Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Prediction of Local Scour Depth Around Cylindrical Bridge Piers in Cohesive Soil Using Advanced Machine Learning Algorithms

Nazila Kardan

n.kardan@azaruniv.ac.ir

Azarbaijan Shahid Madani University

Mehdi Soltani

Azarbaijan Shahid Madani University

Mehdi Majedi Asl

University of Maragheh

Tohid Omidpour

University of Maragheh

Research Article

Keywords: Local scour, Bridge pier, Cohesive bed, Machine learning, QNET, ANN, Froude number, Model validation, GEP

Posted Date: September 1st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10757515/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Iranian Journal of Science and Technology, Transactions of Civil Engineering on November 5th, 2026. See the published version at https://doi.org/10.1007/s40996-026-02298-7.

Prediction of Local Scour Depth Around Cylindrical Bridge Piers in Cohesive Soil Using Advanced Machine Learning Algorithms

Nazila Kardan[1*], Mehdi Soltani[1], Mehdi Majedi Asl[2], Tohid Omidpour[2*]

[1] Department of Civil Engineering, Azarbaijan Shahid Madani University, Tabriz, Iran

[1*] n.kardan@azaruniv.ac.ir

[2] Department of Civil Engineering, Maragheh University, Maragheh, Iran

[2*] tohidomidpour1377@gmail.com

ABSTRACT

Reliable prediction of local scour around cylindrical bridge piers in cohesive beds remains challenging because hydraulic forcing and soil resistance interact nonlinearly. This study evaluates five data-driven approaches—Artificial Neural Network (ANN), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Gene Expression Programming (GEP), and QNET—for predicting the dimensionless maximum scour depth, d_{max}/D, using 122 laboratory records selected from the cohesive-bed database compiled by Choi and Choi (2022). Four predictors were examined: water content (W_c), clay content (CP), pier Froude number Fr = V/$\sqrt{\text{gD}}$, and normalized critical shear strength $\tau_c/(\rho\, V^2)$. The 15 input patterns are the complete non-empty subsets of these four variables. On the reported 70:30 holdout split, QNET achieved the highest test performance (R^2 = 0.9958, RMSE = 0.0111, DC = 0.9836), followed by ANN (R^2 = 0.9206, RMSE = 0.0303, DC = 0.9285). SVM and KNN showed intermediate performance, whereas GEP showed weaker holdout generalization (R^2 = 0.6791) despite yielding an explicit formula. Parameter-deletion results for SVM, ANN, and GEP identified Fr as a highly influential hydraulic predictor, while QNET showed more distributed dependence among inputs. Because the sample is modest and the available analysis is based mainly on a single holdout split, the near-unity QNET performance is interpreted cautiously. The results demonstrate the potential of machine learning for cohesive-bed scour while emphasizing physically consistent predictors, leakage-free validation, and explicit assessment of model applicability.

Keywords: Local scour; Bridge pier; Cohesive bed; Machine learning; QNET; ANN; Froude number; Model validation; GEP.

1. Introduction

Local scour around bridge piers is one of the principal hydraulic processes affecting the stability and serviceability of bridge foundations [1–4]. Flow obstruction by a pier generates a complex three-dimensional flow field characterized by stagnation pressure, downward flow along the upstream face, horseshoe vortices near the bed, flow separation, and wake vortices. These coherent flow structures locally intensify the hydrodynamic forces acting on the bed and promote sediment removal from the vicinity of the foundation. Progressive scour may expose structural elements, reduce foundation embedment, and increase the vulnerability of bridges during flood events [1–6]. Reliable estimation of the maximum scour depth is therefore an essential component of bridge design, inspection, and hydraulic risk assessment. Figure 1 schematically illustrates the principal flow mechanisms and the main hydraulic and geotechnical variables relevant to pier scour in cohesive beds.

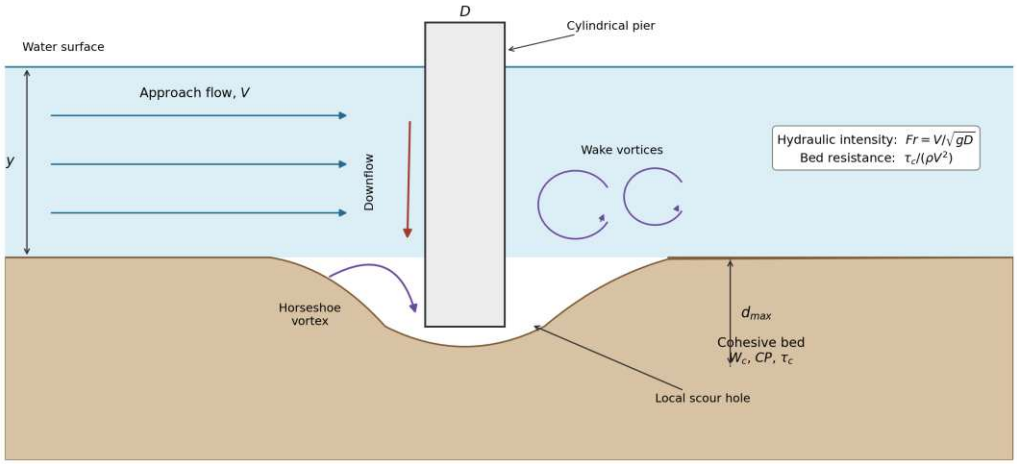


Figure 1. Schematic of local scour around a cylindrical bridge pier in a cohesive bed, showing the principal flow structures and governing variables.

Most established bridge-pier scour relationships have been developed primarily for non-cohesive sediments, in which bed resistance is largely characterized by sediment size, gradation, and threshold conditions for particle motion. Their direct application to cohesive beds is considerably more difficult because cohesive-soil erodibility cannot be described by grain size alone. Clay content, water content, bulk density, and critical bed shear strength affect interparticle bonding, erosion thresholds, and the response of the bed to applied hydraulic forces. Moreover, cohesive material may erode as aggregates, blocks, or flocs rather than as individual grains, resulting in a highly nonlinear interaction between hydrodynamic forcing and geotechnical resistance [14,17,19]. Consequently, predictive frameworks for cohesive-bed scour should account simultaneously for variables describing the flow regime and variables characterizing bed resistance.

The nonlinear nature of these interactions has motivated increasing application of data-driven techniques to scour prediction. Artificial Neural Networks (ANN), Support Vector Machines (SVM), Adaptive Neuro-Fuzzy Inference Systems (ANFIS), Gene Expression Programming (GEP), and related machine-learning methods have demonstrated varying levels of success in modeling scour processes [7,8,10–12,15,16,19]. Choi and Choi [19], for example, investigated the prediction of local scour around bridge piers in cohesive beds using SVM and a database compiled from previously published laboratory investigations. Their results demonstrated the potential of machine learning for cohesive-bed scour prediction while also highlighting the importance of training-data coverage and model generalization. These issues are particularly important when relatively small laboratory datasets are used, because excellent calibration performance does not necessarily guarantee reliable prediction for previously unseen hydraulic or geotechnical conditions.

Recent developments in machine learning have further shifted attention from prediction accuracy alone toward model robustness, interpretability, and transparent validation. Ensemble-learning techniques have increasingly been employed for hydraulic and scour-related prediction problems, while explainable artificial-intelligence approaches have been used to identify the relative influence of hydraulic, geometric, and sediment-related variables [25–29]. These developments indicate that the performance of a predictive model should not be evaluated solely through a high coefficient of determination. The stability of the model between training and testing subsets, sensitivity to the selected predictors, consistency with the underlying physics, and applicability beyond the calibration domain are equally important considerations.

Despite these advances, comparatively few studies have systematically examined several fundamentally different machine-learning paradigms under a common dimensionless framework specifically for cohesive-bed bridge-pier scour. In particular, the combined effects of hydraulic intensity and cohesive-bed resistance are not always examined consistently across alternative model structures, and very high predictive accuracy obtained from limited datasets requires careful interpretation with respect to possible overfitting and data-domain dependence. A comparative assessment that considers both predictive performance and physical consistency is therefore required.

Accordingly, the present study evaluates five machine-learning approaches—ANN, SVM, K-Nearest Neighbors (KNN), GEP, and QNET—for prediction of the dimensionless maximum scour depth around cylindrical bridge piers in cohesive beds. The models are examined using a common set of dimensionless hydraulic and geotechnical predictors, including the pier Froude number, water content, clay content, and normalized critical bed shear-strength parameter. Particular attention is given to the consistency of variable definitions, differences between training and holdout performance, and the physical interpretation of predictor sensitivity.

The specific objectives of this study are to: **(i)** formulate the cohesive-bed pier-scour problem using a physically consistent set of dimensionless hydraulic and geotechnical predictors; **(ii)** compare the predictive performance of five machine-learning approaches under a common data framework; **(iii)** evaluate the effect of different input-variable combinations on model performance; **(iv)** distinguish model fitting capability from holdout generalization and identify possible signs of model instability; and **(v)** discuss the applicability and limitations of laboratory-trained machine-learning models for prediction of bridge-pier scour in cohesive beds.

2. Materials and Methods

2.1. Physical basis and dimensional formulation

The variables governing local scour around bridge piers can generally be classified into three groups: hydraulic variables, sediment or soil properties, and pier-geometry variables. For non-cohesive beds, the equilibrium scour depth, d_{se}, is commonly expressed as a function of approach-flow velocity V, approach-flow depth y, median sediment size d_{50}, critical velocity for sediment motion V_c, and pier diameter D [30]:

$$d_{se} = f(V, y, d_{50}, V_c, D) \quad (1)$$

Equation (1) represents the conventional functional framework for scour in non-cohesive beds under conditions in which additional effects such as pier-shape correction, flow-attack angle, and sediment nonuniformity are either fixed or neglected.

For cohesive beds, however, the resistance of the bed to erosion cannot be characterized adequately by sediment size or critical velocity alone. Cohesive-soil erodibility is additionally governed by geotechnical properties such as water content W_c, clay content CP, and critical bed shear strength τ_c. In this study, W_c consistently denotes water content, whereas CP denotes clay content. Both parameters are expressed as percentages in the source database and are therefore treated as dimensionless descriptors of the soil state.

