International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 19s, 2025
https://theaspd.com/index.php

Application Of Boosting And Bagging Algorithms In Predicting Bridge Scour Under Clear-Water Conditions

Vanshika Bhardwaj[1], Har Amrit Singh Sandhu[2], and Baldev Setia[3]
[1,2]Department of Civil Engineering, Punjab Engineering College, Chandigarh, India
[3]Department of Civil Engineering, National Institute of Technology, Kurukshetra, Haryana, India
vanshikabhardwaj.phd23civil@pec.edu.in[1], hassandhu@pec.edu.in[2], and setia@nitkkr.ac.in[3]

Abstract: *Accurate prediction of scour depth around bridge abutments is critical for ensuring the safety and longevity of hydraulic structures, especially under clear-water conditions. This study investigated the predictive performance of two ensemble machine learning models including Extreme Gradient Boosting (XGBoost) and Random Forests (RF) to estimate scour depth for three distinct abutment geometries: Vertical Wall, 45° Wing Wall, and Semicircular Abutments. A laboratory-generated experimental dataset was employed with a 70:30 split for training and testing respectively. The results demonstrated that XGBoost consistently outperformed RF across all abutment types achieving higher determination coefficients (R^2) and lower error metrics (Root Mean Square Error and Mean Absolute Error). It was observed that XGBoost achieved an R^2 of 0.9707 for Vertical Wall abutments, compared to 0.8721 by RF. The superior performance of XGBoost is attributed to its gradient-boosting framework and regularization capabilities, which enhance its generalization ability on complex, nonlinear datasets. This study confirms the effectiveness of XGBoost as a reliable and accurate tool for predicting scour depth, outperforming traditional ensemble methods. The findings highlight the potential of advanced machine learning approaches in improving hydraulic design and risk assessment practices.*

Keywords: *Sour Depth, Machine Learning, XGBoost, Random Forests*

INTRODUCTION:

Scouring is the process of removal of sediments from around a hydraulic structure. This process is a concern for hydraulic engineers, especially around bridge piers, abutments, or embankments, because it compromises structural stability. The scour is mainly of three types: General scour, contraction scour, and local scour. The general scour occurs due to changes in river bed over a length of river, thus independent of the presence of structure across the river, whereas, on the other hand, the contraction and local scour are dependent on the presence of hydraulic structure across the river. The contraction scour occurs when the flow area is reduced, leading to increased flow velocity causing erosion. The local scour occurs around specific structures like piers and abutments due to the formation of vortices. Among all three, local scour is the most critical and thus is studied widely.

The mechanism of scour is a complex process caused due to flow obstructed by the pier or abutment on the upstream side causing downflow. This downflow hits the sediment bed with force, causing removal of sediment leading to formation of vortex. The vortex formed has a characteristic shape of a horseshoe, thus named horseshow vortex. This horseshoe vortex wraps around the base of the pier or abutment and transports the sediment away from the foundation. According to Melville and Coleman (2000), the horseshoe vortex system is the most significant factor contributing to development of local scour at the front of the pier exerting high shear stress on the bed material. In addition to the horseshoe vortex, secondary vortex called wake vortex is form around the pier. This wake vortex acts like a vacuum pump that carries the sediment from around the pier or abutment to downstream side, further eroding the sediment bed. These vortices fluctuate in time and space, thus, increasing the turbulence intensity and contributing to sediment entrainment behind the pier (Ettema et al.,1998). The mechanism of scour is influenced by various factors like flow velocity, pier shape, orientation of pier, size of sediment, gradation of sediment, depth of flow, and composition of bed. Chiew and Melville (1987) showed through laboratory studies that sharper pier shapes generate stronger vortices and deeper scour holes. The flow separation points are more pronounced for non-cylindrical piers, causing enhanced turbulence and erosion of sediment. Dey and Raikar (2007) also emphasized that equilibrium scour depth depend on the sediment transport capacity and critical shear stress of the bed material. For cohesive sediments, time and cohesion plays important role in resisting scour. Thus, the interaction of hydrodynamic forces and sediment resistance governs the scour depth and extent of scour highlighting the importance of predicting scour depth.

