Contents lists available at SCCE

Journal of Soft Computing in Civil Engineering

Journal homepage: www.jsoftcivil.com

An Optimized Stacking Machine Learning Estimator for Sediment Transport Estimation

Hojjat Emami [1,*]; Somayeh Emami [2]; Mohammed Achite [3,4]; Aseel Smerat [5,6]

1. Associate Professor, Computer Engineering Department, Faculty of Engineering, University of Bonab, Bonab, Iran
2. Ph.D., Department of Water Science and Engineering, University of Tabriz, Tabriz, Iran
3. Professor, Laboratory of Water and Environment, Faculty of Nature and Life Sciences, Hassiba Benbouali University of Chlef, 02180 Chlef, Algeria
4. University of Oran 2 Mohamed Ben Ahmed, Georessources, Environment, and Natural Risks Laboratory, Oran, Algeria
5. Faculty of Educational Sciences, Al-Ahliyya Amman University, Amman 19328, Jordan
6. Centre for Research Impact and Outcome, Chitkara University, Punjab, India

* Corresponding author: *emami@ubonab.ac.ir*

https://doi.org/10.22115/SCCE.2025.2151

ARTICLE INFO

Article history:
Received: 25 August 2025
Revised: 21 September 2025
Accepted: 07 October 2025

Keywords:
Sediment transport estimation;
Ensemble learning;
Stacking model;
Data analysis.

ABSTRACT

Accurate estimation of sediment transport is essential for effective water resources management, as sedimentation can decrease dam reservoir capacity, impact water quality, and cause significant changes to river ecosystems. This study introduces a two-level optimized ensemble machine learning model to estimate sediment transport at Algeria's Oued El Abtal station. The proposed model combines four machine learning algorithms—random forest regression, categorical boosting, gradient boosting regression, and k-nearest neighbor—as base learners, with ridge regression serving as a meta-learner to improve accuracy. Multivariate inputs, including height, instantaneous concentration, and water discharge, are used to estimate sediment discharge. Testing on a dataset of 8,792 hydrological records shows that the stacking model significantly outperforms individual learners and alternative methods. The model achieved superior results on the test data, with an R^2 of 0.973, an RMSE of 671.147 kg/s, and an MAE of 68.887 kg/s. These results demonstrate a clear advantage over the next best method (DNN), which had an R^2 of 1.00 and an MAE of 102.175 kg/s. This research highlights the potential of stacking ensembles to accurately estimate sediment transport, providing a dependable method for sustainable water management.

How to cite this article: Emami H, Emami S, Achite M, Smerat A. An Optimized Stacking Machine Learning Estimator for Sediment Transport Estimation. Journal of Soft Computing in Civil Engineering. 2026;10(3):2151. https://doi.org/10.22115/scce.2025.2151

1. Introduction

Many water resource challenges, including dam reservoir design, river restoration, stable channel design, fish and wildlife habitat conservation, watershed management, and environmental impact assessment, require an estimate and simulation of sediment transport. Direct measurement is known as the most accurate strategy for computing sediment load; however, it is time-consuming and does not allow for frequent sampling [1]. Furthermore, enhancing river modeling can be greatly aided by precise assessment and forecasting of the river sediment load. Sediment load prediction is necessary for addressing various issues, including river sediment transport, dam dead storage design, stable channel design, estimation of bridge pier aggradation and degradation, impacts of gravel and sand mining on river-bed equilibrium, environmental impact assessment, and determination of dredging requirements. Additionally, sediment load is a significant contaminant and a transporter of chemicals, nutrients, and insecticides [2].

Numerous studies on the use of artificial neural networks (ANNs) in sediment yield forecasts and simulations have been presented in recent decades. ANNs are well-known intelligent models with extensive application in different domain [3,4]. An integrated stage–discharge–sediment concentration relation is established by Jain [5] for two Mississippi River stations using the ANN technique. This study demonstrates that the ANN produces superior results compared to the traditional method. In the north of England, Cigizoglu [6] used ANNs to measure suspended sediment concentrations and compared ANNs with sediment rating curves for two rivers with remarkably similar catchment regions and features. The results of the ANN model for non-steady state sheet sediment transport were compared with those of physically based models by Tayfur [7]. To predict the daily suspended sediment concentration, Kisi (2004) employed a variety of ANN methodologies. The results showed that the multi-layer perceptron (MLP) outperformed the counterpart models. Cigizoglu and Alp [8] used generalized regression neural networks (GRNN) and feed-forward back propagation (FFBP) to predict the yield of river sediment. Tayfur and Guldal [9] used ANNs and a non-linear black box model based on two-dimensional unit sediment graph theory (2DUSGT) to measure the daily total suspended sediment in natural rivers based on precipitation data. For a small agricultural watershed, Raghuwanshi et al. [10] created ANN models to forecast sediment output and runoff on a daily and weekly basis. Regression models were also used to forecast sediment output and runoff on a daily and weekly basis. The ANN outperformed the regression-based models in every instance. Alp and Cigizoglu [11] used two different neural network types— MLP and radial basis function (RBF)—to model suspended sediment load.

Melesse et al. [12] utilized ANNs with back propagation training algorithms to forecast the suspended sediment load. They then examined the performance of the model with three alternative methods: multiple linear regressions (MLR), auto-regressive integrated moving average (ARIMA), and multiple non-linear regressions (MNLR). The suspended sediment discharge of the Pari River near Silibin in Peninsular Malaysia was forecasted by Mustafa et al. [13,14] using an MLP feed-forward neural network with four distinct training procedures. Panahi et al. [15] used the support vector machine (SVM) and adaptive neuro-fuzzy interface system (ANFIS) to measure the sediment transport of two main branches of the Telar River, located in the north of Iran. To improve the accuracy of the SVM and ANFIS, they used the black widow optimization (BWO) algorithm. Among the models in the Telar River, it was determined that the ANFIS-BWO generated the highest determination coefficient (R^2). In the testing level for the Kasilian River, the MAE of the ANFIS-BWO is 899.12 (Ton/day). Emami et al. [1] addressed the task of estimating sediment transport in the Zarinehrood River. To achieve this, researchers developed a hybrid model combining the whale optimization algorithm (WOA) with an ANFIS. Using a decade of flow discharge data as input, the WOA-ANFIS model demonstrated exceptionally high accuracy, with an R^2 of 0.962 and a low error rate (NRMSE=0.051), proving it to be a highly capable method for this