Accordingly, the dimensional functional relationship for the maximum scour depth in a cohesive bed can be written as:

$$d_{max} = f(V, y, d_{50}, D, \rho, g, W_c, CP, \tau_c) \quad (2)$$

Where ρ is the water density and g is gravitational acceleration. The use of maximum scour depth dmax, rather than equilibrium scour depth d_{se}, is consistent with the cohesive-bed database employed in the present analysis.

The dimensional-to-dimensionless formulation adopted here follows previously published cohesive-bed scour frameworks, particularly those used by Choi and Choi [19] and Debnath and Chaudhuri [31]. Therefore, Equations (3) and (4) are adopted and reorganized from established formulations and are not proposed as new dimensional-analysis relationships in the present study. This clarification is important because the purpose of the present work is to compare machine-learning models using a physically consistent predictor framework rather than to introduce a new scour-scaling law.

Using D as the characteristic length scale and V as the characteristic velocity scale, the maximum dimensionless scour depth may be expressed as:

$$\frac{d_{max}}{D} = f\left[\frac{V}{\sqrt{gD}}, \frac{y}{D}, \frac{d_{50}}{D}, W_c, CP, \frac{\tau_c}{\rho V^2}\right] \quad (3)$$

Where

$$Fr = \frac{V}{\sqrt{gD}}$$

Fr is the pier Froude number. The term

$$\frac{\tau_c}{\rho V^2}$$

represents the ratio of critical cohesive-bed resistance to the characteristic dynamic-pressure scale of the approach flow. Retaining this parameter ensures that the mechanical resistance of the cohesive bed remains explicitly represented in the predictive framework.

It should be emphasized that τ_c represents the critical bed shear strength or erosion-resistance threshold, whereas the applied approach-bed shear stress, τ_0, represents the hydraulic forcing acting on the bed. A consistent value of τ_0 was not available across all source datasets used in the compiled database. Consequently, τ_0 could not be introduced as an additional independent predictor without substantially reducing the usable dataset or introducing estimated values. This limitation is explicitly considered when interpreting the physical scope of the developed models.

Previous studies have shown that the influence of relative sediment size becomes comparatively small when D/d_{50} is sufficiently large, while the influence of relative approach depth becomes limited above specified y/D ranges. For cohesive-bed pier scour, previous formulations have commonly adopted simplifications in which the effects of d_{50}/D and y/D are neglected once the corresponding relative-dimension criteria are satisfied [31,32]. On this basis, and consistent with the predictor structure adopted in the source cohesive-bed database, the reduced functional relationship used for the machine-learning analysis is:

$$\frac{d_{max}}{D} = f\left[Fr, W_c, CP, \frac{\tau_c}{\rho V^2}\right] \quad (4)$$

This reduction should be interpreted strictly as a modeling assumption valid within the domain represented by the available laboratory data. It does not imply that flow depth or sediment size are universally irrelevant to cohesive-bed pier scour. Outside the parameter ranges represented in the database, these variables may again exert significant influence and should not automatically be excluded.

2.2. Database Compilation, Data Partitioning, and Input-Subset Design

The model-development database consists of 122 laboratory records reported in the cohesive-bed scour dataset used by Choi and Choi [19]. Each record contains hydraulic, geometric, and geotechnical information relevant to maximum local scour around cylindrical bridge piers. The available variables include approach-flow velocity (V), pier diameter (D), approach-flow depth (y), median sediment size (d_{50}), critical bed shear strength (τ_c), clay content (CP), water content (W_c), and maximum scour depth (d_{max}). The dimensionless maximum scour depth, d_{max}/D, was used as the response variable in the machine-learning analyses.

For model development, the 122 records were divided into approximately 70% training data and 30% holdout testing data, consistent with the partition reported in the Results section. The term holdout test subset is used consistently throughout the revised manuscript to avoid the ambiguity created by the term validation in the original Table 1. The holdout subset was reserved for evaluation of predictive performance and should not be interpreted as part of model calibration or hyperparameter selection. The reported ranges of the variables in the two subsets are summarized in Table 1. The original manuscript reports somewhat different overall ranges in the accompanying narrative; therefore, the revised manuscript uses the more detailed partition-specific ranges reported in the original table rather than reconstructing unverified aggregate statistics.

Table 1. Reported variable ranges for the training and holdout subsets of the 122-record model-development database.

Data subset	(V) (m/s)	(D) (m)	(y) (m)	(d_{max}) (mm)	(τ_c) (kPa)	(CP) (%)	(W_c) (%)	(d_{max}) (m)
Training (~70%)	0.199–0.830	0.025–0.210	0.160–0.600	0.0006–0.550	1.10–44.13	2–65	19.20–39.28	0.0142–0.270
Holdout test (~30%)	0.141–0.827	0.100–0.120	0.300–0.450	0.0055–0.182	5.70–35.60	8–52	10.70–45.92	0.0239–0.229

The source manuscript additionally identifies two independent published datasets containing 35 records from Gudavalli [33] and 25 records from Kothyari et al. [14]. These datasets should be clearly distinguished from the 122-record model-development database and were not part of the 70:30 partition. Their importance lies in the fact that several of their variable ranges extend beyond the domain represented by the training subset, thereby providing potentially useful datasets for assessing model transferability.

However, the available manuscript does not provide a complete and internally consistent set of external-validation performance statistics for these 60 records. Accordingly, the revised manuscript refers to them as external-domain datasets rather than claiming that a complete independent external validation has been demonstrated. Table 2 summarizes their role and the clearest reported domain extensions relative to the model-development training data.

Table 2. Additional external-domain datasets identified for assessment of model transferability.

Source dataset	No. of records	Intended role	Principal ranges extending beyond the model-development training domain
Gudavalli [33]	35	Independent external-domain assessment	(CP=75–100%), entirely above the training range; (W_c=2.26–39.30%), extending substantially below the training minimum
Kothyari et al. [14]	25	Independent external-domain assessment	(V=0.64–1.33) m/s; (d_{50}=1.271–2.49) mm; (W_c=6.15–17.09%), extending beyond the corresponding training ranges

Following the dimensional formulation introduced in Section 2.1, four predictors were retained in the reduced machine-learning framework: water content (W_c), clay content (CP), pier Froude number (Fr), and normalized critical bed shear strength $\tau_c/(\rho V^2)$. The response variable for every model was d_{max}/D.

Rather than selecting input patterns arbitrarily or through a separate statistical significance test, all possible non-empty subsets of the four predictors were considered. For four predictors, the total number of non-empty subsets is 24−1=15. Consequently, the 15 combinations listed in Table 3 constitute an exhaustive input-subset design. This structure allows the influence of predictor removal to be examined systematically and provides a common input framework for ANN, SVM, KNN, GEP, and QNET.

Table 3. Exhaustive input-predictor combinations evaluated by the five machine-learning models.

Combination	Predictors included	No. of predictors	Interpretation
1	$Fr, W_c, CP, \tau_c/\rho V^2$	4	Full predictor set
2	Fr, W_c, CP	3	Without normalized shear resistance
3	$Fr, CP, \tau_c/\rho V^2$	3	Without (W_c)
4	$Fr, W_c, \tau_c/\rho V^2$	3	Without (CP)
5	$W_c, CP, \tau_c/\rho V^2$	3	Without (Fr)
6	Fr, CP	2	Two-predictor subset
7	Fr, W_c	2	Two-predictor subset
8	W_c, CP	2	Two-predictor subset
9	$Fr, \tau_c/\rho V^2$	2	Two-predictor subset
10	$CP, \tau_c/\rho V^2$	2	Two-predictor subset
11	$W_c, \tau_c/\rho V^2$	2	Two-predictor subset
12	Fr	1	Single-predictor model
13	CP	1	Single-predictor model
14	W_c	1	Single-predictor model
15	$\tau_c/\rho V^2$	1	Single-predictor model

This exhaustive subset structure also provides a transparent basis for the subsequent parameter-deletion analyses. In particular, Combinations 2–5 represent systematic removal of one predictor from the full four-variable model, whereas Combinations 6–11 and 12–15 evaluate progressively reduced two-variable and single-variable representations, respectively. Consequently, differences in predictive performance among these combinations can be interpreted as model-specific evidence of predictor sensitivity rather than as direct measures of physical causality.

2.3. Predictor Space and Machine-Learning Model Development

Following the dimensional formulation described in Section 2.1, four predictors were used for model development: water content (W_c), clay content (CP), pier Froude number (Fr), and normalized critical bed shear strength $\tau_c/(\rho V^2)$. The dimensionless maximum scour depth, d_{max}/D, was considered as the model output. The ranges of the response variable and predictors are summarized in Table 4.

Table 4. Range of variables used in the machine-learning analysis.

Variable	Definition	Min	Max
d_{max}/D	Dimensionless maximum scour depth	0.18	3.04
(W_c) (%)	Water content	10.70	45.92
(CP) (%)	Clay content	2.00	65.00
(Fr)	Pier Froude number	0.19	1.68
($\tau_c/(\rho V^2)$)	Normalized critical bed shear strength	17.31	443.52

The predictor variables were scaled prior to model development to reduce the influence of differences in numerical magnitude. To avoid information leakage, preprocessing parameters should be determined from the training subset and subsequently applied to the holdout data. The retained modeling documentation confirms that normalization/standardization was performed, although the exact archived scaling coefficients are not available.

Five machine-learning approaches were evaluated: Artificial Neural Network (ANN), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Gene Expression Programming (GEP), and QNET. ANN was implemented using a multilayer-perceptron structure with nonlinear activation and gradient-based optimization; SVM employed a radial basis function kernel; KNN used Euclidean-distance-based regression; and GEP was applied as an evolutionary symbolic-regression technique. QNET was used as an alternative neural-network modeling framework. These algorithms represent different learning strategies and therefore provide a useful basis for comparing predictive accuracy and model behavior under identical input combinations.

Hyperparameter tuning was carried out using the training data. For SVM, the kernel parameter γ was searched within 0.001–10, with an optimum value of approximately 8 for the best-performing configuration. The complete final settings for ANN, KNN, GEP, and QNET were not preserved in the available modeling records; therefore, unsupported parameter values are not reconstructed in the revised manuscript. Model performance was subsequently evaluated on the holdout subset using R2, RMSE, and Willmott's index of agreement (DC).