Many studies have contributed to understand the local scour depth around bridge abutments through experimental, analytical, and data-driven approaches. Laursen and Toch (1956) conducted foundational laboratory experiments on local scour around abutments which was limited because of the absence of flow visualization tools. Melville (1992) overcame this limitation by expanding the experimental dataset by varying flow depth, abutment geometry, and sediment properties, while, Kwan and Melville (1994) highlighted the significance of vortex formed in driving scour development. Dey and Barbhuiya (2004) studied scour under clear water conditions for various abutment geometries and proposed empirical relationships for equilibrium scour depth. Similarly, Fael et al. (2006) applied regression techniques to describe scour morphology and Abou-Seida et al. (2012) developed predictive equations for scour depth in cohesive sediment bed. Barbhuiya and Mazumder (2014) experimented behavior of scour for varying sediment size and abutment geometries resulting in a refined empirical model for predicting scour depth. A study conducted by Singh et al. (2020) introduced analytical expressions for time-dependent and equilibrium scour depth under short contractions. The physical models have advanced understanding, but they are inherently time consuming and resource intensive that lacks scalability to complex, real-world conditions. Numerical models (Afzal, 2013; Ahmad et al., 2015; Gautam et al., 2021) have provided a detailed understanding of sediment transport and scour process, but requires significant computational demand. Due to these limitations, Machine Learning (ML) approaches have emerged as efficient data-driven alternatives that have the capability of handling non-linearity and large datasets.

Advancements in ML models in predicting scour depth have significant increase. Muzammil (2008) applied artificial neural networks (ANN), which highlights the performance of ANN over traditional empirical models. Muzzammil (2010) also depicted the same results using adaptive-network-based fuzzy inference systems (ANFIS). Azamathulla et al (2010) employed gene expression programming (GEP) that resulted in better efficiency than ANN and regression models. Najafzadeh et al. (2013a, 2013b) introduced group method of data handling (GMDH) and found it effective in case of both clear-water and live-bed scour conditions. Recent introduction of hybrid and optimized ML approaches have made a significant breakthrough. Azimi et al. (2017,2019) developed a Pareto-evolutionary ANFIS model, while, Parsaie et al. (2019) an Ebtehaj et al. (2018) showed a better performance of support vector machine (SVM) and extreme learning machine (ELM) in terms of accuracy and computational efficiency. Bonakdari et al. (2020) reported better performance of ELM in predicting clear-water scour. Genetic algorithms (Pandey et al., 2020), firefly and grasshopper optimization (Kaveh et al., 2021; Kohansarbaz et al., 2021) and hybrid models like RC-Kstar and WIHW-Kstar (Khosravi et al., 2021) enhanced prediction accuracy of scour depth around different abutment geometries.

This study employs ensembled ML models including Extreme Gradient Boosting (XGBoost) and Random Forests (RF) for predicting scour depth around various abutment geometries under clear-water conditions using MATLAB. These ML models were selected because of their advantage in handling non-linear relationships, high dimensional data, and overfitting issues due to hydrodynamic properties of scouring. To evaluate their predictive performance, statistical indices including coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE) were used. The results provide a comparative analysis of the models employed highlighting their strengths and limitations in predicting scour depth.

METHODOLOGY:

In the present study, the data for scour depth prediction under unsteady flow conditions were collected from published datasets from Dey and Barbhuiya (2005) (Table 1), for predicting scour depth using ML models such as XGBoost and RF. The datasets were divided in 70/30 ratio, where 70 % is used as training and 30 % for testing the models. The efficiency of the models developed was analysed using statistical error indices such as Root Mean Square Error (RMSE), Coefficient of Determination (R^2), and, Mean Absolute Error (MAE) (Figure 2). The model with higher R^2 value and lower RMSE and MAPE values is considered to be thebest (Kumar et al., 2023).

