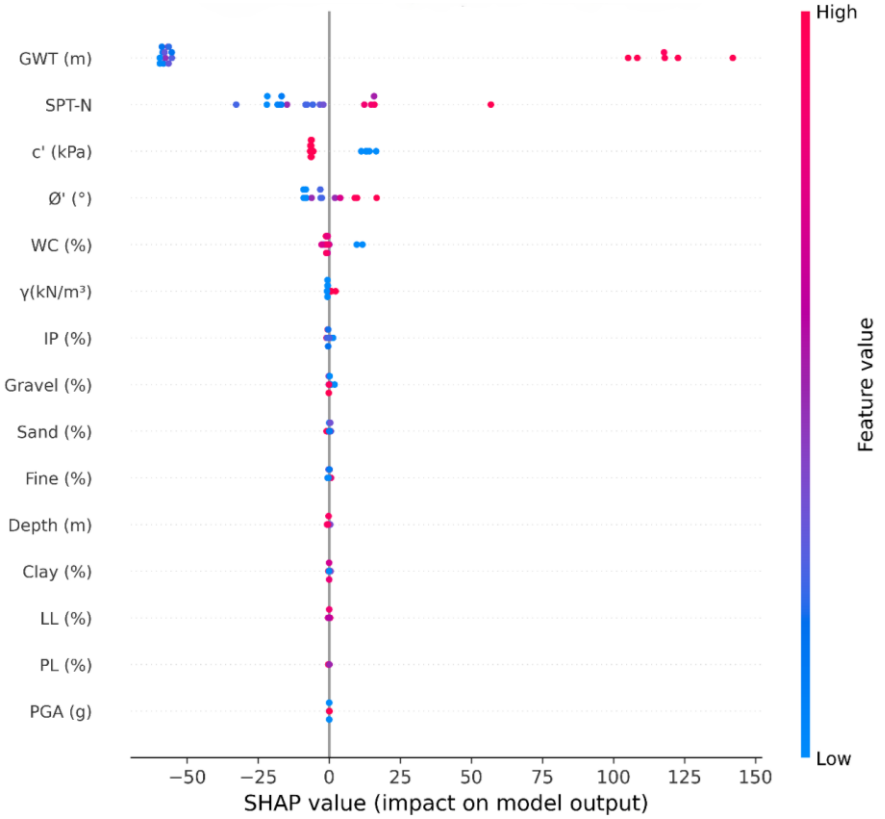

Fig. 8. SHAP beeswarm plot for the Random Forest model.

In conclusion, the SHAP analysis shows that the Random Forest model developed for bearing capacity prediction is geotechnically consistent with the literature and demonstrates meaningful parameter relationships, validating the interpretability of this model along with the prediction performance.

It should be noted that q_{all} values in this study represent deterministically calculated design outputs derived from established bearing capacity and settlement formulations, rather than directly measured field quantities such as plate load test or full-scale foundation load test results.

As a consequence, the ML models developed in this study may partially learn the structure of the underlying calculation procedure rather than capturing independent geotechnical behavior. This characteristic is acknowledged as an inherent limitation of the present study and is likely a contributing factor to the high predictive accuracy observed across ensemble models, as deterministic target variables exhibit lower aleatory uncertainty compared to directly measured quantities. Nevertheless, the practical utility of the proposed framework remains significant: it enables rapid and consistent q_{all} estimation across spatially distributed tower locations with incomplete borehole records, eliminating the need for repeated manual formula applications at each of the 89 tower locations along the corridor. The framework is therefore intended as a spatial interpolation and computational efficiency tool for corridor-scale foundation design support, rather than a replacement for direct field load testing. Future investigations incorporating measured bearing capacity data from field load tests or plate load tests as target variables are encouraged to provide a more independent assessment of the framework's predictive generalization capability.