application. For the Brahmani River basin in Odisha State, India, Samantaray et al. [16] developed a robust model that computes suspended sediment load at the Jenapur, Jaraikela, Gomlai, and Tilga stations using the SVM and a unique sparrow search algorithm (SVM-SSA). The creation of the model takes into account five distinct scenarios. Using the R^2, mean absolute error (MAE), root mean squared error (RMSE), and Nash–Sutcliffe efficiency (ENS), the performance of models is evaluated. Three hybrid models, including the SVM-grasshopper optimization algorithm (SVM-GOA), SVM-Bat algorithm (SVM-BA), and SVM-butterfly optimization algorithm (SVM-BOA), are used to compare the results of the SVM-SSA model. With ENS = 0.96481, MAE = 15.3926 kg/s, and RMSE = 15.5287 kg/s, the results showed that the SVM-SSA model outperforms its counterparts in accurately estimating suspended sediment load. Bezak et al. [17] examined two models, including eXtreme gradient boosting (XGBoost) and a sediment rating curve (SRC), for forecasting the suspended sediment load in the Slovenian Sora River basin. The models were evaluated using historical data from 2016 to 2021, and the results showed that XGBoost outperformed the SRC and produced a smaller bias (about 15%).

Kundu et al. [18] examined machine learning (ML) models for sediment load prediction in the Godavari River Basin. When the dataset was separated into pre-1990 (1969–1990) and post-1990 (1990–2020) periods, it was found that anthropogenic factors, including the dam construction and changes in land-use/land-cover (LULC), significantly reduced the mean annual sediment load from 136.85 to 62.38 million tons after 1990. So, Tree-based learning models were created and assessed using R^2, RMSE, and MAE metrics to enhance sediment load prediction. Though the gradient boosting regressor (GBR) and random forest regressor (RFR) produced competitive results, the extra trees regressor (ETR) showed the highest prediction accuracy with the lowest errors. The results show how human changes affect sediment transport and that ensemble tree-based models provide a reliable way to anticipate sediment load. Under shifting hydrological conditions, this work offers important insights for modeling sustainable sediment transport and river basin management. In Algeria, in the eastern part, Tamrabet et al [19] showed through a comparative study using ANFIS and ANN models at the five largest basins in eastern Algeria: the Soummam, Kébir-Rhumel, Highlands, Seybouse, and Constantinois Coastal basins, which are characterized by high water erosion. The ANN and ANFIS models showed high accuracy, confirmed by excellent R values ranging from 0.77 to 0.98. The NSE ranged from 0.67 to 0.97. The error values were very good; the MAE varies from 0.004 g/L to 0.028 g/L for both ANFIS and ANN models. For the western part, Achite et al. [20] shown the effectiveness of many ML, particularly deep learning (DL) models for predicting sediment transport in the Mina Basin at two hydrometric stations: Sidi Abdelkader Djillali and Oued Abtal. Among the models are the convolutional neural network (CNN), MLP, deep neural network (DNN), category boosting (CBR), sediment rating curve (SRC), and gated recurrent unit (GRU). The DNN model continuously outperformed the others in this regard. The DNN obtained NSE = 0.99, MAE = 102.17 kg/s, PBIAS = 6.81%, and RMSE = 243.72 kg/s, for the Oued Abtal station. It reported NSE = 0.99, MAE = 46.51 kg/s, PBIAS = 38.06%, and RMSE = 91.27 kg/s at the Sidi Abdelkader Djillali station. According to error analysis, the DNN model provides the most accurate predictions.

ML models, particularly ensemble learning models, have shown remarkable success not only in sediment transport estimation but also in solving a wide range of complex problems across environmental and civil engineering. For instance, in water quality management, Anaraki et al. [21] proposed a stacking ensemble ML to predict rainfall under climate change. Similarly, Kadkhodazadeh et al. [22] introduced a new algorithm, LSSVM-AO-PSO, which combines least squares support vector machine (LSSVM) with hybrid Aquila optimization-particle swarm optimization (AO-PSO). Their model is designed to predict minimum, maximum, and mean monthly air temperatures, considering climate change over time. In more recent work, Farzin et al. [23] proposed a novel approach using a CNN, specifically CNN4, to construct

missing runoff data in Iran by combining basin characteristics, geographical, temporal, and river flow statistics. This CNN-based method, further refined with quantile mapping for bias correction, showed superior performance compared to counterpart models, successfully reconstructing runoff time series even for stations with incomplete data.

After this short review, we conclude that the existing approaches achieved encouraging results; however, their results are not ideal. This issue emphasizes that there is a need to improve the existing methods or propose new methods to bring the performance closer to the ideal state. The goal of this work is to fill this gap by examining several ML algorithms, as the Stacking model integrates four base learners, including random forest regression (RFR), CatBoost (CBR), gradient boosting regression (GBR), and k-nearest neighbors (KNN), and a ridge regression (RR) as a meta-learner to enhance the prediction of the sediment transport in the Wadi Mina basin at the Oued Abtal hydrometric station. In contrast to recent studies that use effective but single-model learners (e.g., DNN [18]) or homogeneous ensembles (e.g., ETR [16]), our proposed model's innovation lies in its strategic integration of diverse base learners. This strategy allows the stacking model to capture a set of underlying patterns in the hydraulic data. The proposed model enhances estimation performance through combining these diverse predictions, where each model identifies specific patterns, and their intelligent integration by an optimized meta-learner improves overall robustness, stability, and performance.