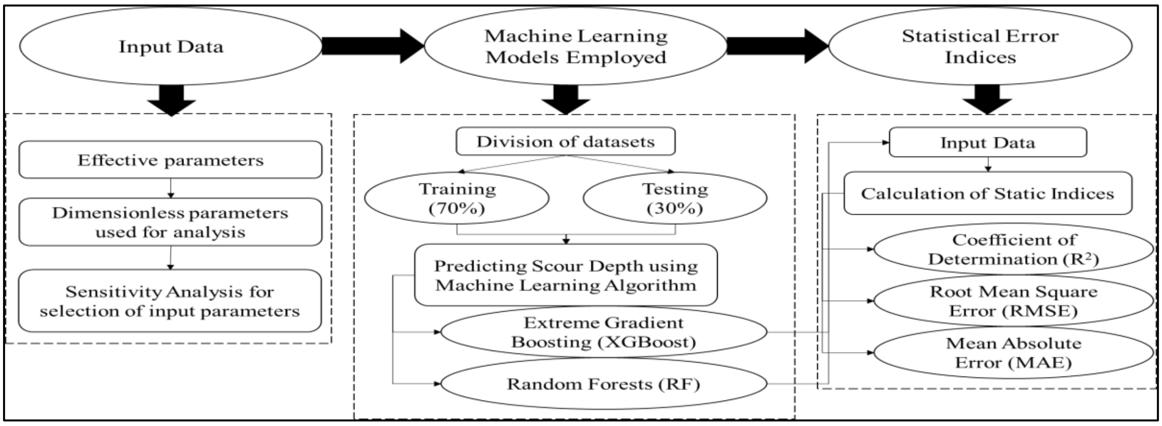

Figure 1. Flowchart of Methodology

Data Used:

A total of 297 experimental dataset used in this study was obtained from Dey and Barbhuiya (2005) which predicted scour depth under clear-water conditions for three different abutment geometry including vertical wall, 45° wing wall, and semicircular abutment (Table 1). The dataset includes hydraulic and geometric parameters like mean sediment size (d_{50}), flow depth (h), approach flow velocity (U), transverse abutment length (L), width of abutment across the channel (b), Froude number (F_r), and geometric standard deviation (σ_g). They measured scour depth (d_s) using Vernier point gauge with an accuracy of ±0.1mm.

Table 1. Dataset obtained from Dey and Barbhuiya (2005):

Parameters	Vertical Wall		45° Wing Wall		Semicircular Abutment	
	Max	Min	Max	Min	Max	Min
d_{50} (mm)	3.1	0.26	3.1	0.26	3.1	0.26
h (m)	0.25	0.058	0.25	0.059	0.25	0.059
U (m/s)	0.67	0.219	0.67	0.22	0.67	0.22
L (cm)	12	4	12	4	13	4
b (cm)	24	8	36	12	26	8
F_r	0.8327	0.1571	0.8327	0.1571	0.8327	0.1571
σ_g	1.38	1.17	1.38	1.17	1.38	1.17
d_s (m)	0.293	0.068	0.278	0.053	0.28	0.055

Sensitivity Analysis:

For performing sensitivity analysis on the dataset, Principal Component Regression (PCR) was employed to address the multicollinearity among the input variables to predict scour depth. The analysis retained three principal components which altogether explains a substantial proportion of variance in the dataset. The resulting PCR model highlights the strong predictive capability with a R^2 of 0.9539 and a low RMSE of 0.1415, which shows high model accuracy. The first principal component (PC1) captured the largest proportion of variance that was predominantly influenced by the excess abutment Froude number (F_r) and the relative sediment size (d_{50}/L), both contributing significantly and positively. The geometric standard deviation (σ_g) exhibited a strong negative loading on PC1 suggesting that greater sediment heterogeneity tends to reduce scour depth. In contrast, the effect of relative flow depth (h/L) and sediment-size-to-depth ratio (d_{50}/h) were relatively less. These findings reinforce the major effect of F_r and d_{50}/L in scour processes (Khosravi et al., 2021; Kumar et al., 2023), thus validating PCR as a better method for sensitivity analysis and prediction in hydrodynamics case where multicollinearity among predictors is present.

Extreme Gradient Boosting (XGBoost):

Extreme Gradient Boosting (XGBoost) is an ensemble machine learning algorithm based on decision tree boosting due to its scalability and efficiency in prediction. It operates by sequentially building an ensemble regression trees where every new tree attempts to correct the errors from previous iterations using a gradient descent optimization algorithm. In the context of hydraulic and geotechnical engineering, XGBoost has shown promising results in modeling complex nonlinear phenomena, such as scour depth around bridge piers, due to its ability to handle multicollinearity, missing values, and noisy data (Chen and Guestrin, 2016). Recent studies have demonstrated its superiority over traditional regression models and basic machine learning techniques in predicting scour depth under varying flow conditions (Khosravi et al., 2021; Wang et al., 2023). The algorithm's regularization mechanisms help prevent overfitting, making it particularly suitable for datasets with limited observations but high dimensionality, such as those derived from flume experiments or field measurements of pier scour.