The complete modeling procedure is illustrated in Figure 2, including database preparation, formation of the 15 input combinations, data partitioning and preprocessing, development of the five machine-learning models, holdout evaluation, and sensitivity analysis.

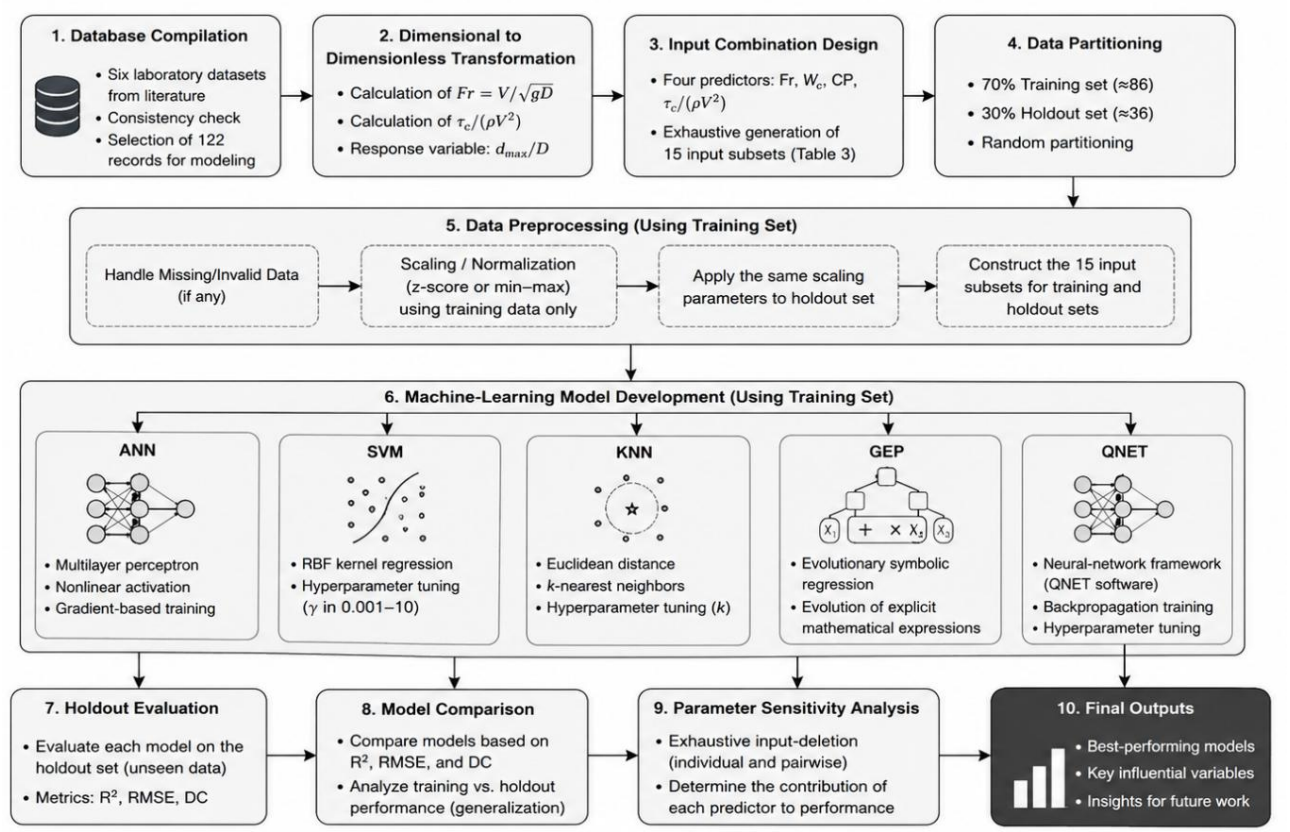


Figure 2. Machine-learning workflow for prediction of dimensionless local scour depth around bridge piers in cohesive beds.

2.4. Performance Evaluation Metrics

The predictive performance of the models was evaluated using three statistical indices: the coefficient of determination (R^2), root-mean-square error ($RMSE$), and Willmott's index of agreement (DC). These metrics provide complementary information on goodness of fit, prediction error, and agreement between observed and predicted values. Higher values of R^2 and DC, together with lower $RMSE$, indicate better predictive performance.

The coefficient of determination was calculated as

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(O_i - P_i)^2}{\sum_{i=1}^{N}(O_i - \bar{O})^2} \tag{5}$$

Where O_i and P_i denote the observed and predicted values, respectively, $\bar{O}$ is the mean of the observed values, and N is the number of observations. An R^2 value approaching unity indicates that a large proportion of the observed variability is reproduced by the model.

The root-mean-square error was calculated as

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(O_i - P_i)^2} \tag{6}$$

RMSE quantifies the overall magnitude of prediction errors and gives greater weight to relatively large deviations. Because the response variable in the present study is the dimensionless scour depth d_{max}/D, the reported $RMSE$ values are also dimensionless.

Willmott's index of agreement was computed as:

$$DC = 1 - \frac{\sum_{i=1}^{N}(O_i - P_i)^2}{\sum_{i=1}^{N}(|P_i - \bar{O}| + |O_i - \bar{O}|)^2} \tag{7}$$

DC varies from 0 to 1, with values closer to 1 indicating stronger agreement between predicted and observed values. In the present study, the three indices were interpreted jointly rather than relying on a single performance measure. Particular emphasis was placed on the holdout-test statistics because training-set metrics primarily represent goodness of fit, whereas holdout performance provides a more meaningful indication of model generalization.

3. Results and Discussion

The results are evaluated using the performance statistics retained in the original model outputs. As established in Section 2.4, R^2 corresponds to the squared Pearson correlation coefficient. The observed-versus-predicted series embedded in the manuscript reproduce the reported R^2 values, whereas the exact scaling/evaluation basis used for some archived $RMSE$ and DC values cannot be reconstructed consistently from the plotted series. Accordingly, $RMSE$ and DC are referred to as reported model-output statistics and are interpreted jointly with R^2. For cross-model comparison, the full four-predictor configuration is used as the common reference because it contains the same physical information for each algorithm; it is not assumed to be the numerical optimum for every individual metric or every reduced-input subset.

3.1. Support Vector Machine (SVM) Performance

For the full-input SVM reference configuration, the reported training statistics were R^2 = 0.8380, RMSE = 0.0185, and DC = 0.8999, whereas the holdout values were R^2 = 0.8008, RMSE = 0.0308, and DC = 0.7562. The reduction from training to holdout performance indicates a moderate generalization gap. Because some reduced-input SVM configurations improve one metric while degrading others, the full-input configuration is retained here as the physically complete reference rather than being described as uniquely optimal. The corresponding reference statistics are summarized in Table 5.

Table 5. Reported performance of the selected full-input SVM reference configuration.

Model	Input set	γ	R^2 Train	$RMSE$ Train	DC Train	R^2 Holdout	$RMSE$ Holdout	DC Holdout
SVM	Full four-predictor set	≈ 8	0.8380	0.0185	0.8999	0.8008	0.0308	0.7562

Figure 3 summarizes the reported training and holdout statistics. Figure 4 presents the observed-versus-predicted series retained in the source manuscript. The holdout scatter confirms useful predictive skill but also shows non-negligible deviations from the 1:1 line. From an engineering perspective, deviations at the upper end of the scour-depth range are especially important because underprediction of large scour depths may be unconservative.

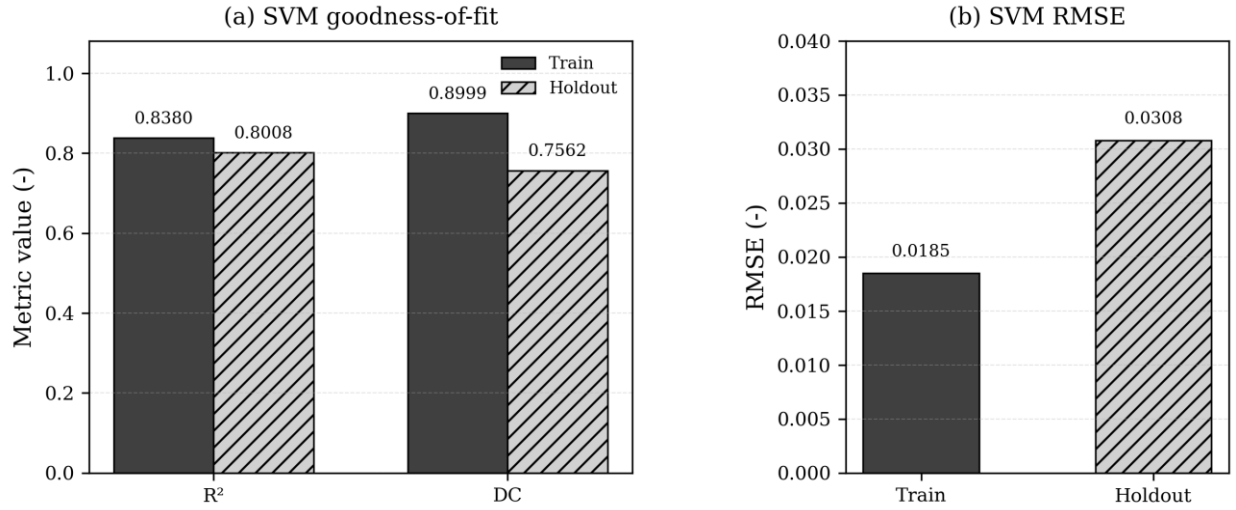


Figure 3. Reported training and holdout performance of the selected full-input SVM reference configuration: (a) R^2 and DC and (b) RMSE.