To summarize, our contributions are as follows:

- Modeling the sediment transport prediction as a supervised machine learning task. We proposed a robust two-layer stacking approach to estimate the sediment transport. This approach effectively improves the estimation performance of the individual regression algorithms and outperforms state-of-the-art models in the field.
- Tuning the control parameters of the base and meta-model in the stacking model using the randomized grid search algorithm to boost the estimation performance.
- Conducting extensive experiments to examine the estimation performance of the models. The experimental results clearly show the superiority of the proposed stacking model compared to its counterparts. The proposed model achieved the best results compared to other counterparts by obtaining a value of R^2: 0.973, RMSE: 671.147 kg/s, and MAE: 68.887 kg/s on the test dataset. The results confirmed the stability and robustness of the proposed stacking model.

The rest of the paper is formed as follows. Section 2 describes the study area and its characteristics. Section 3 presents the working principle of the proposed stacking approach and its components. Section 4 is devoted to experiments and discussions. Finally, Section 5 concludes the paper and outlines some future research directions.

2. Methodology

2.1. Study area

Figure 1 shows the study area discussed in this research. The Wadi Mina sub-basin at the hydrométric station of the Wadi Mina basin, located in the North-Western part of Algeria, is the main branch of the Wadi Cheliff, the major river of Algeria. As listed in Table 1, the basin area of Wadi Mina is 4126 km^2, the length of its main stream is 56 km, and the drainage density is 0.42 km km^{-2} (Table 1). The basin supplies the Sidi Mohamed Benaouda (SMB) reservoir, whose basin area is 4900 km^2. The study area has a Mediterranean climate with a distinct seasonal hydrologic imbalance. A protracted dry season is met

with highly limited annual precipitation (~250 mm), resulting in a severe hydraulic deficit due to potential evapotranspiration rates (~1525 mm) that are six times greater.

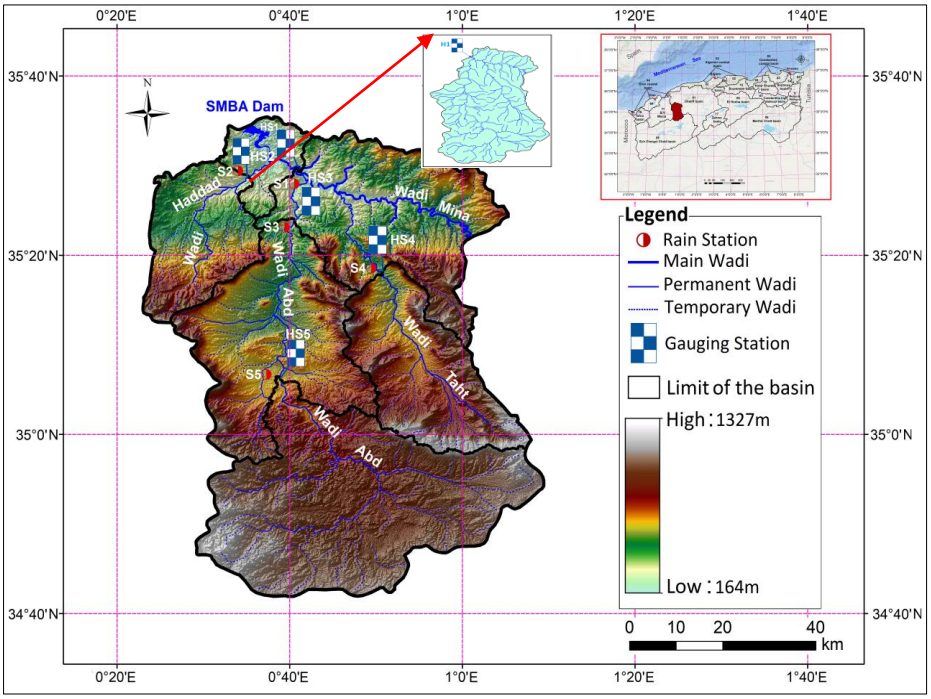


Fig. 1. Study area: Wadi Mina basin at Oued Abtal hydrometric station, northwest Algeria [20].

Table 1
Main characteristics of the investigated sub-basins in the Mina basin.

Type	Parameters	Symbol	Unit	Parameter value
Morphometric characteristics	Surface area	(S)	Km^2	4126.00
	Perimeter	(P)	Km	330.44
	Gravelius compactness index	(Kc)	/	1.44
	Length of equivalent rectangle	(L)	Km	133.03
	Width of equivalent rectangle	(l)	Km	31.01
Topographical characteristics	Maximum altitude	(Hmax)	m	1156.00
	Minimum altitude	(Hmin)	m	230.00
	Average altitude	(Hmean)	m	815.63
	Median altitude	(H50%)	m	755.00
	Altitude at 5% of the surface	(H5%)	m	1200.00
	Altitude at 95% of the surface	(H95%)	m	400.00
	Global slope index	(Ig)	m/km	6.01
	Specific elevation	(Ds)	m	386.27
Hydrological characteristics	Drainage density	(Dd)	Km/Km^2	0.42
	Length of the main wadi	(L_{mw})	Km	56.00
	Concentration time	(Tc)	Hours	17.61

2.2. The proposed model

Figure 2 demonstrates the working principles of the proposed approach. Five main steps in the proposed framework are as follows:

- **Pre-processing:** Before dividing the data into training and testing sets, three preprocessing steps were applied: handling missing values by replacing them with the mean values of each feature, and normalization to guarantee that the features are on a comparable scale. The dataset was randomly split into two sets: 70% for training and 30% for testing.
- **Model training:** ML models are trained using a training set. In each model, the input training data is randomly partitioned into five folds, where four folds are considered for training and one fold is considered as validation data.
- **Validation:** Machine learning models are evaluated on the testing set using performance indexes.
- **Prediction:** In this phase, the model receives the testing data, processes it, and generates the outcomes.
- **Evaluation:** The performance of the ML models is examined using performance metrics, and the performance indexes are reported.