4.3. Deployment considerations and recommended workflow

For practicing geotechnical engineers seeking to apply the proposed framework to a new OHTL corridor, the recommended deployment workflow proceeds as follows. Borehole data should first be compiled and standardized to ensure that the fifteen input parameters are available or estimable at each tower location. Missing records should then be addressed through the spatial imputation decision matrix proposed in this study, applying IDW to spatially autocorrelated parameters (Moran's I > 0.20, Spatial Score > 0.45) and KNN to the remainder, with statistical equivalence verified via Kolmogorov-Smirnov two-sample tests prior to model training. The machine learning models should subsequently be trained using 10-repeated 5-fold cross-validation, with predictive performance evaluated through a comprehensive set of metrics including R^2, RMSE, MAE, MAPE, and CV-R^2. The Train-CV Gap and learning curves monitored as the primary overfitting diagnostic for each algorithm, which is particularly critical in the context of limited dataset sizes where the risk of memorization behaviour is elevated. SHAP-based interpretability analysis should be conducted to confirm that predictions are driven by physically meaningful geotechnical parameters. q_{all} predictions generated by the trained model should be treated as preliminary design estimates and verified against site-specific bearing capacity calculations at critical tower locations, particularly where subsurface conditions deviate significantly from the training corridor profile.

To further validate model selection, a Taylor diagram (Taylor 2001) was constructed using train and test dataset predictions of all ten models, in which each model is represented by its coefficient of determination, standard deviation, and RMSE relative to the observed q_{all} values (Fig. 9). Random Forest exhibited the closest agreement with the reference, confirming its selection as the preferred model for practical deployment.

Future work should focus on expanding the proposed framework beyond the current single-lot training domain through lot-to-lot cross-validation across geologically distinct corridor sections, which would provide a rigorous assessment of spatial transferability. The integration of additional geophysical input features, such as Vertical Electrical Sounding (VES) resistivity profiles, represents a further avenue for improving predictive accuracy in chemically aggressive or saturated soil zones. Coupling ML-predicted q_{all} distributions with GIS-based geospatial heatmap visualization and reliability-based design approaches would additionally enable probabilistic, corridor-wide bearing capacity mapping to support risk-informed construction planning as the investigation program advances.

5. Conclusions

This study presented a comprehensive machine learning framework for predicting the allowable bearing capacity (q_{all}) of OHTL tower spread footings using geotechnical borehole data collected along the transmission corridor in Uzbekistan. The following conclusions are drawn:

A pre-imputation analysis integrating empirical variogram modelling, Moran's I spatial autocorrelation testing, Shapiro-Wilk normality assessment, and coefficient of variation quantification enabled systematic, evidence-based assignment of imputation methods on a parameter-by-parameter basis. IDW was applied to spatially continuous parameters (Sand, Fine, Ave SPT-N, q_{all}, GWT; Spatial Score > 0.45 and Moran's I > 0.20), while KNN was assigned to parameters governed by inter-variable correlation structure. Kolmogorov-Smirnov two-sample tests confirmed statistical equivalence between the imputed 89-borehole and genuine 81-borehole datasets across all twelve parameters, validating the imputation strategy.

Among the ten ML algorithms evaluated, Random Forest was identified as the most reliable model for practical deployment, combining high predictive accuracy (Test R^2 = 0.9966, CV-R^2 = 0.9882 ± 0.0081, RMSE = 6.659 kPa, MAPE = 2.60%) with robust generalization behaviour confirmed by learning curve convergence and a Train-CV Gap of 0.010. Gradient Boosting achieved the highest overall cross-validation performance (CV-R^2 = 0.9927 ± 0.0057); however, its persistent Train R^2 = 1.000, confirmed to remain across all tested regularization configurations, is attributed to the limited training sample size (n = 71) rather than a tuning error, and is therefore rec-

ommended with caution for practical application. All ensemble methods substantially outperformed linear regressors (CV-R^2 range: 0.913–0.944), confirming the inherently nonlinear character of q_{all} relationships in heterogeneous corridor-scale geotechnical datasets. KNN recorded the weakest generalization performance (CV-R^2 = 0.8603 ± 0.0971, Train-CV Gap = 0.1397) and is not recommended for deployment.