Random Forests (RF):

Random Forest (RF) is a widely used ensemble learning method that builds multiple decision trees using random samples of the data and a random selection of features. These trees then work together, with their results combined either by averaging (for regression) or voting (for classification) to produce a more accurate prediction. Originally proposed by Breiman in 2001, RF is particularly effective in tackling complex, high-dimensional problems and is known for its strong resistance to overfitting. In hydraulic modeling and scour depth prediction, RF has gained attention for its ability to manage noisy datasets, handle multicollinearity, and capture nonlinear relationships between variables without the need for extensive preprocessing. Its effectiveness has been demonstrated in various water resources applications, such as modeling sediment transport, estimating flow discharge, and predicting scour depth around bridge piers (Pal et al., 2013; Habib et al., 2024). Overall, compared to traditional models and single decision trees, Random Forest consistently delivers higher accuracy and more reliable results, making it a strong choice for data-driven scour prediction in both clear-water and live-bed scenarios.

RESULTS AND DISCUSSION:

In this study, the predictive capabilities of two ensemble machine learning models (Extreme Gradient Boosting (XGBoost) and Random Forest (RF)), were evaluated for estimating clear-water scour depth around three abutment geometries i.e. Vertical Wall, 45° Wing Wall, and Semicircular Abutments. A 70:30 split was adopted for training and testing the dataset, ensuring sufficient data for learning while preserving a meaningful validation subset for assessing generalization performance. As presented in Table 2, XGBoost demonstrated superior performance across all abutment types when compared to RF. For the Vertical Wall abutment, XGBoost achieved a high coefficient of determination (R^2) of 0.9707, with low error metrics (RMSE = 0.1821, MAE = 0.1362). In comparison, RF showed a significant drop in predictive quality (R^2 = 0.8721, RMSE = 0.3850, MAE = 0.2524). This large difference highlights the stronger modeling capacity of XGBoost for cases with complex hydraulic behavior and more distinct scour patterns.

For the 45° Wing Wall, the trend remained consistent. XGBoost recorded an R^2 of 0.9293, while RF achieved 0.9186. Although the difference in R^2 here was smaller, XGBoost still outperformed RF in terms of both RMSE (0.1023 vs. 0.2491) and MAE (0.0815 vs. 0.2029), indicating more precise and reliable predictions. The geometry of the 45° Wing Wall may introduce angled flow separation and vortex interactions, which are more effectively captured by the sequential tree structure and boosting logic of XGBoost.

In the case of Semicircular Abutments, XGBoost again led with R^2 = 0.9307, outperforming RF's 0.8689. While RF still showed reasonable performance, its higher RMSE (0.2029) and MAE (0.2381) compared to XGBoost's (RMSE = 0.1812, MAE = 0.1325) reinforce the notion that RF may struggle to model more subtle, nonlinear interactions such as those induced by curvature-related flow patterns around semicircular abutments. These findings align with several recent studies that have demonstrated XGBoost's advantages over other ensemble models in hydraulic and geotechnical applications. For example, Chen and Guestrin (2016) emphasized the strength of XGBoost's gradient boosting framework with built-in regularization, which prevents overfitting and efficiently handles complex feature interactions. Similarly, Habib et al. (2024) applied XGBoost for sediment transport and scour modeling, achieving higher accuracy and generalizability compared to Support Vector Machines and Random Forests.

Random Forest, though slightly less accurate in this context, remains a valuable tool, particularly for its robustness, low sensitivity to overfitting, and ease of interpretation. Introduced by Breiman (2001), RF is widely used in hydrologic modeling and environmental engineering for tasks involving uncertainty and noisy data. Studies such as Pal et al. (2013) have shown its effectiveness in land-use classification and runoff estimation. However, RF's ensemble averaging nature often causes it to miss finer, compound interactions between input variables, an area where boosting algorithms like XGBoost excel. Furthermore, XGBoost's iterative learning approach prioritizes correcting errors from previous trees, which is particularly useful in datasets like this one, where scour depth is influenced by multiple interrelated hydraulic variables (velocity, flow depth, sediment size). RF, by contrast, builds all trees independently and then averages their outputs, making it more resistant to noise but potentially less responsive to subtle patterns in the data.

From a practical standpoint, the better performance of XGBoost suggests its greater potential in real-world applications of scour prediction, especially where field data may be limited or exhibit nonlinear behaviors. For engineers and decision-makers, such precision can translate into more reliable infrastructure design and risk mitigation strategies in flood-prone or erosion-susceptible regions. In summary, XGBoost not only provided more accurate predictions of scour depth but also displayed stronger generalization across varying abutment geometries. Its ability to adaptively learn complex relationships within the dataset makes it a more suitable and advanced tool for modeling clear-water scour scenarios in hydraulic engineering applications.