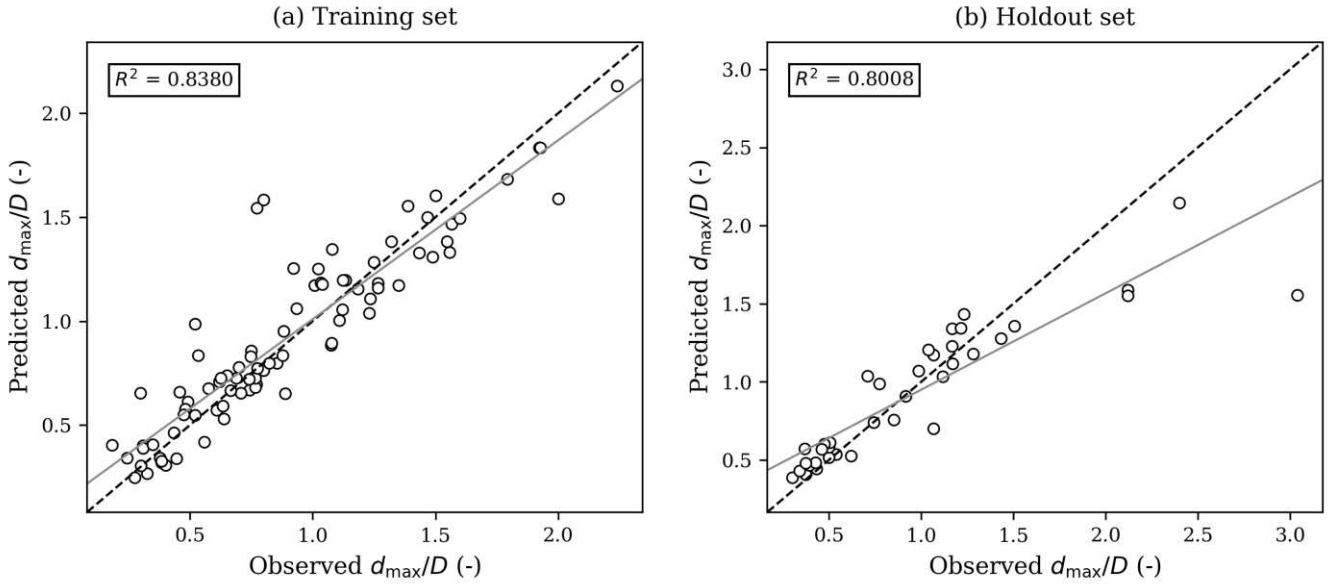
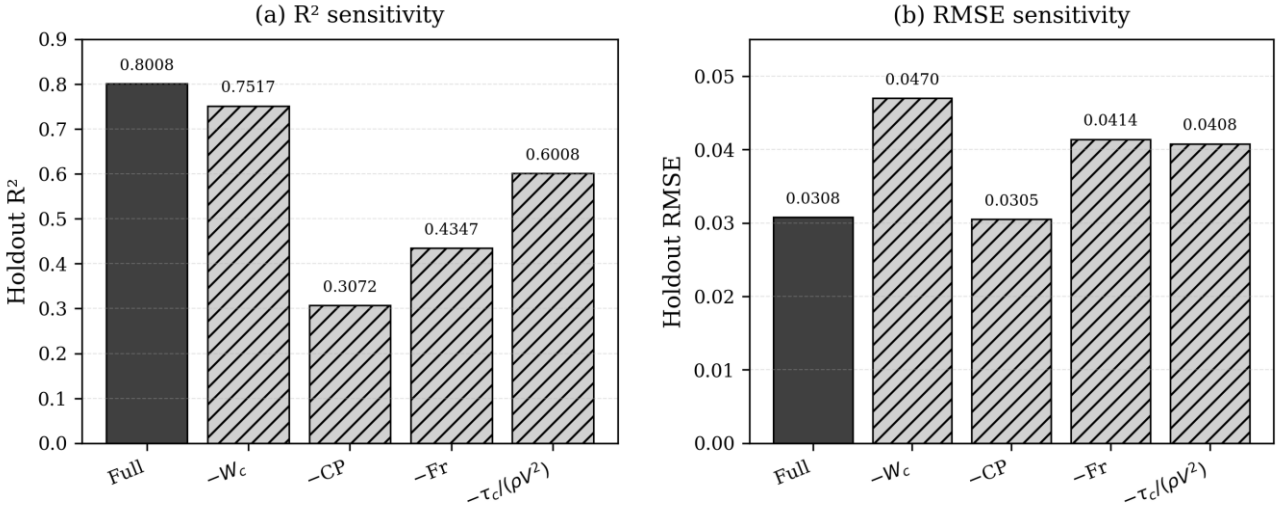


Figure 4. Observed-versus-predicted values available for the SVM training and holdout subsets; the dashed line denotes perfect agreement.

3.2. Sensitivity Analysis of the SVM Model

The one-at-a-time deletion analysis is consistent with the input combinations defined in Table 3. Removing CP caused the largest reduction in holdout R^2, from 0.8008 to 0.3072. Removing Fr caused the strongest reduction in DC, from 0.7562 to 0.3279, whereas removing W_c produced the largest reported increase in $RMSE$, from 0.0308 to 0.0470. The sensitivity pattern therefore depends on the selected metric and does not support a universal single-variable dominance claim for SVM. Instead, the results indicate complementary dependence on hydraulic intensity and cohesive-bed properties. The holdout sensitivity pattern is summarized in Figure 5.



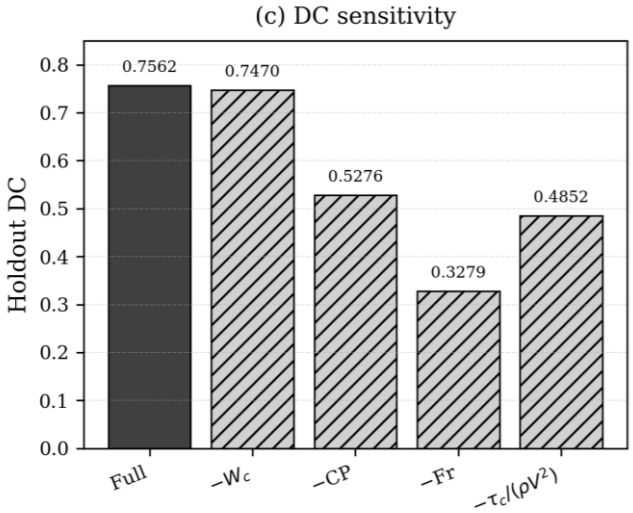


Figure 5. SVM sensitivity to one-at-a-time predictor deletion b7ased on the reported holdout metrics: (a) R^2, (b) $RMSE$, and (c) DC.

3.3. QNET Performance

QNET produced the strongest reported full-input performance among the five evaluated algorithms. The training statistics were R^2 = 0.9975, $RMSE$ = 0.0073, and DC = 0.9973, while the holdout statistics were R^2 = 0.9958, $RMSE$ = 0.0111, and DC = 0.9836. The small train-holdout difference is encouraging for the retained partition; however, the near-unity R^2 should be interpreted as evidence for this specific split rather than proof of split-independent or field-scale generalization. The reported reference statistics are listed in Table 6.

Table 6. Reported performance of the selected full-input QNET reference configuration.

Model	Input set	R^2 Train	$RMSE$ Train	DC Train	R^2 Holdout	$RMSE$ Holdout	DC Holdout
QNET	Full four-predictor set	0.9975	0.0073	0.9973	0.9958	0.0111	0.9836

Figure 6 summarizes the reported performance statistics, and Figure 7 shows the archived observed-versus-predicted series. The parity plots corroborate the very high reported R^2 values, with most points concentrated close to the line of perfect agreement. Because the complete QNET project file and repeated resampling history are unavailable, stronger claims regarding robustness are intentionally avoided.

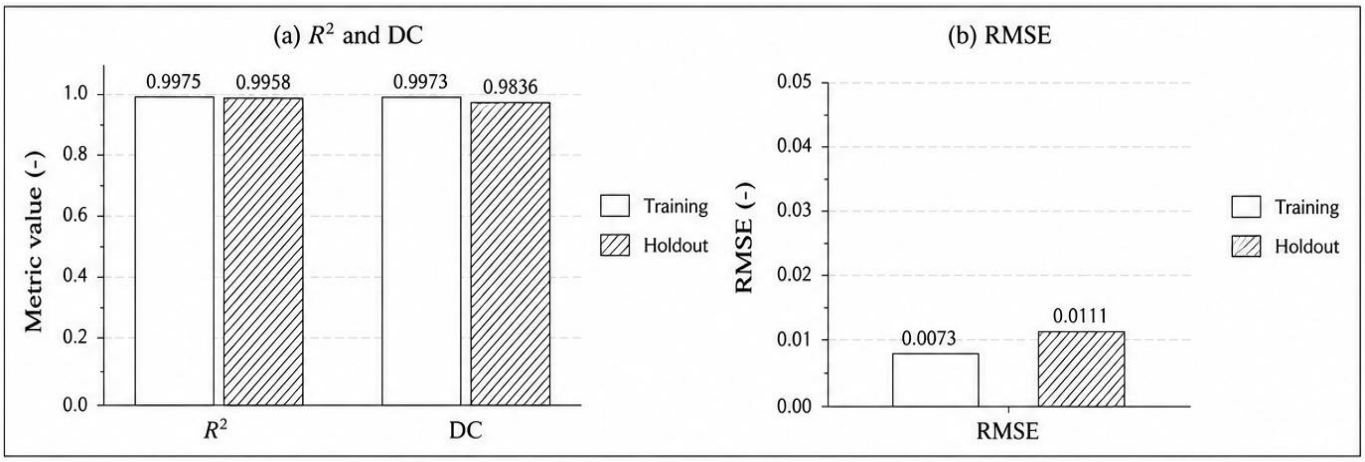


Figure 6. Reported training and holdout performance of the selected full-input QNET reference configuration: (a) R^2 and DC and (b) $RMSE$.