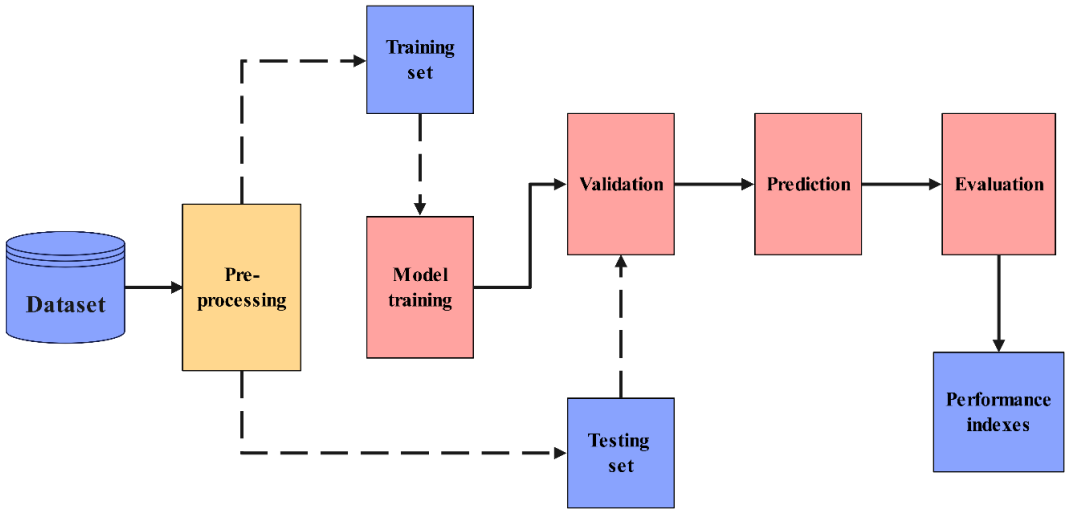


Fig. 2. The general structure of the proposed approach.

The main motivation of the proposed stacking model is to combine the predictions of multiple machine learning models to improve performance indexes. The stacking model is composed of five learners (four base models and one meta-model) that are integrated in a two-level structure. The base learners are composed of four algorithms, as follows.

- **Category boosting regression (CBR)** [24] is a tree-based model that effectively handles categorical features. It uses gradient boosting to integrate weak models and uses regularization methods, like data permutation, to mitigate overfitting. It also increases prediction accuracy and speed with a symmetric tree structure and robustness to noise and outliers.
- **Gradient boosting regression (GBR)** [25] is a boosting tree-based ML model that covers complex nonlinear relations. It gradually improves the prediction accuracy by iteratively combining decision trees and focusing on reducing the residual errors of previous trees. GBR is very effective for managing a variety of data and reducing overfitting.
- **Random forest regression (RFR)** [26] is a bagging tree-based algorithm that increases prediction accuracy by combining the outcomes of multiple decision trees built from random subsets of data and features. It is suitable for complex and high-dimensional data sets by reducing overfitting and providing insight into the importance of features.
- **K-nearest neighbor (KNN)** [27] is a non-parametric method that identifies the *K* nearest neighbors to a new data point, determining the class or value of that point based on majority voting or the average of the neighbors. It works well with local patterns.

These models are integrated using a ridge regression (RR) [28] as a meta-learner. RR is a simple and stable model that prevents overfitting in stacking predictions of the base learners. The rationale behind the model selection is to maximize diversification and minimize overfitting. The proposed setup for the stacking model ensures diversity, stability, and optimizes generalization. As illustrated in Fig. 3, the working principle of the proposed model is as follows:

1. Training base learners (CBR, GBR, RFR, and KNN) on the input dataset
2. Generating predictions on a held-out validation set using base learners
3. Create a feature set to train the meta-model by combining the predictions of base learners
4. The RR then learns the optimal way to merge the output of base models

To improve the performance of each base learner and the overall performance of the stacking model, a two-level randomized grid search optimization was used. This approach involved independently tuning the hyperparameters for each base and meta learner. The RandomizedSearchCV method from the Scikit-learn library in Python was used, which samples a constant number of parameter settings from the pre-defined search space. For each model, the algorithm examined 100 random combinations of control parameters from the predefined search spaces presented in Table 2. The performance of each combination was evaluated using 5-fold cross-validation on the training data, with the R^2 used as the target metric to maximize. The hyperparameter permutation that generated the highest average R^2 score across the folds was selected as the optimal setting for that model.

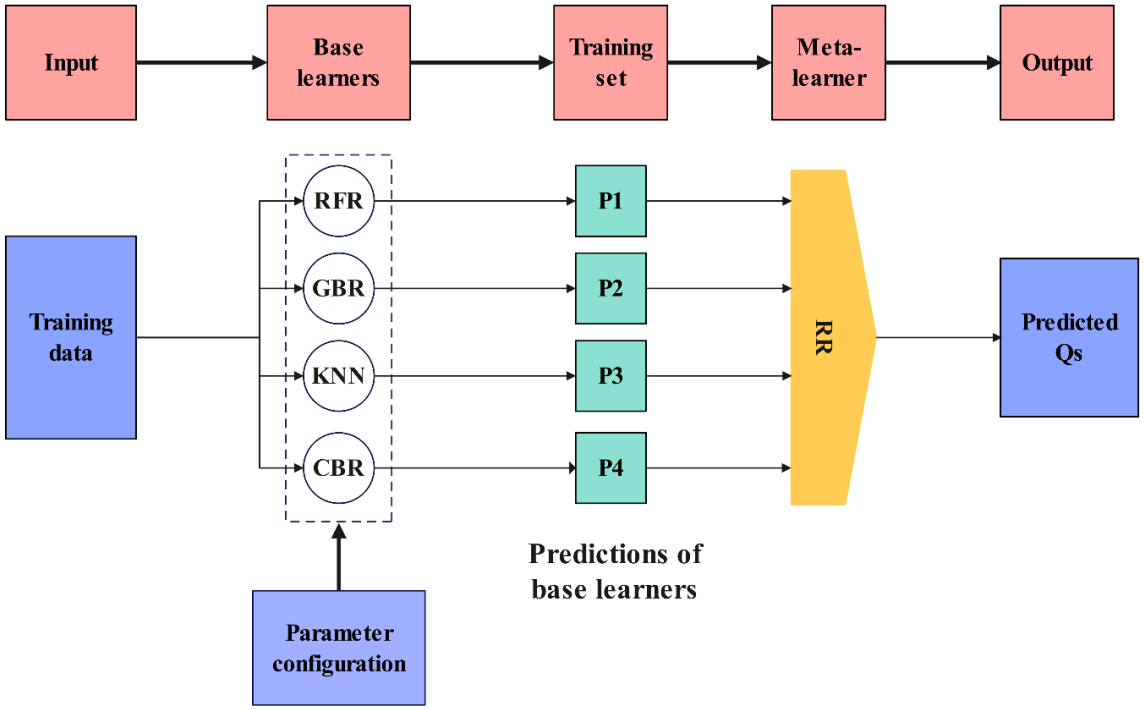


Fig. 3. The working principle of the proposed stacking model.