Table 2. Summary of statistical indices of ML models:

ML models	**XGBoost**			**RF**		
	Vertical Wall	45° Wing Wall	Semicircular Abutments	Vertical Wall	45° Wing Wall	Semicircular Abutments
R^2	0.9707	0.9293	0.9307	0.8721	0.9186	0.8689
RMSE	0.1821	0.1023	0.1812	0.3850	0.2491	0.2029
MAE	0.1362	0.0815	0.1325	0.2524	0.2029	0.2381

CONCLUSION:

This study set out to evaluate how well two powerful ensemble machine learning models could predict scour depth under clear-water conditions around different bridge abutment shapes. The results showed that XGBoost consistently delivered better predictions than Random Forest across all three abutment types. It achieved higher R^2 scores and lower error values (RMSE and MAE), showing that it was better at capturing the complex, nonlinear relationships between hydraulic parameters and scour depth. This makes sense, given that XGBoost is designed to improve predictions step by step while also controlling for overfitting. Random Forest also performed reasonably well but was less accurate, especially in the case of vertical wall abutments. Its averaging approach across many decision trees works well for general patterns, but it can sometimes miss the fine details that XGBoost is able to pick up. In short, XGBoost proved to be a more reliable and precise tool for estimating scour depth under varying geometric conditions. Its performance highlights the potential of modern machine learning techniques in solving real-world hydraulic engineering problems. Going forward, incorporating more diverse datasets or trying out hybrid or deep learning models could further improve accuracy and support better-informed design decisions in bridge and river engineering.

REFERENCES:

1. Melville, B. W., and Coleman, S. E. (2000). Bridge scour. Water Resources Publication.
2. Ettema, R., Melville, B. W., and Barkdoll, B. (1998). Scale effect in pier-scour experiments. Journal of Hydraulic Engineering, 124(6), 639-642.
3. Chiew, Y. M., and Melville, B. W. (1987). Local scour around bridge piers. Journal of hydraulic research, 25(1), 15-26.
4. Dey, S., and Raikar, R. V. (2007). Clear-water scour at piers in sand beds with an armor layer of gravels. Journal of Hydraulic Engineering, 133(6), 703-711.
5. Laursen, E. M. and Toch, A. (1956). Scour Around Bridge Piers and Abutments. Iowa Highway Research, Board Ames, IA.
6. Melville, B. W. (1992). Local scour at bridge abutments. Journal of Hydraulic Engineering 118 (4), 615–631.