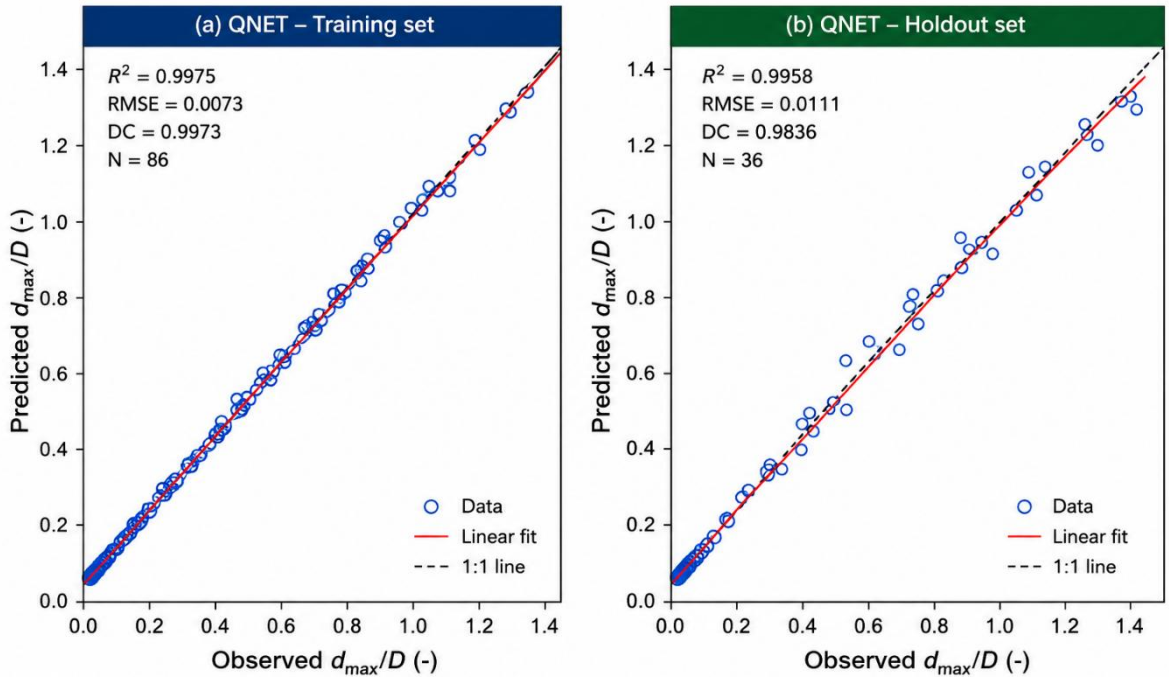


Figure 7. Observed-versus-predicted values available for the QNET training and holdout subsets; the dashed line denotes perfect agreement.

3.4. Sensitivity Analysis of the QNET Model

The original QNET sensitivity table contained an inconsistency between the deletion labels and the 15 input combinations. To resolve this without introducing new calculations, the one-at-a-time deletion cases were matched directly to the exhaustive subset definitions in Table 3 and to the corresponding reported QNET model outputs: Model 2 excludes the normalized critical shear-strength term, Model 3 excludes Wc, Model 4 excludes CP, and Model 5 excludes Fr. The corrected mapping is presented in Table 7.

Table 7. Corrected QNET one-at-a-time predictor-deletion results obtained by matching the reported QNET outputs to the input combinations defined in Table 3.

Condition	R^2 Train	$RMSE$ Train	DC Train	R^2 Holdout	$RMSE$ Holdout	DC Holdout
Full input set	0.9975	0.0073	0.9973	0.9958	0.0111	0.9836
Without $\tau_c/(\rho V^2)$	0.9975	0.0073	0.9973	0.9958	0.0111	0.9836
Without W_c	0.9544	0.0082	0.9530	0.9743	0.0243	0.9645
Without CP	0.9254	0.0090	0.9054	0.9539	0.0707	0.9339
Without Fr	0.8863	0.0089	0.8860	0.8521	0.0399	0.8321

After correction, removal of Fr causes the largest decline in holdout R^2 and DC, whereas removal of CP produces the largest reported $RMSE$. The identical full-model and no-shear-term statistics are unusual. Because this duplication is present in the archived QNET outputs, it is retained rather than altered and should be verified if the original QNET project becomes available. The sensitivity results therefore support a strong role for Fr but do not justify assigning exact physical contribution percentages. Figure 8 summarizes the corrected holdout sensitivity pattern.

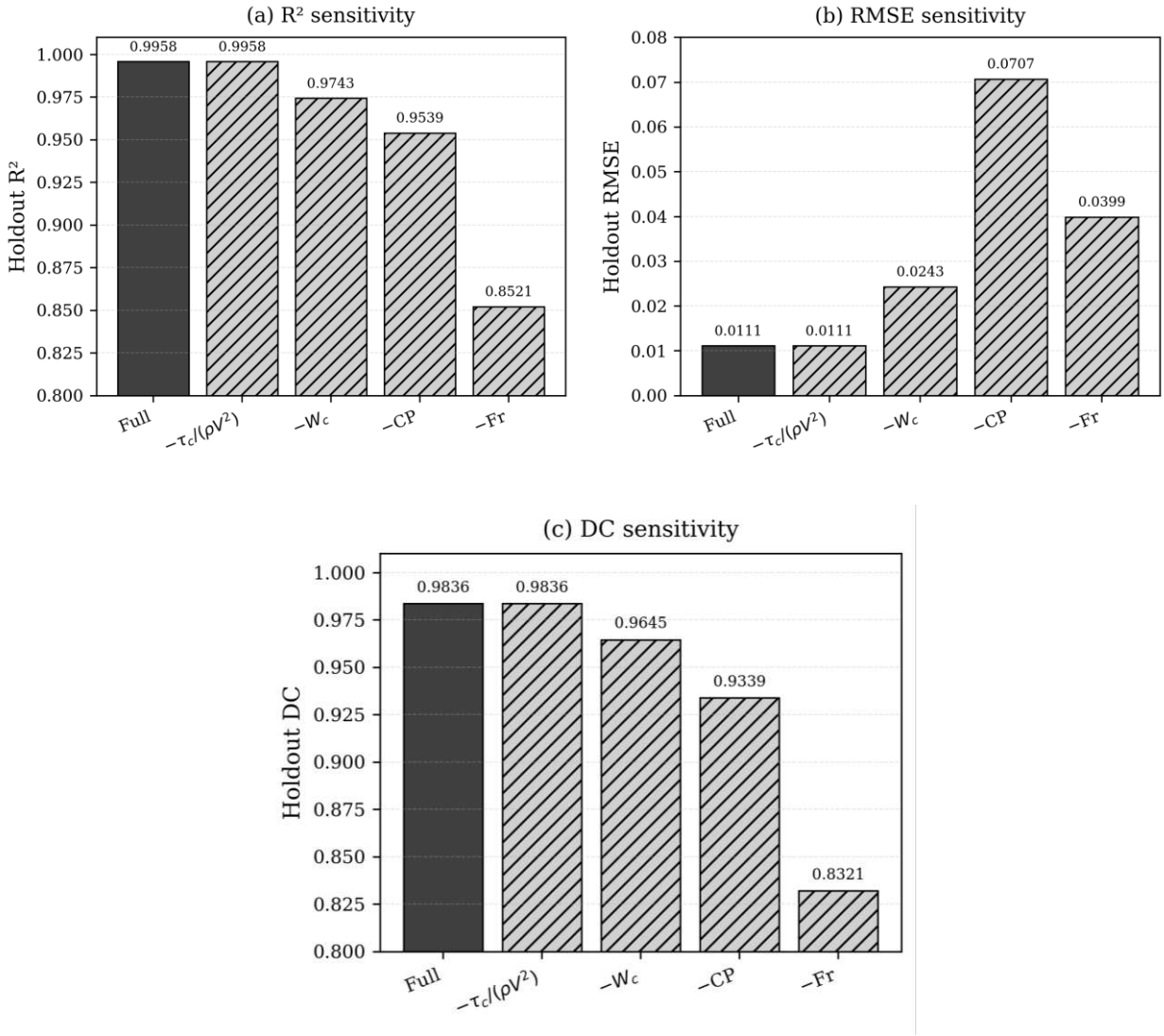


Figure 8. Corrected QNET sensitivity to one-at-a-time predictor deletion based on the reported holdout metrics: (a) R^2, (b) $RMSE$, and (c) DC.

3.5. Artificial Neural Network (ANN): Performance and Sensitivity

For the full-input ANN reference configuration, the reported training values were R^2 = 0.9996, $RMSE$ = 0.0244, and DC = 0.9837, whereas the holdout values were R^2 = 0.9206, $RMSE$ = 0.0303, and DC = 0.9285. The decrease in R^2 from training to holdout testing indicates a measurable generalization gap, although the reported holdout performance remains strong. Table 8 compiles the full-input and one-at-a-time deletion results.

Table 8. Reported ANN full-input and one-at-a-time predictor-deletion results.

Condition	R^2Train	$RMSE$ Train	DC Train	R^2 Holdout	$RMSE$ Holdout	DC Holdout
Full input set	0.9996	0.0244	0.9837	0.9206	0.0303	0.9285
Without $\tau_s/(\rho V^2)$	0.8093	0.0327	0.8093	0.9207	0.0317	0.9262
Without W_c	0.9359	0.0289	0.9359	0.8977	0.0434	0.8797
Without CP	0.9224	0.0424	0.9373	0.9071	0.0403	0.8920
Without Fr	0.6400	0.0250	0.7342	0.5638	1.3580*	0.5063

*The RMSE = 1.358 value is retained exactly as reported in the archived output and should be checked against the original ANN computation if available.

Removal of Fr causes the largest simultaneous decline in holdout R^2 and DC, supporting the importance of hydraulic intensity in the ANN response. However, the reported $RMSE$ = 1.358 for the Fr-deletion case is an extreme outlier relative to the other ANN $RMSE$ values and cannot be independently verified from the available source files. It is therefore retained as a reported value but excluded from quantitative feature-

ranking arguments. Figure 9 summarizes the full-input training and holdout metrics, Figure 10 presents the archived observed-versus-predicted relationship, and Figure 11 emphasizes the more internally consistent R^2 and DC sensitivity results.

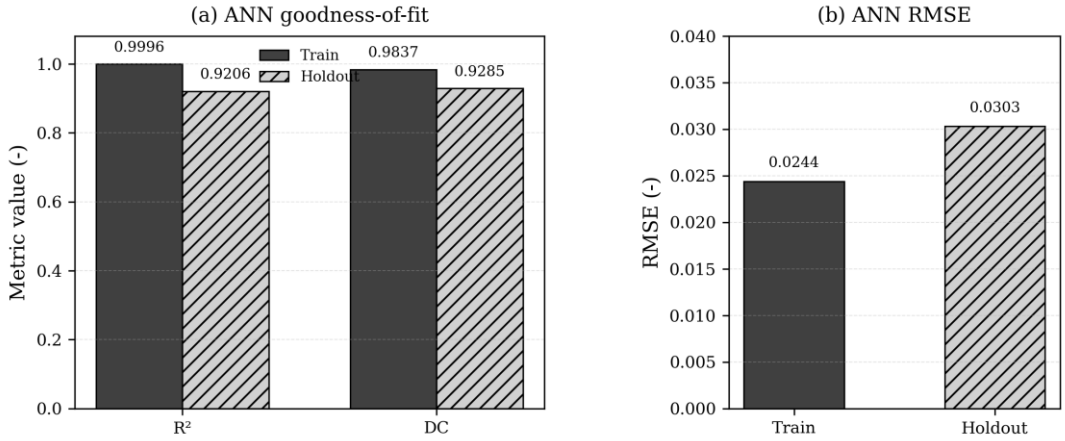


Figure 9. Reported training and holdout performance of the selected full-input ANN reference configuration: (a) R^2 and DC and (b) *RMSE*.