3. Experimental results

3.1. Dataset

Table 3 shows the characteristics of the benchmark dataset. The dataset contains 8792 records with three hydraulic variables: height (H), instantaneous concentration (C), and instantaneous water discharge (Ql), and a target instantaneous sediment discharge (Qs). The time period of the data (sampling years) is September 1970 to Aug. 2010.

Table 2
Control parameters, search space, and optimal values generated by the parameter tuning component.

Model	Hyperparameter	Search Space	Optimal value
RFR	n_estimators	[100, 200, 500]	500
	max_depth	[None, 10, 20, 30]	20
	min_samples_split	[2, 5, 10]	5
	max_features	['sqrt', 'log2']	'sqrt'
GBR	n_estimators	[100, 200, 500]	500
	learning_rate	[0.01, 0.05, 0.1]	0.05
	max_depth	[3, 5, 7]	5
	min_samples_split	[2, 5, 10]	5
KNN	n_neighbors	[3, 5, 7, 9, 11, 13]	7
	weights	['uniform', 'distance']	'distance'
	p	[1 (Manhattan), 2 (Euclidean)]	1 (Manhattan)
CBR	iterations	[500, 1000, 1500]	1500
	learning_rate	[0.01, 0.03, 0.05, 0.1]	0.03
	depth	[4, 6, 8, 10]	6
	l2_leaf_reg	[1, 3, 5, 10]	3
RR	alpha	[0.1, 0.5, 1.0, 2.0, 5.0, 10.0]	2.0
	fit_intercept	[True, False]	True
	solver	['auto', 'svd', 'cholesky', 'lsqr']	'auto'

Table 3
Descriptive statistics of the input and target variables in the benchmark dataset.

Metric	Variables			Target
	H	*C*	*Ql*	*Qs*
Count	8792	8792	8792	8792
Mean	112.9888	25.09814	18.93328	892.7361
Std. Dev.	69.71512	33.60104	53.59307	3683.383
Minimum	-33	0	0.001	0
25th Percentile	73	7	0.28	3.16075
Median (50th)	97	16.3	2.11	19.9875
75th Percentile	135	32	13.825	277.511
Maximum	760	1602	960.3	75383.55
Units: *H (m), C (g/L), Ql (m³/s), Qs (kg/s)*				

Figure 4 shows the pairwise plot matrix of the benchmark dataset variables, including the distribution of features (on the diagonal), pairwise scatterplots (lower triangle), and Pearson correlation coefficients (upper triangle). Correlation analysis shows that “Ql” (0.88) has the strongest positive correlation, and “C” (0.23) has the minimum correlation with “Qs”.

Shojaeezadeh et al. [29] found that flow variables have the greatest impact on suspended sediment load when using hydro-climatic variables and ML models. This finding is similar to the present correlation analysis, in which *Ql* was determined as the most important feature, with a correlation coefficient of 0.88. Stull and Ahmari [30] using satellite data and machine learning, showed that the correct selection of features significantly increases R^2. This observation aligns with the results of the SHAP analysis in the present study, which identified *Ql* as the most important variable.

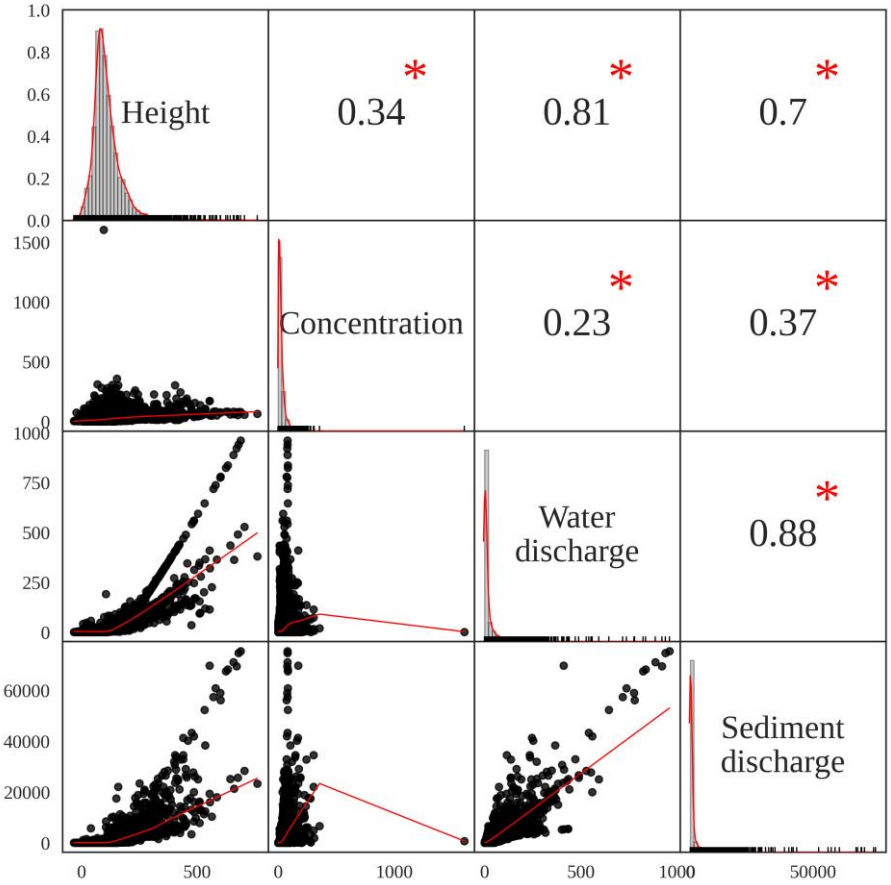


Fig. 4. Correlation heat-map matrix for the benchmark dataset.