7. Kwan, R. T. F. and Melville, B. W. (1994). Local scour and flow measurements at bridge abutments. Journal of Hydraulic Research 32 (5), 661–673.
8. Dey, S. and Barbhuiya, A. K. (2004). Clear water scour at abutments. Proceedings of the Institution of Civil Engineers: Water Management 157 (2), 77–97.
9. Fael, C. M. S., Simarro-Grande, G., Mart'in-Vide, J.-P. and Cardoso, A. H. (2006). Local scour at vertical-wall abutments under clear-water flow conditions. Water Resources Research 42 (10), 1–12.
10. Abou-Seida, M. M., Elsaeed, G. H., Mostafa, T. M. and Elzahry, E. F. (2012). Local scour at bridge abutments in cohesive soil. Journal of Hydraulic Research Taylor and Francis 50 (2), 171–180.
11. Barbhuiya, A. K. and Mazumder, M. H. (2014). Live-bed local scour around vertical-wall abutments. ISH Journal of Hydraulic Engineering 20 (3), 339–351. Taylor and Francis.
12. Singh, R. K., Pandey, M., Pu, J. H., Pasupuleti, S. and Villuri, V. G. K. (2020). Experimental study of clear-water contraction scour. Water Supply 20 (3), 943–952. IWA Publishing.
13. Afzal, M. S. (2013). 3D Numerical Modelling of Sediment Transport Under Current and Waves. Norwegian University of Science and Technology, Høgskoleringen 7A, Trondheim, Norway.
14. Ahmad, N., Afzal, S., Bihs, H. and Arntsen, Ø. A. (2015). Three-dimensional numerical modeling of local scour around a non-slender cylinder under varying wave conditions. In 36th IAHR World Congress, Hague, The Netherlands.
15. Gautam, S., Dutta, D., Bihs, H. and Afzal, M. S. (2021). Three-dimensional computational fluid dynamics modelling of scour around a single pile due to combined action of the waves and current using level-set method. Coastal Engineering 170, 104002.
16. Muzzammil, M. 2008 Application of neural networks to scour depth prediction at the bridge abutments. Engineering Applications of Computational Fluid Mechanics 2 (1), 30–40. Taylor and Francis
17. Muzzammil, M. 2010 ANFIS approach to the scour depth prediction at a bridge abutment. Journal of Hydroinformatics 12 (4), 474–485. IWA Publishing
18. Azamathulla, H. M., Ghani, A. A., Zakaria, N. A. and Guven, A. 2010 Genetic programming to predict bridge pier scour. Journal of Hydraulic Engineering 136 (3), 165
19. Najafzadeh, M., Barani, G.-A. and Kermani, M. R. H. 2013a GMDH based back propagation algorithm to predict abutment scour in cohesive soils. Ocean Engineering 59, 100–106. Elsevier.
20. Najafzadeh, M., Barani, G.-A. and Kermani, M. R. H. 2013b Abutment scour in clear-water and live-bed conditions by GMDH network. Water Science and Technology 67 (5), 1121–1128. IWA Publishing.
21. Azimi, H., Bonakdari, H., Ebtehaj, I., Talesh, S. H. A., Michelson, D. G. and Jamali, A. 2017 Evolutionary Pareto optimization of an ANFIS network for modeling scour at pile groups in clear water condition. Fuzzy Sets and Systems 319, 50–69.
22. Azimi, H., Bonakdari, H., Ebtehaj, I., Shabanlou, S., Ashraf Talesh, S. H. and Jamali, A. 2019 A Pareto design of evolutionary hybrid optimization of ANFIS model in prediction abutment scour depth. Sadhna Springer 44, 1–14.
23. Parsaie, A., Haghiabi, A. H. and Moradinejad, A. 2019 Prediction of scour depth below river pipeline using support vector machine. KSCE Journal of Civil Engineering Springer 23, 2503–2513.
24. Ebtehaj, I., Bonakdari, H., Moradi, F., Gharabaghi, B. and Khozani, Z. S. 2018 An integrated framework of extreme learning machines for predicting scour at pile groups in clear water condition. Coastal Engineering 135, 1–15.
25. Bonakdari, H., Moradi, F., Ebtehaj, I., Gharabaghi, B., Sattar, A. A., Azimi, A. H. and Radecki-Pawlik, A. 2020 A non-tuned machine learning technique for abutment scour depth in clear water condition. Water 12 (1), 301
26. Pandey, M., Zakwan, M., Khan, M. A. and Bhave, S. 2020 Development of scour around a circular pier and its modelling using genetic algorithm. Water Supply 20 (8), 3358–3367.
27. Kaveh, K., Mai, D. N., Pham, Q. B. and Anh, D. T. 2021 A hybrid feed-forward neural network with grasshopper optimization for observing pattern of scour depth around bridge piers. Arabian Journal of Geosciences 14, 1–11.
28. Kohansarbaz, A., Kohansarbaz, A., Yaghoubi, B., Izadbakhsh, M. A. and Shabanlou, S. 2021 An integration of adaptive neuro-fuzzy inference system and firefly algorithm for scour estimation near bridge piers. Earth Science Informatics 14 (3), 1399–1411.
29. Khosravi, K., Khozani, Z. S. and Mao, L. 2021 A comparison between advanced hybrid machine learning algorithms and empirical equations applied to abutment scour depth prediction. Journal of Hydrology Elsevier 596, 126100.
30. Kumar, A., Baranwal, A., and Das, B. S. (2023). Modelling of clear water scour depth around bridge piers using M5 tree and ANN-PSO. AQUA—Water Infrastructure, Ecosystems and Society, 72(8), 1386-1403.
31. Kumar, L., Afzal, M. S., and Ghosh, S. (2023). A novel-tuned Custom ensemble machine learning model to predict abutment scour depth in clear water conditions. AQUA—Water Infrastructure, Ecosystems and Society, 72(5), 798-813.
32. Chen, T., and Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794).
33. Wang, T., Reiffsteck, P., Chevalier, C., Zhu, Z., Chen, C. W., and Schmidt, F. (2023). A novel extreme gradient boosting algorithm-based model for predicting the scour risk around bridge piers: application to French railway bridges. European Journal of Environmental and Civil Engineering, 27(3), 1104-1122.
34. Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.
35. Habib, M. A., Abolfathi, S., O'Sullivan, J. J., and Salauddin, M. (2024). Efficient data-driven machine learning models for scour depth predictions at sloping sea defences. Frontiers in Built Environment, 10, 1343398.
36. Pal, M., Singh, N. K., and Tiwari, N. K. (2013). Pier scour modelling using random forest regression. ISH Journal of Hydraulic Engineering, 19(2), 69-75.