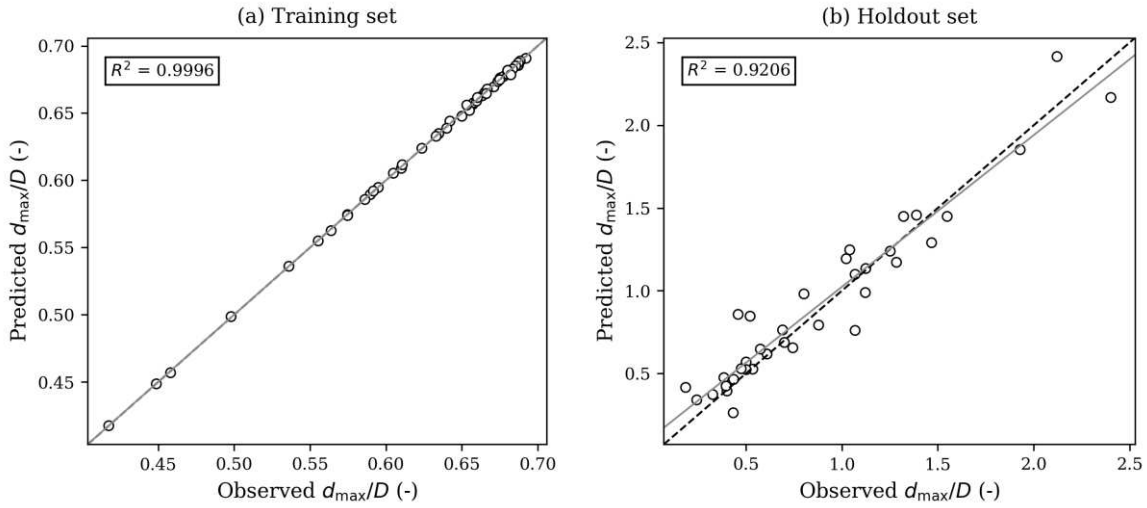


Figure 10. Observed-versus-predicted values retained in the source ANN plots; the dashed line denotes perfect agreement.

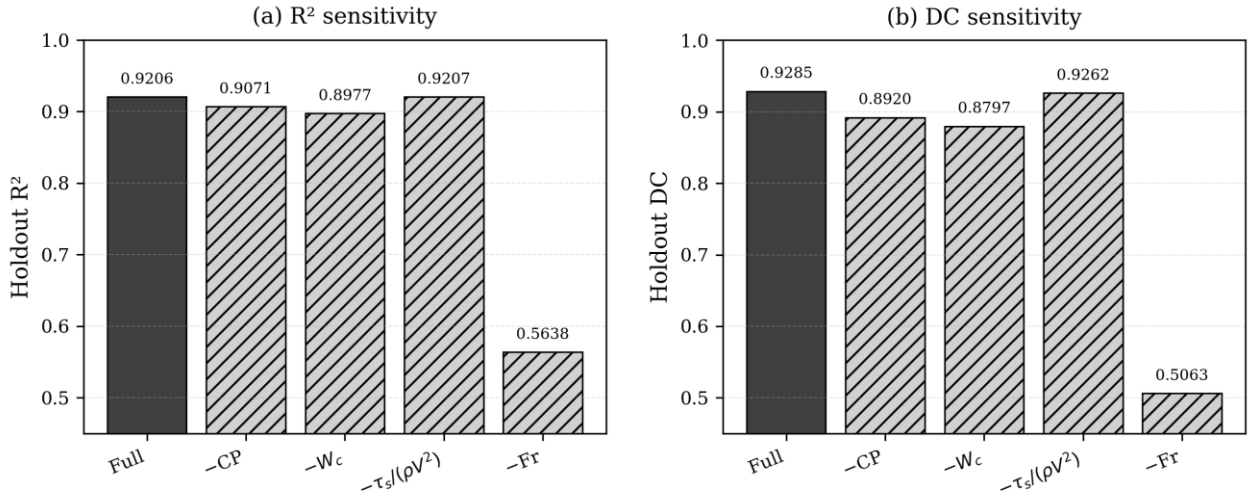


Figure 11. ANN sensitivity to one-at-a-time predictor deletion based on the reported holdout (a) R^2 and (b) *DC*. *RMSE* is not used in this figure because the reported *Fr*-deletion *RMSE* is anomalous.

3.6. K-Nearest Neighbors (KNN) Performance

The retained KNN results report holdout R^2 = 0.8104, $RMSE$ = 0.1529, and DC = 0.8669. These statistics indicate intermediate predictive capability, with a substantially larger reported RMSE than SVM or ANN. Complete KNN training statistics and a full 15-subset performance record were not preserved in the available manuscript files; consequently, no unsupported statement is made regarding the unique optimum KNN configuration. The available holdout statistics are summarized in Table 9.

Table 9. Reported KNN holdout performance retained in the manuscript.

Model	R^2 Holdout	*RMSE* Holdout	DC Holdout	Interpretation
KNN	0.8104	0.1529	0.8669	Intermediate reported holdout performance; complete training/subset record unavailable

Figure 12 presents the archived KNN holdout observed-versus-predicted series. The larger scatter relative to QNET and ANN is consistent with the lower reported predictive precision. Because KNN is a local distance-based method, its transferability is expected to be particularly sensitive to feature scaling and to the density of observations in predictor space.

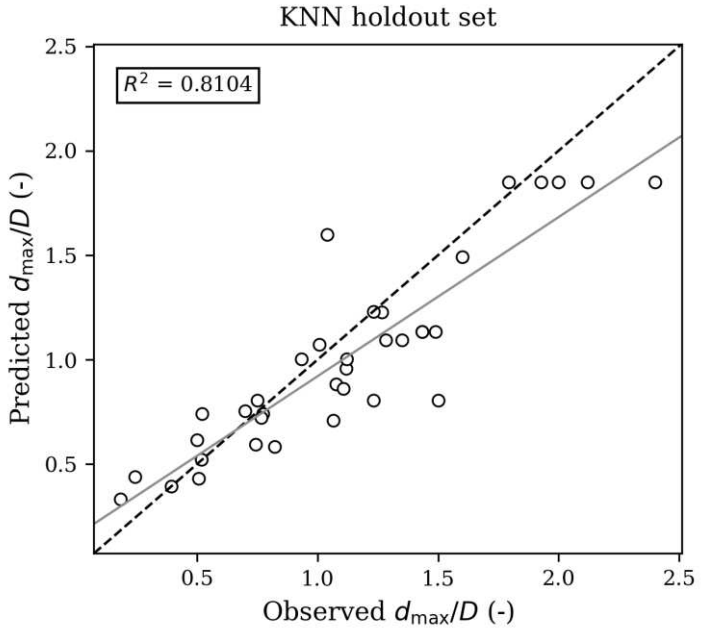


Figure 12. Observed-versus-predicted values available for the reported KNN holdout subset; the dashed line denotes perfect agreement.

3.7. Gene Expression Programming (GEP): Performance, Sensitivity, and Symbolic Interpretation

For the full-input GEP reference configuration, the reported training statistics were R^2 = 0.9091, $RMSE$ = 0.1739, and DC = 0.9020, while the holdout statistics were R^2 = 0.6791, $RMSE$ = 0.2498, and DC = 0.6030. The original manuscript described the full-input GEP configuration as the best model, but the archived 15-subset output does not support a unique optimum across all metrics. For example, a reduced W_c-Fr configuration reports a higher holdout R^2 and lower $RMSE$, whereas another reduced configuration reports the highest DC. The full-input GEP result is therefore retained only as the common physically complete reference and as the basis of the reported symbolic expression. Table 10 summarizes the full-input and one-at-a-time deletion statistics.

Table 10. Reported GEP full-input and one-at-a-time predictor-deletion results.

Condition	R^2 Train	$RMSE$ Train	DC Train	R^2 Holdout	$RMSE$ Holdout	DC Holdout
Full input set	0.9091	0.1739	0.9020	0.6791	0.2498	0.6030
Without $\tau_s/(\rho V^2)$	0.7052	0.3555	0.7334	0.4624	0.4522	0.8745
Without W_c	0.5214	1.4600	0.4729	0.3906	1.3492	0.4601
Without CP	0.7817	0.3448	0.4893	0.5352	1.3255	0.4166
Without Fr	0.3214	1.4600	0.3727	0.2906	1.3592	0.3501

Within the one-at-a-time deletion analysis, removing Fr produces the largest decrease in holdout R^2. However, the reported DC after removing the normalized shear-resistance term increases despite deterioration in R^2 and $RMSE$, demonstrating that the archived GEP metrics are not fully concordant. The GEP sensitivity results are therefore interpreted qualitatively and are not converted into exact feature-contribution percentages.

The principal advantage of GEP is the explicit symbolic expression obtained from evolutionary search. Nevertheless, nested trigonometric operators and fitted constants are data-driven curve-fitting structures rather than independent hydraulic laws. Given the modest full-input holdout R^2, extrapolation beyond the experimental domain should be avoided without independent validation. The full-input training-holdout contrast is shown in Figure 13.