3.2. Results

The performance of predictive models is examined using five performance metrics, including mean absolute error (MAE), coefficient of determination (R^2), root mean squared error (RMSE), Nash–Sutcliffe efficiency (NSE), and percentage bias (PBIAS). The R^2 index indicates the degree of fit between the simulated and observed data, with values closer to one representing higher model accuracy. [31]. The RMSE and MAE indices quantify the level of model prediction error, with RMSE being more sensitive to large errors than MAE [32,33]. The NSE index expresses the overall performance of the model by comparing observed and simulated values, and values close to one indicate an excellent fit [34]. The PBIAS index indicates the average tendency of the model to underestimate or overestimate the observed data [32]. The mathematical formulations of these criteria are as follows:

$$R^2 = \left[\frac{\sum_{i=1}^{n} \left(Y_i - \overline{Y}\right)(X_i - \overline{X})}{\sum_{i=1}^{n} \sqrt{(Y_i - \overline{Y})^2} \sum_{i=1}^{n} \sqrt{(X_i - \overline{X})^2}} \right]^2 \tag{1}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Y_i - X_i)^2} \tag{2}$$

$$NSE = EF = 1 - \frac{\sum_{i=1}^{n} (Y_i - X_i)^2}{\sum_{i=1}^{n} (Y_i - \overline{X})^2} \tag{3}$$

$$MAE = \sum_{i=1}^{n} \frac{1}{n} |X_i - Y_i| \tag{4}$$

$$PBIAS = \frac{\sum_{i=1}^{n} (Y_i - X_i)}{\sum_{i=1}^{n} Y_i} \times 100 \tag{5}$$

Table 4 summarizes the results of predictive models on the training and validation sets. The KNN model with the highest R^2=1 and the lowest error (MAE=0 kg/s, RMSE=0 kg/s) generated the best results compared with their counterparts in the training phase. However, KNN is likely suspected of overfitting because it shows poor results in validation and test datasets. The proposed stacking, CBR, and GBR performed reasonably well in the training phase. RR performed poorly, with a very high error (RMSE=1615.771 kg/s, and MAE=576.143 kg/s) in the training phase, likely due to underfitting. Regarding NSE and PBIAS, the CBR and stacking models are the top performers, with near-perfect validation NSE scores and low PBIAS. However, stacking demonstrates slightly better generalization due to the smaller gap between validation and training scores. The RFR and GBR predictors are also strong. The KNN model generated severe overfitting, and the RR model is the weakest predictor.

Table 4
Results of estimators on training and validation datasets.

Model	Training					Validation				
	R^2	MAE	RMSE	NSE	PBIAS (%)	R^2	MAE	RMSE	NSE	PBIAS (%)
KNN	**1.000**	**0.000**	**0.000**	**0.997**	0.276	0.928	126.721	911.969	0.926	4.311
RFR	0.977	61.761	534.115	0.979	**0.125**	0.970	82.464	586.574	0.970	**1.298**
GBR	0.987	169.018	399.394	0.986	-1.404	0.970	198.115	591.168	0.967	-1.191
CBR	0.998	60.611	162.232	**0.997**	-0.587	0.989	82.392	**361.914**	0.991	-2.072
RR	0.790	576.143	1615.771	0.790	-2.076	0.828	565.575	1413.357	0.829	-3.023
Stacking	0.998	43.390	139.049	0.995	1.305	**0.991**	**65.492**	318.131	**0.993**	2.575

Table 5 lists the results of predictive models on the testing datasets. The analysis of the results in validation and testing datasets shows that the stacking model generated the best performance compared with other models, producing the lowest error and the highest R^2 value (0.991 in validation and 0.973 in the testing phase). In terms of MAE, with a value of 68.887, the stacking model generated the best performance in the testing phase. GBR and CBR are also in the next positions, while KNN and RR perform poorly. The significant difference between the stacking and the other models shows the outstanding performance of the stacking model. The results also confirm that tree-based learning models (such as CBR and GBR) outperform linear models (e.g., RR), which indicates that stacking improved performance by intelligently combining various models. Regarding the NSE metric, the CBR obtained the highest NSE (0.976), and with a negligible difference, the stacking model possesses the second rank. All tree-based models, including GBR, CBR, and RFR, significantly outperformed the simpler models like KNN and RR. In terms of PBIAS, the most unbiased model is GBR (1.127%), followed by CBR (1.836%). With a slight change, the stacking model generated acceptable results with low underestimation bias. While individual models such as CBR and GBR generated superior values in terms of NSE and PBIAS criteria, the stacking model aims to combine the strengths of different learners to create a more generalizable and robust predictor. Its performance shows that the stacking model is successful. The model performs best in terms of R^2, MAE, and RMSE criteria while also presenting very low bias. This makes it a reliable tool for deployment in real-world scenarios.

Table 5
Results of estimators on testing datasets.