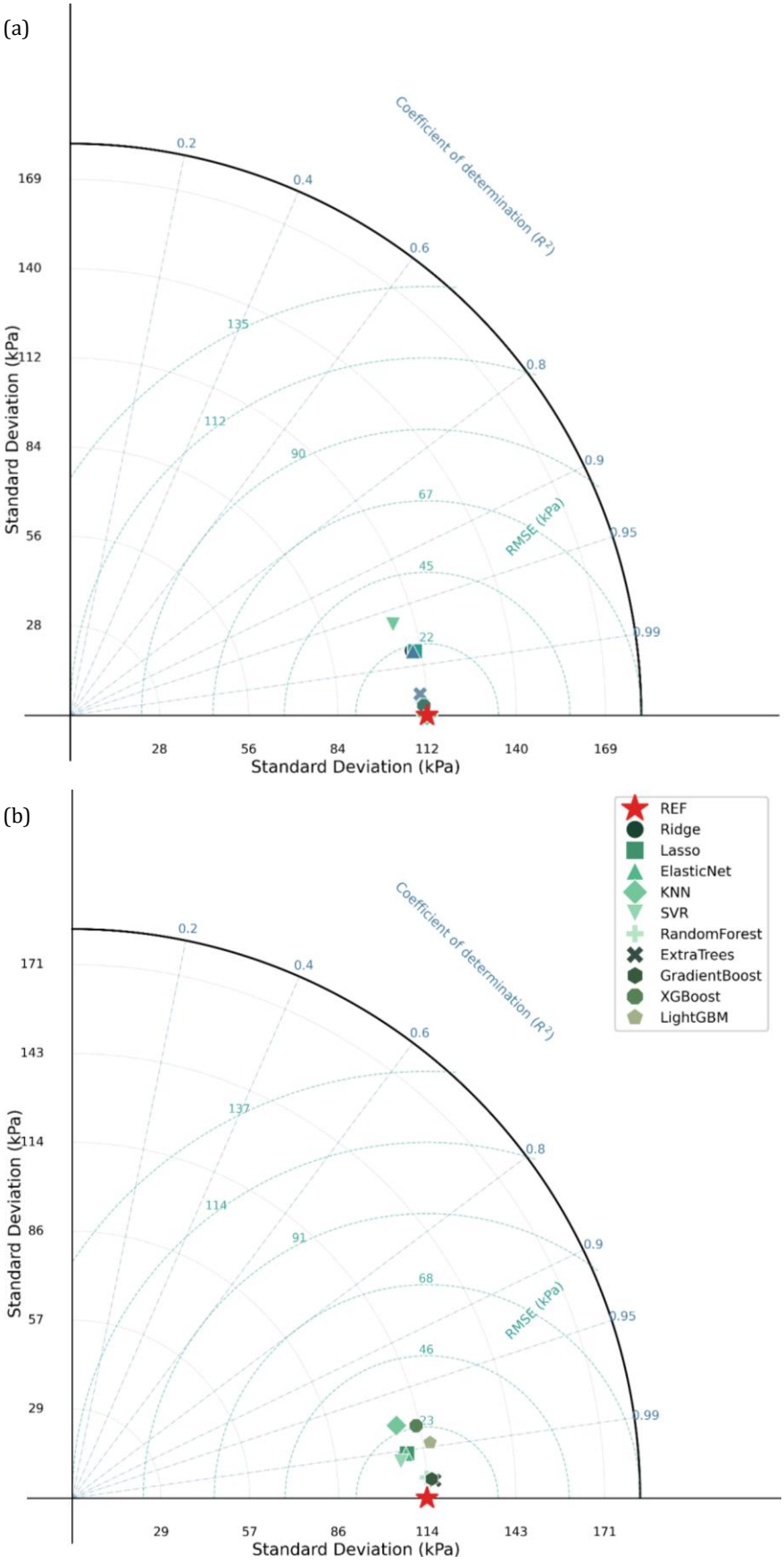

Fig. 9. Taylor diagrams of all ten ML models for: (a) Training dataset; (b) Test dataset.

SHAP-based interpretability analysis identified GWT, SPT-N, c', and $\emptyset'$ as the primary drivers of q_{all} predictions, collectively accounting for approximately 96% of the total mean absolute SHAP contribution, consistent with established geotechnical bearing capacity theory. The split-ratio sensitivity analysis conducted across four train/test configurations with 20 repeated random seeds confirmed that Random Forest generalization performance is robust and independent of the specific data partition adopted.