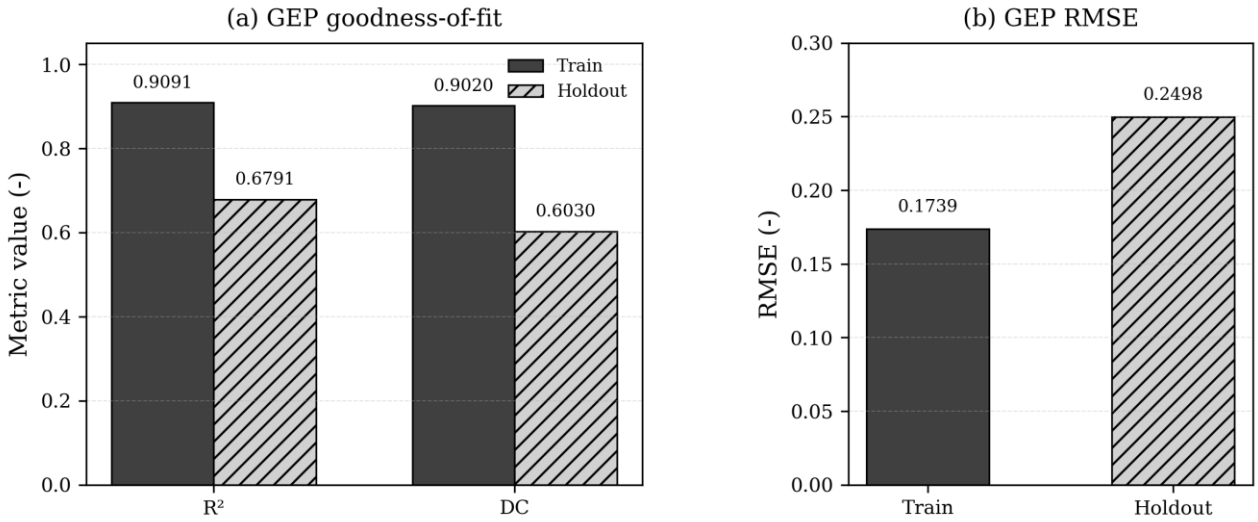


Figure 13. Reported training and holdout performance of the full-input GEP reference configuration: (a) R^2 and DC and (b) $RMSE$.

3.8. Cross-Model Comparison and Physical Interpretation

Table 11 compares the reported holdout statistics for the common full-input/reference configurations. QNET and ANN provide the strongest reported results. SVM and KNN show intermediate performance, while the full-input GEP configuration is weaker. Because several archived RMSE and DC values cannot be independently reconstructed from the plotted observed-versus-predicted series, small cross-model differences in these two metrics should not be overinterpreted; R^2 provides the most directly reproducible cross-check from the retained charts.

Table 11. Reported holdout statistics for the common full-input/reference model comparison.

Model	R^2 Holdout	*RMSE* Holdout	DC Holdout	Interpretation
QNET	0.9958	0.0111	0.9836	Strongest reported full-input holdout results
ANN	0.9206	0.0303	0.9285	Strong reported holdout performance
KNN	0.8104	0.1529	0.8669	Intermediate; complete training/subset record unavailable
SVM	0.8008	0.0308	0.7562	Intermediate; reduced subsets improve some individual metrics
GEP	0.6791	0.2498	0.6030	Full-input reference; reduced subsets may improve individual metrics

The combined sensitivity evidence supports a physically coherent interpretation: Fr is repeatedly influential, but it does not eliminate the need for geotechnical descriptors. *Fr* characterizes hydraulic intensity, whereas W_c, CP, and the normalized critical shear-strength term represent complementary aspects of cohesive-bed state and erosion resistance. The most defensible conclusion is therefore that predictive performance benefits from representing hydraulic forcing and bed resistance together.

No direct superiority claim over HEC-18 or other empirical scour equations is made in this section because the retained files do not contain a reproducible empirical-equation calculation on the same holdout records. Likewise, recent ensemble methods reported in the literature [25-28] are discussed as relevant benchmarks but are not assigned numerical performance in the present study because they were not rerun on the archived database. Figure 14 provides a compact visual comparison of the reported full-input/reference holdout statistics.

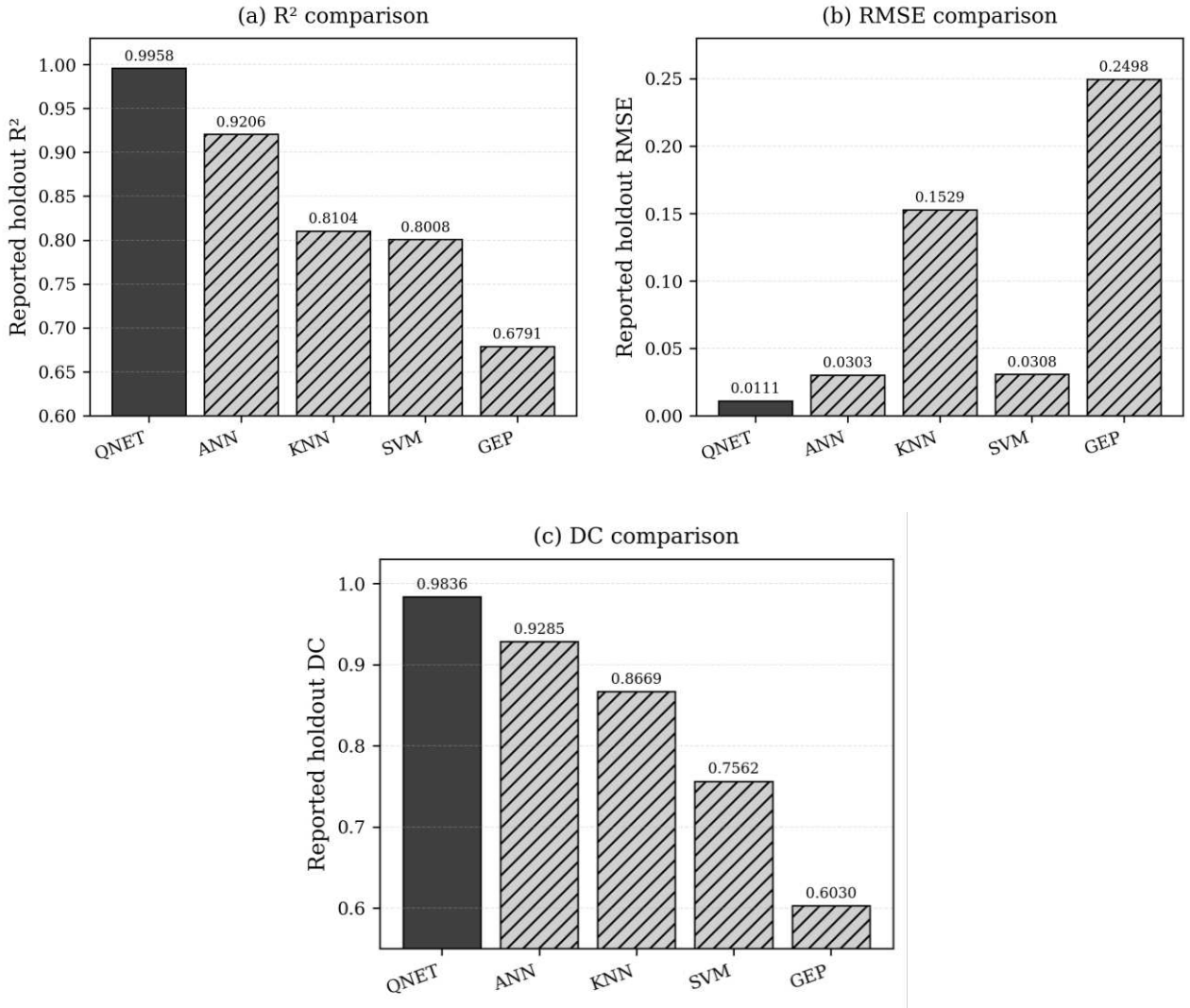


Figure 14. Comparison of the reported holdout statistics for the common full-input/reference configurations: (a) R^2, (b) *RMSE*, and (c) DC.

3.9. Consistency Limitations, Engineering Implications, and Future Validation

Several limitations must be considered when interpreting the results. First, the development database contains 122 laboratory records and the principal evidence is based on a single reported 70:30 partition. Second, complete random-seed, preprocessing, and hyperparameter logs are unavailable for every model. Third, the archived plotted sample counts are not identical across all algorithms, which may reflect incomplete chart ranges rather than different computational partitions; the original model files are required to resolve

this point. Fourth, some archived *RMSE* and DC values are not independently reproducible from the plotted series, although the reported R^2 values are consistent with squared Pearson correlation.

These issues do not negate the qualitative model comparison, but they limit the strength of claims regarding universal superiority or field-scale generalization. A definitive rerun should use leakage-free repeated or nested k-fold cross-validation, preserve all random seeds and hyperparameters, calculate all performance measures from one verified observed-predicted dataset, and evaluate the Gudavalli and Kothyari et al. datasets strictly as independent external tests. Residual distributions, MAE, bias, uncertainty intervals, and performance at high scour depths should also be reported.

For future model development, a well-tuned ensemble benchmark such as XGBoost, coupled with SHAP-based interpretation, would provide a useful contemporary comparison and a more transparent assessment of predictor effects. For engineering application, however, field validation and direct characterization of applied bed shear stress, bulk density, erosion rate, and cohesive strength remain more important than adding numerous additional algorithms. The present models should therefore be regarded as data-driven prediction tools within the represented laboratory domain rather than stand-alone design equations.

4. Conclusions and Recommendations

This study evaluated five machine-learning approaches—Artificial Neural Network (ANN), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Gene Expression Programming (GEP), and QNET—for predicting the dimensionless maximum scour depth around cylindrical bridge piers in cohesive beds. A common four-predictor framework was adopted using the pier Froude number (Fr), water content (W_c), clay content (CP), and normalized critical bed shear strength $\tau_C/(\rho V^2)$, thereby representing both hydraulic forcing and cohesive-bed resistance within the same modeling structure.

For the common full-input/reference configuration, QNET produced the strongest reported holdout performance (R^2 = 0.9958, *RMSE* = 0.0111, and DC = 0.9836), followed by ANN (R^2 = 0.9206, *RMSE* = 0.0303, and DC = 0.9285). SVM and KNN showed intermediate predictive performance, whereas the full-input GEP configuration exhibited weaker holdout generalization despite its advantage of providing an explicit symbolic expression. These results should be interpreted as reported performance for the retained data partition rather than as definitive evidence of universal model superiority.