Model	R^2	MAE (kg/s)	RMSE (kg/s)	NSE	PBIAS (%)
KNN	0.932	108.837	1063.287	0.930	3.580
RFR	0.946	91.517	949.225	0.953	3.342
GBR	0.961	203.664	805.219	0.965	**1.127**
CBR	0.968	89.854	725.476	**0.976**	1.836
RR	0.819	584.897	1737.584	0.821	-4.736
Stacking	**0.973**	**68.887**	**671.147**	0.968	3.209

Bezak et al. [17], using the XGBoost model to measure sediment transport under climate change, demonstrated that machine learning models outperform traditional methods. This finding aligns with the results of the present study, which showed the superiority of the stacking model over linear models such as RR. Mir & Patel [35], in a comprehensive review, it was stated that machine learning models outperform empirical equations in complex hydrodynamic environments. This observation is consistent with the better performance of the stacking model over RR in the present complex dataset. The findings of the present study, in addition to the high accuracy of the stacking model in estimating sediment load, can also be useful in macro-management of water resources. Recent studies have shown that accurate hydrological forecasts play an important role in supporting stakeholders' decision-making. The use of game theory-based models to determine the optimal exploitation of groundwater resources showed that designing reliable predictive models can help in multilateral decision-making and resolving conflicts between operators and the environment [36]. On the other hand, examining the sensitivity of variables in hydrological forecasting models is of great importance. Using sensitivity analysis methods (such as RF–GA) enhances the accuracy of forecasts [37]. Recent advances in the field of hybrid metaheuristics, such as the ELM–GWO model for predicting the compressive strength of concrete, demonstrate the potential of advanced search and optimization algorithms in enhancing the performance of ML models [38].

Figure 5 shows the prediction error plots generated by predictive models on the testing dataset. The plots examine the alignment between predictions and observations. Data points on the *x* and *y*-coordinates show the observations and predictions, respectively. The plots show that most of the models have promising alignment between the observations and predictions, and their results fall almost on the diagonal line (perfect alignment). The best alignment is achieved by the stacking model with R^2: 0.973, in testing data. RR with R^2: 0.819 produced the worst results with some outliers in the training and testing datasets. CBR, GBR, and RFR performed well on the testing set. KNN with R^2: 0.932 shows moderate results with some outliers.

Figure 6 shows partial dependency (PDP) plots for the features in the dataset. The plots clearly show how changes in the values of the features affect the response of the predictive model. These plots also act as a validation tool, showing how well the model results are consistent with the domain knowledge. As the values of *C* and *Ql* increase, the value of the target variable increases continuously. Variable *H* initially has a neutral effect on the target variable, and after reaching the number 150, it has a positive effect on the target variable. Overall, variables *C* and *Q1* are key drivers with threshold effects, while *H* has a lower linear effect. Delcey et al. [39] using PINN networks proved that machine learning can explain physical aspects, such as particle settling velocity, in addition to prediction. This result is similar to the present PDP and SHAP analysis, which clearly revealed the relationships between features and the target variable.

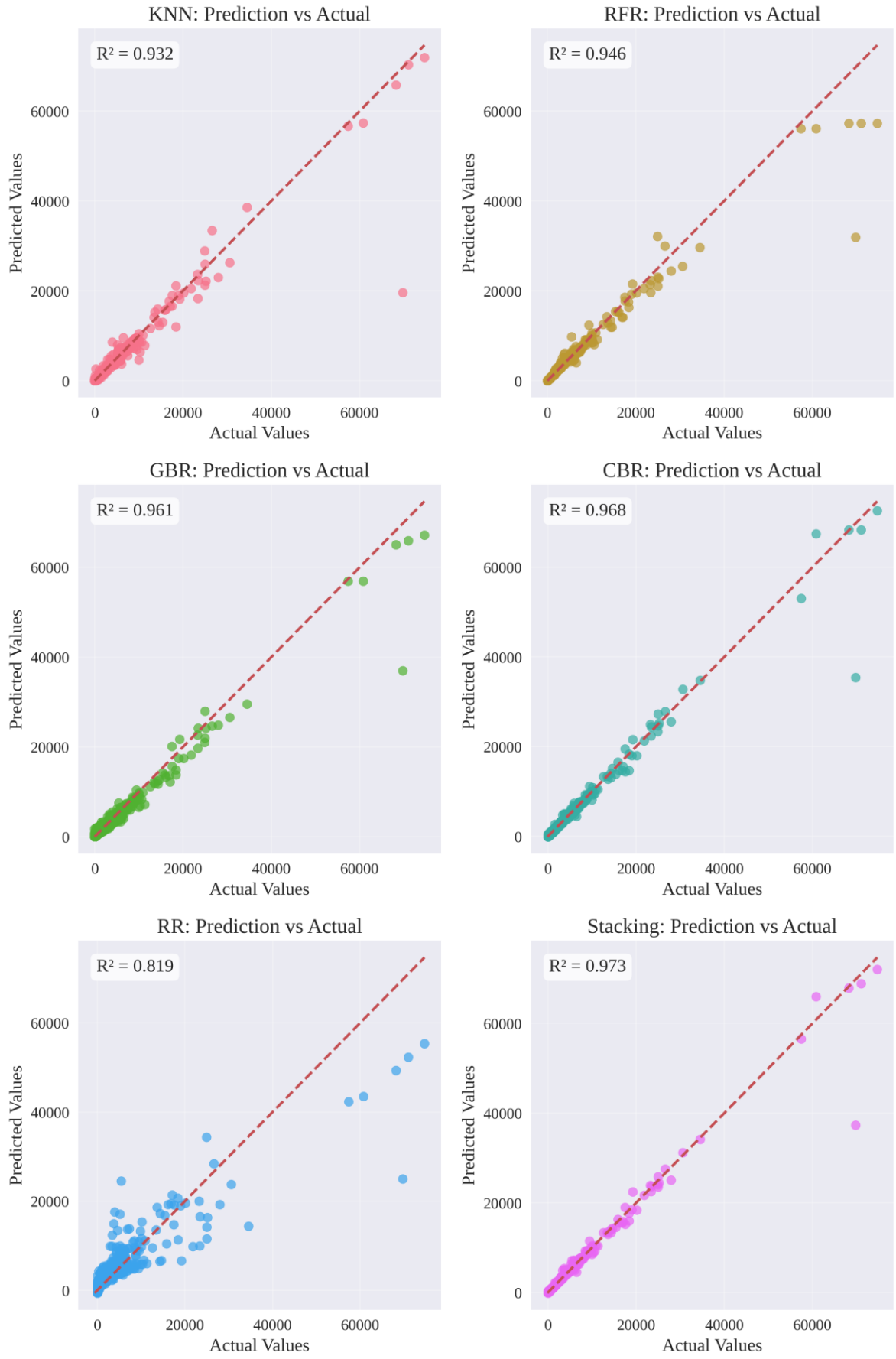


Fig. 5. Prediction error plots generated by the proposed stacking model and its constituent models on the testing data.