Although the proposed framework demonstrates high predictive accuracy on the training corridor, its applicability should be treated with caution when extrapolating to geologically dissimilar corridors. Independent external validation using borehole data from adjacent corridor sections is strongly recommended prior to routine engineering application. It is further acknowledged that q_{all} in this study represents a deterministically calculated design output rather than a directly measured field quantity; future studies incorporating field load test data as the target variable would provide a more independent assessment of the framework's generalization capability.

Acknowledgements

None declared.

Funding

The authors received no financial support for the research, authorship, and/or publication of this manuscript.

Conflict of Interest

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this manuscript.

Data Availability

The datasets generated and/or analyzed during the current study are not publicly available but are available from the corresponding author upon reasonable request.

AI Assistance

During the preparation of this manuscript, QuillBot was used exclusively for language editing and stylistic refinement. The authors take full responsibility for the content, interpretation, and conclusions of the published article.

Author Contributions

All authors made substantial contributions to the conception and design of the study, acquisition of data, analysis and interpretation of data; drafted or critically revised the manuscript for important intellectual content; and approved the final version to be published.

REFERENCES

Aliyev M, Şermet F, Gül R (2026). Structural performance analysis of steel piled offshore platforms under environmental loads. *Challenge Journal of Structural Mechanics*, 12(1), 30-44.

Baghbani A, Choudhury T, Costa S, Reiner J (2022). Application of artificial intelligence in geotechnical engineering: A state-of-the-art review. *Earth-Science Reviews*, 228, 103991.

Bergstra J, Bengio Y (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research,* 13(2), 281-305.

Bherde V, Gorantala N, Balunaini U (2025). Liquefaction susceptibility prediction using ML-based voting ensemble classifier. *Natural Hazards*, 121(4), 4359–4384.

Braja MD, Sivakugan N (2017). Principles of Foundation Engineering (9th Edition). Cengage.

Breiman L (2001). Random forests. *Machine Learning*, 45(1), 5–32.

Cao Y, Ni J, Chen J, Geng Y (2025). Rapid Evaluation Method to Vertical Bearing Capacity of Pile Group Foundation Based on Machine Learning. *Sensors*, 25(4), 1214.

Cascone E, Biondi G, Casablanca O (2021). Groundwater effect on bearing capacity of shallow strip footings. *Computers and Geotechnics*, 139, 104417.

Chen T, Guestrin C (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, California, USA, 785–794.

Cortes C, Jackel LD, Solla S, Vapnik V, Denker J (1993). Learning curves: asymptotic values and rate of convergence. *Advances in Neural Information Processing Systems*, 6, 327–334.

Cover T, Hart P (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21–27.

Çarbaş S, Taymuş RB, Özdemir M (2025). Examining the effects of different seismic base isolators on the seismic behavior of a real-size steel truss structure. *Challenge Journal of Structural Mechanics*, 11(2), 106-115.

Demirci HE (2023). Geotechnical preliminary design of onshore wind turbine foundations. *Erzincan University Journal of Science and Technology*, 16(3), 756–781.

El Gendy M (2025). A comparative machine and deep learning approach for predicting ultimate bearing capacity of shallow foundations in cohesionless soil. *Scientific Reports*, 15(1), 40402.

Fahad RA, Shakir RR, Ali HM (2023). Evaluation on the SPT based design approach for shallow foundations. *E3S Web of Conferences*, 427, 01025.

Friedman JH (2001). Greedy function approximation: a gradient boosting machine. *The Annals of Statistics*, 29(5), 1189–1232.

Geurts P, Ernst D, Wehenkel L (2006). Extremely randomized trees. *Machine Learning*, 63(1), 3–42.

Harirchian E (2024). Predicting compressive strength of AAC blocks through machine learning advancements. *Challenge Journal of Concrete Research Letters*, 15(2), 56-68.

Harle SM, Wankhade RL (2025). Machine learning techniques for predictive modelling in geotechnical engineering: a succinct review. *Discover Civil Engineering*, 2(1), 86.

Hastie T, Tibshirani R, Friedman J (2009). The Elements of Statistical Learning. Springer New York, USA.