The predictor-deletion analyses showed that feature influence was model- and metric-dependent. For SVM, CP caused the largest reduction in holdout R^2, Fr caused the largest reduction in DC, and W_c produced the largest reported increase in RMSE. After correcting the QNET deletion mapping to match the input combinations defined in Table 3, removal of Fr produced the largest decline in holdout R^2 and DC, whereas removal of CP generated the largest reported RMSE. ANN also showed a pronounced reduction in R^2 and DC when Fr was removed. The GEP results similarly indicated a strong influence of Fr, although inconsistencies among its reported sensitivity metrics preclude quantitative attribution of exact feature contributions. Overall, the results support the combined use of hydraulic and geotechnical predictors rather than a single-variable description of cohesive-bed scour.

Several limitations must be considered. The model-development database contains only 122 laboratory records, the main evaluation is based on a single reported 70:30 partition, and complete preprocessing, random-seed, and hyperparameter records are not available for every algorithm. In addition, some archived *RMSE* and DC values cannot be independently reconstructed from the plotted observed–predicted series, although the reported R^2 values are consistent with the squared Pearson correlation formulation adopted in Section 2.4. Consequently, the developed models should be regarded as data-driven prediction tools within the represented laboratory domain and not as unrestricted field design equations.

Funding: No external funding was received for this research.

Conflicts of interest: The authors declare no known competing financial interests or personal relationships that could have influenced the work reported in this manuscript.

Authors' Contribution Statement:

Nazila Kardan: Methodology, Data curation, Original Draft.

Mahdi Soltani Sotobadi: Conceptualization, Supervision, Validation, Writing.

Mahdi Majedi Asl: Formal Analysis, Visualization, Software.

Tohid Omidpour Alavian: Writing - Review & Editing, Investigation, Project Administration, Resources.

References

[1] Melville, B. W., & Sutherland, A. J. (1988). Design method for local scour at bridge piers. Journal of Hydraulic Engineering, 114(10), 1210–1226. https://doi.org/10.1061/(ASCE)0733-9429(1988)114:10(1210)
[2] Breusers, H. N. C., & Raudkivi, A. J. (1991). Scouring (Hydraulic Structures Design Manual Series, Vol. 2). London: CRC Press. https://doi.org/10.1201/9781003079477
[3] Raudkivi, A. J. (1998). Loose boundary hydraulics. CRC Press. https://doi.org/10.1201/9781003077800
[4] Melville, B. W., & Coleman, S. E. (1999). Bridge scour. Water Resources Publications. ISBN 1-887201-18-1.
[5] Bozkus, Z., & Yildiz, O. (2004). Effects of inclination of bridge piers on scouring depth. Journal of Hydraulic Engineering, 130(8), 827–830. https://doi.org/10.1061/(ASCE)0733-9429(2004)130:8(827)
[6] Sheppard, D. M., & Miller Jr., W. (2006). Live-bed local pier scour experiments. Journal of Hydraulic Engineering, 132(7), 635–642. https://doi.org/10.1061/(ASCE)0733-9429(2006)132:7(635)
[7] Choi, S. U., & Cheong, S. H. (2006). Prediction of local scour around bridge piers using artificial neural networks. Journal of the American Water Resources Association, 42(2), 487–494. [8] Bateni, S. M., Borghei, S. M., & Jeng, D.-S. (2007). Neural network and neuro-fuzzy assessments for scour depth around bridge piers. Engineering Applications of Artificial Intelligence, 20(3), 401–414.
[9] Haltigin, T. W., Biron, P. M., & Lapointe, M. F. (2007). Predicting equilibrium scour-hole geometry near angled stream deflectors using a three-dimensional numerical flow model. Journal of Hydraulic Engineering, 133(8), 983–993. https://doi.org/10.1061/(ASCE)0733-9429(2007)133:8(983)
[10] Azamathulla, H. Md., Ab Ghani, A., Zakaria, N. A., & Guven, A. (2009). Genetic programming to predict bridge pier scour. Journal of Hydraulic Engineering, 136(3). https://doi.org/10.1061/(ASCE)HY.1943-7900.0000133
[11] Muzzammil, M., & Alam, J. (2011). ANFIS-based approach to scour depth prediction at abutments in armored beds. Journal of Hydroinformatics, 13(4), 699–712. https://doi.org/10.2166/hydro.2010.006
[12] Ghazanfari-Hashemi, S., Etemad-Shahidi, A., Kazeminezhad, M. H., & Mansoori, A. R. (2011). Prediction of pile group scour in waves using support vector machines and ANN. Journal of Hydroinformatics, 13(4), 609–620. https://doi.org/10.2166/hydro.2010.107
[13] Maddison, B. (2012). Scour failure of bridges. Proceedings of the Institution of Civil Engineers – Forensic Engineering, 165(1), 39–52. https://doi.org/10.1680/feng.2012.165.1.39
[14] Kothyari UC, Kumar A, Jain RK (2014) Influence of cohesion on river bed scour in the wake region of piers. Journal of Hydraulic Engineering 140(1):1-13, DOI: 10.1061/(ASCE)HY.1943-7900.0000793
[15] Najafzadeh, M., Etemad-Shahidi, A., & Lim, S. Y. (2016). Scour prediction in long contractions using ANFIS and SVM. Ocean Engineering, 111, 128–135. https://doi.org/10.1016/j.oceaneng.2015.10.053
[16] Parsaie, A., Haghiabi, A. H., & Moradinejad, A. (2019). Prediction of scour depth below river pipeline using support vector machine. KSCE Journal of Civil Engineering, 23(5), 2503–2513. https://doi.org/10.1007/s12205-019-0504-6
[17] Mahalder, B., Schwartz, J. S., Palomino, A. M., & Zirkle, J. (2021). Scour hole development in natural cohesive bed sediment around cylinder-shaped piers subjected to varying sequential flow events. Water, 13(22), 3289. https://doi.org/10.3390/w13223289

[18] Bordbar, A., Sharifi, S., & Hemida, H. (2021). Investigation of the flow behaviour and local scour around single square-shaped cylinders at different positions in live-bed. Ocean Engineering, 238, 109772. https://doi.org/10.1016/j.oceaneng.2021.109772
[19] Choi, S.-U., & Choi, S. (2022). Prediction of local scour around bridge piers in the cohesive bed using support vector machines. KSCE Journal of Civil Engineering, 26(5), 2174–2182. https://doi.org/10.1007/s12205-022-1803-9
[20] Rafiqii, M. A., Lone, M. A., & Tantray, M. A. (2023). Exploring the challenges and solutions for river scour in dynamic environments: A comprehensive review. World Journal of Engineering, 21(6), 1045–1063. https://doi.org/10.1108/WJE-04-2023-0113
[21] Shams, G., Haghighatjou, P., & Hatefi, S. M. (2025). CFD and machine learning for scour prediction at square bridge piers: South Platte, Mississippi, Choptank, and Pearl rivers. Aerospace Traffic and Safety, 2(3–4), 145–156. https://doi.org/10.1016/j.aets.2025.11.001
[22] Huynh, N. T., Phan, A. T. T., Huynh, Q. D., Nguyen, H. H. D., & Nguyen, T. N. (2025). Machine learning-based prediction of scour depth around complex bridge piers with model interpretability. Ocean Engineering, 340, 122479. https://doi.org/10.1016/j.oceaneng.2025.122479
[23] Soltani Sotobadi, M., Omidpour Alavian, T., & Kardan, N. (2025). Forecasting the Local Scour Depth around Cylindrical Bridge Foundations on Cohesive Soil Utilizing Meta Models. Iranian Journal of Irrigation and Drainage, 19(2), 275–290.
[24] Kanishkavardhan, P., Maheswari, N., Nithilan, B., & Aman, A. M. (2026). Scour depth prediction using machine learning and explainable AI: assessment of bridge vulnerability. Frontiers in Built Environment, 12, 1790274. https://doi.org/10.3389/fbuil.2026.1790274.
[25] Ghumman, A. R., Haider, H., Al Salamah, I. S., Shafiquzzaman, M., Alodah, A., Alresheedi, M., et al. (2024). Local scour depth around bridge piers: performance evaluation of dimensional analysis-based empirical equations and AI techniques. KSCE Journal of Civil Engineering, 28, 3270–3285. https://doi.org/10.1007/s12205-024-1161-x
[26] Kim, T., Shahriar, A. R., Lee, W.-D., & Gabr, M. A. (2024). Interpretable machine learning scheme for predicting bridge pier scour depth. Computers and Geotechnics, 170, 106302. https://doi.org/10.1016/j.compgeo.2024.106302
[27] Kumar, S., Goyal, M. K., Deshpande, V., & Agarwal, M. (2023). Estimation of time dependent scour depth around circular bridge piers: Application of ensemble machine learning methods. Ocean Engineering, 270, 113611. https://doi.org/10.1016/j.oceaneng.2022.113611
[28] Eini, N., Bateni, S. M., Jun, C., Heggy, E., & Band, S. S. (2023). Estimation and interpretation of equilibrium scour depth around circular bridge piers by using optimized XGBoost and SHAP. Engineering Applications of Computational Fluid Mechanics, 17(1), 2244558. https://doi.org/10.1080/19942060.2023.2244558
[29] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30, 4765–4774.
[30] Choi, S. U., Choi, B., & Choi, S. (2015). Improving predictions made by ANN model using data quality assessment: An application to local scour around bridge piers. Journal of Hydroinformatics, 17(6), 977–989. https://doi.org/10.2166/hydro.2015.097
[31] Debnath, K., & Chaudhuri, S. (2010). Bridge pier scour in clay–sand mixed sediments at near-threshold velocity for sand. Journal of Hydraulic Engineering, 136(9), 597–609. https://doi.org/10.1061/(ASCE)HY.1943-7900.0000221
[32] Ettema, R., Kirkil, G., & Muste, M. (2006). Similitude of large-scale turbulence in experiments on local scour at cylinders. Journal of Hydraulic Engineering, 132(1), 33–40. https://doi.org/10.1061/(ASCE)0733-9429(2006)132:1(33)
[33] Gudavalli, S. R. (1997). Prediction model for scour rate around bridge piers in cohesive soil on the basis of flume tests. PhD Thesis, Texas A&M University, College Station, TX, USA.