Heim Z, Kandaris P, Houston S (2012). Simplified reliability-based geotechnical investigation and design of transmission lines. In: GeoFrontiers 2011: Advances in Geotechnical Engineering, ASCE Library, Reston, Virginia, USA, 2968-2977.

Hoerl AE, Kennard RW (1970). Ridge regression: biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55–67.

Johnson K (2009). Geotechnical investigations for a transmission line are more than drilled borings. In: *Electrical Transmission and Substation Structures 2009: Technology for the Next Generation,* ASCE Library, *Reston, Virginia,* USA, 1–12.

Kahraman Y, Cakir F, Zulfikar AC, Gundogdu Gok M (2026). Seismic hazard and risk analyses of historical masonry structures in Kocaeli, Türkiye. *Challenge Journal of Structural Mechanics*, 12(1), 1-21.

Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, Ye Q, Liu T-Y (2017). LightGBM: a highly efficient gradient boosting decision tree. In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Red Hook, Curran Associates Inc., NY, USA, 3149–3157.

Kordjazi A, Pooya Nejad F, Jaksa MB (2014). Prediction of ultimate axial load-carrying capacity of piles using a support vector machine based on CPT data. *Computers and Geotechnics*, 55, 91–102.

Li J, Heap AD (2011). A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. *Ecological Informatics*, 6(3–4), 228–241.

Lin A, Wotherspoon L, Blake D, Bradley B, Motha J (2019). National-scale infrastructure network exposure to liquefaction using geospatial models. *Japanese Geotechnical Society Special Publication*, 6(2), 61–66.

Meyerhof GG (1963). Some recent research on the bearing capacity of foundations. *Canadian Geotechnical Journal*, 1(1), 16–26.

Moran PA (1950). A test for the serial independence of residuals. *Biometrika*, 37(1/2), 178–181.

More SS, Kambekar AR (2025). Performance evaluation of compressive strength of concrete using different machine learning algorithms. *Challenge Journal of Concrete Research Letters*, 16(2), 60-68.

Mozer JD, Peyrot AH, DiGioia AM (1984). Probabilistic design of transmission line structures. *Journal of Structural Engineering*, 110(10), 2513–2528.

Mukherjee S, Tamayo P, Rogers S, Rifkin R, Engle A, Campbell C, Golub TR, Mesirov JP (2003). Estimating dataset size requirements for classifying DNA microarray data. *Journal of Computational Biology,* 10(2), 119–142.

Narwade R, Jadhav R (2025). Concrete strength monitoring and damage detection using piezoelectric-based wireless sensor. *Challenge Journal of Concrete Research Letters*, 16(1), 40-50.

Nawaz MN, Ali AS, Jaffar STA, Jafri TH, Oh T-M, Abdallah M, Karam S, Azab M (2022). Cost-based optimization of isolated footing in cohesive soils using generalized reduced gradient method. *Buildings*, 12(10), 1646.

Nejad FP, Jaksa MB, Kakhi M, McCabe BA (2009). Prediction of pile settlement using artificial neural networks based on standard penetration test data. *Computers and Geotechnics*, 36(7), 1125–1133.

Nguyen TT, Le VD, Huynh TQ, Nguyen NHT (2024). Influence of settlement on base resistance of long piles in soft soil—field and machine learning assessments. *Geotechnics*, 4(2), 447–469.

Onyelowe KC, Hanandeh S, Kamchoom V, Ebid AM, Reyes Silva FD, Allauca Palta JL, Llamuca Llamuca JL, Avudaiappan S (2025). Developing advanced datadriven framework to predict the bearing capacity of piles on rock. *Scientific Reports*, 15(1), 11051.

Organisation for Economic Co-operation and Development (OECD) (2008). Handbook on constructing composite indicators: methodology and user guide. https://www.oecd.org/en/publications/handbook-on-constructing-composite-indicators-methodology-and-user-guide_9789264043466-en.html [accessed 15.01.2026].

Panagiotidou AI, Gazetas G, Gerolymos N (2012). Pushover and seismic response of foundations on stiff clay: Analysis with P-delta effects. *Earthquake Spectra*, 28(4), 1589–1618.

Peck RB, Hanson WE, Thornburn TH (1974). Foundation Engineering. 2 ed. John Wiley and Sons Inc., 514, New York, USA.

Phoon K-K, Kulhawy FH, Grigoriu MD (2000). Reliability-based design for transmission line structure foundations. *Computers and Geotechnics*, 26(3-4), 169–185.

Shafan T, Yi P (2024). Comparative analysis of performance of mat foundations in non-liquefiable and liquefiable soil. *East African Journal of Engineering,* 7(1), 187–198.

Shepard D (1968). A two-dimensional interpolation function for irregularly-spaced data. *Proceedings of the 1968 23rd ACM National Conference*, New York, USA, Association for Computing Machinery, 517–524.

Shooshpasha I, Hasanzadeh A, Taghavi A (2013). Prediction of the axial bearing capacity of piles by SPT-based and numerical design methods. *GEOMATE Journal*, 4(8), 560–564.

Soysal BF (2025). A discrete element method for evaluating the seismic performance of concrete gravity dam-reservoir systems under main shock-aftershock events. *Challenge Journal of Structural Mechanics*, 11(4), 229-244.

Stroud M (1974). The standard penetration test in insensitive clays and soft rocks. *Proceedings of the 1st European Symposium on Penetration Testing, Stockholm, Sweden*, 367–375.

Taylor KE (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, 106(D7), 7183–7192.

Telló M, Pulz LT, Telló VB, Ferraz RG, Vidor FF, Gazzana DS (2021). An enhanced soil characterization study supported by resistivity data processing and standard penetration test. *IEEE Transactions on Industry Applications*, 58(1), 49–58.

Terzaghi K (1943). Theoretical Soil Mechanics. 1 ed. Wiley, New York, USA.

Terzaghi K, Peck RB, Mesri G (1996). Soil Mechanics in Engineering Practice 3rd ed. John Wiley and Sons, New York, USA.

Thirmanpalli S, Kommu S, Kadali S (2024). Design of shallow foundation on clayey strata using ground improvement techniques – a numerical study. *Journal of Physics: Conference Series*, 2779(1), 012064.

Tibshirani R (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1), 267–288.

Troyanskaya O, Cantor M, Sherlock G, Brown P, Hastie T, Tibshirani R, Botstein D, Altman RB (2001). Missing value estimation methods for DNA microarrays. *Bioinformatics*, 17(6), 520–525.

U.S. Department of Energy Office of Scientific and Technical Information (2005). LWST phase i project conceptual design study: evaluation of design and construction approaches for economical hybrid steel/concrete wind turbine towers. https://www.osti.gov/biblio/15011444 [accessed 15.01.2026].

Vabalas A, Gowen E, Poliakoff E, Casson AJ (2019). Machine learning algorithm validation with a limited sample size. *PloS One*, 14(11), e0224365.

Vapnik V, Golowich S, Smola A (1996). Support vector method for function approximation, regression estimation and signal processing. *Proceedings of the 10th International Conference on Neural Information Processing Systems,* Denver, Colorado, 281-287.

Vesić AS (1973). Analysis of ultimate loads of shallow foundations. *Journal of the Soil Mechanics and Foundations Division*, 99(1), 45–73.

Webster R, Oliver MA (2007). Geostatistics for Environmental Scientists. John Wiley and Sons, New York, USA.

Wu H, Aguilar V, Montgomery J (2025). Applications of machine learning to foundation design for transportation structures. *Transportation Research Record*, 2679(1), 556–568.

Yaghoubi E, Yaghoubi E, Khamees A, Vakili AH (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*(21), 12655-12699.

Yilmaz M, Schubert S, Tinjum JM, Fratta D (2014). Foundation soil response to wind turbine generator loading. In: Geo-Congress 2014: Geo-Characterization and Modeling for Sustainability, ASCE Library, *Reston, Virginia,* USA, 1493–1502.

Zeini HA, Seno ME, Shehab EQ, Abood EA, Imran H, Bernardo LFA, Ribeiro TP (2025). Predicting the bearing capacity of shallow foundations on granular soil using ensemble machine learning models. *Geotechnics*, 5(3), 57.

Zou H, Hastie T (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical* Methodology, 67(2), 301–320.