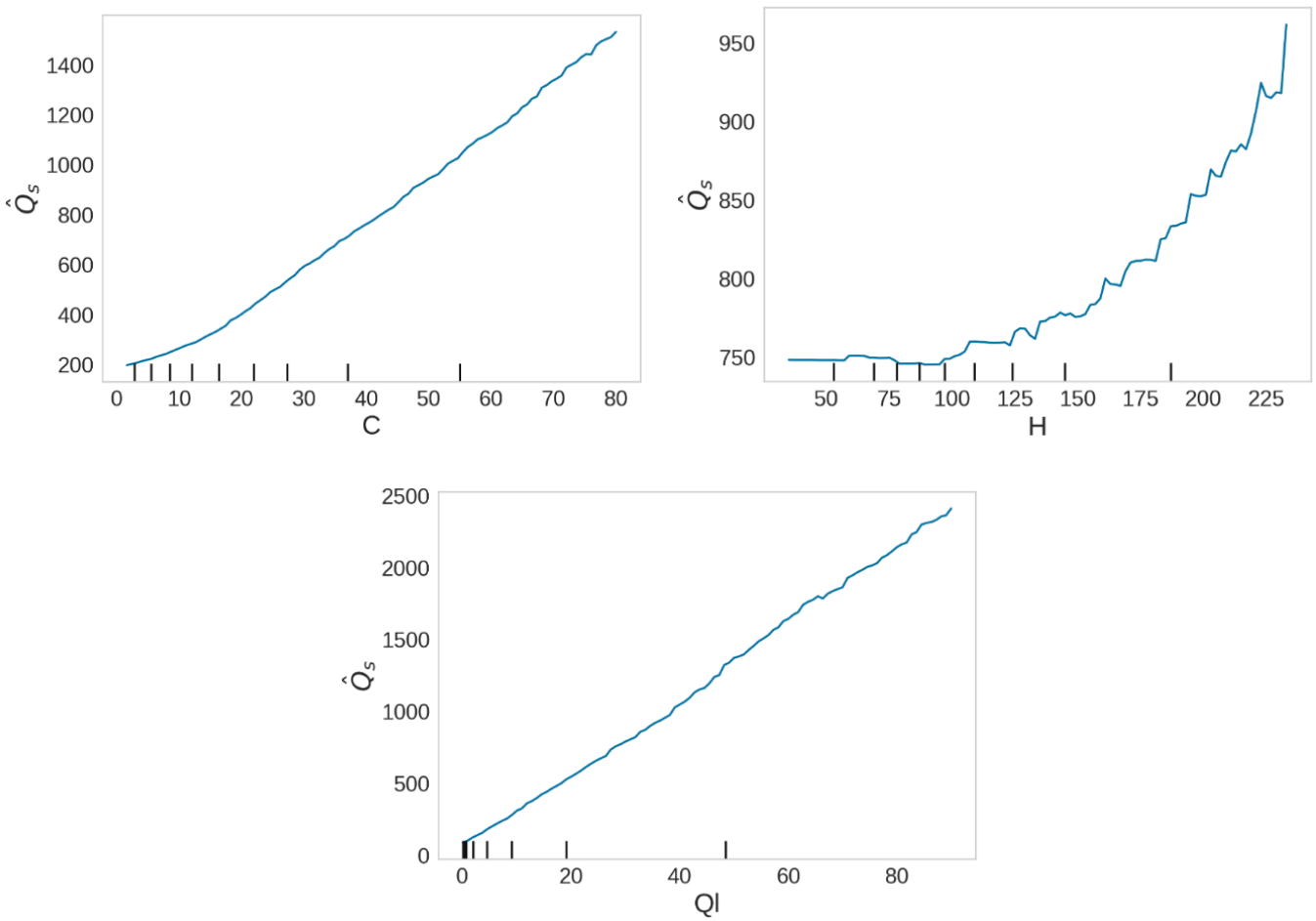


Fig. 6. Dependency between features generated by the stacking model in the training data.

The results of five-fold cross-validation in terms of the R^2 metric for predictive models in training data are shown in Fig. 7. In this way, the input data is randomly partitioned into five equal-sized partitions. One partition is used for validation, while the other four sets are used to train the learner. This process is iterated five times, with each partition using the validation data exactly once. By computing the average of the results of these iterations, a reliable validation of the model's performance is obtained. The plots confirm that the proposed Stacking, CBR, and GBR models show stable behavior across all folds.

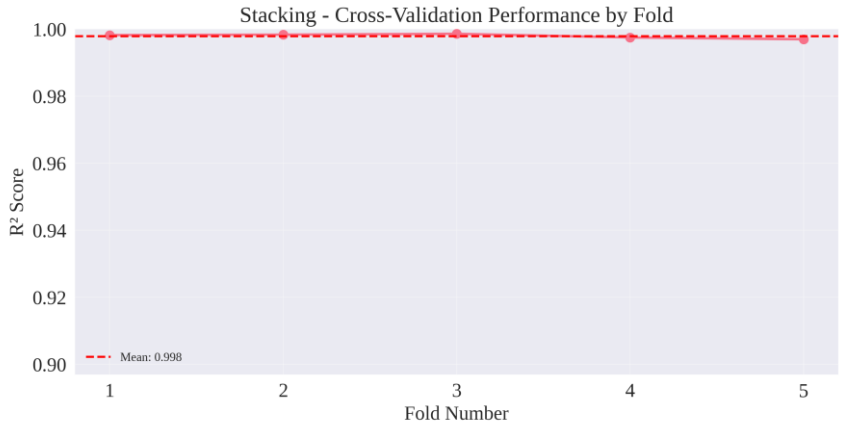


Fig. 7. The R^2 values obtained by predictive models on the training data computed by 5-fold cross-validation.

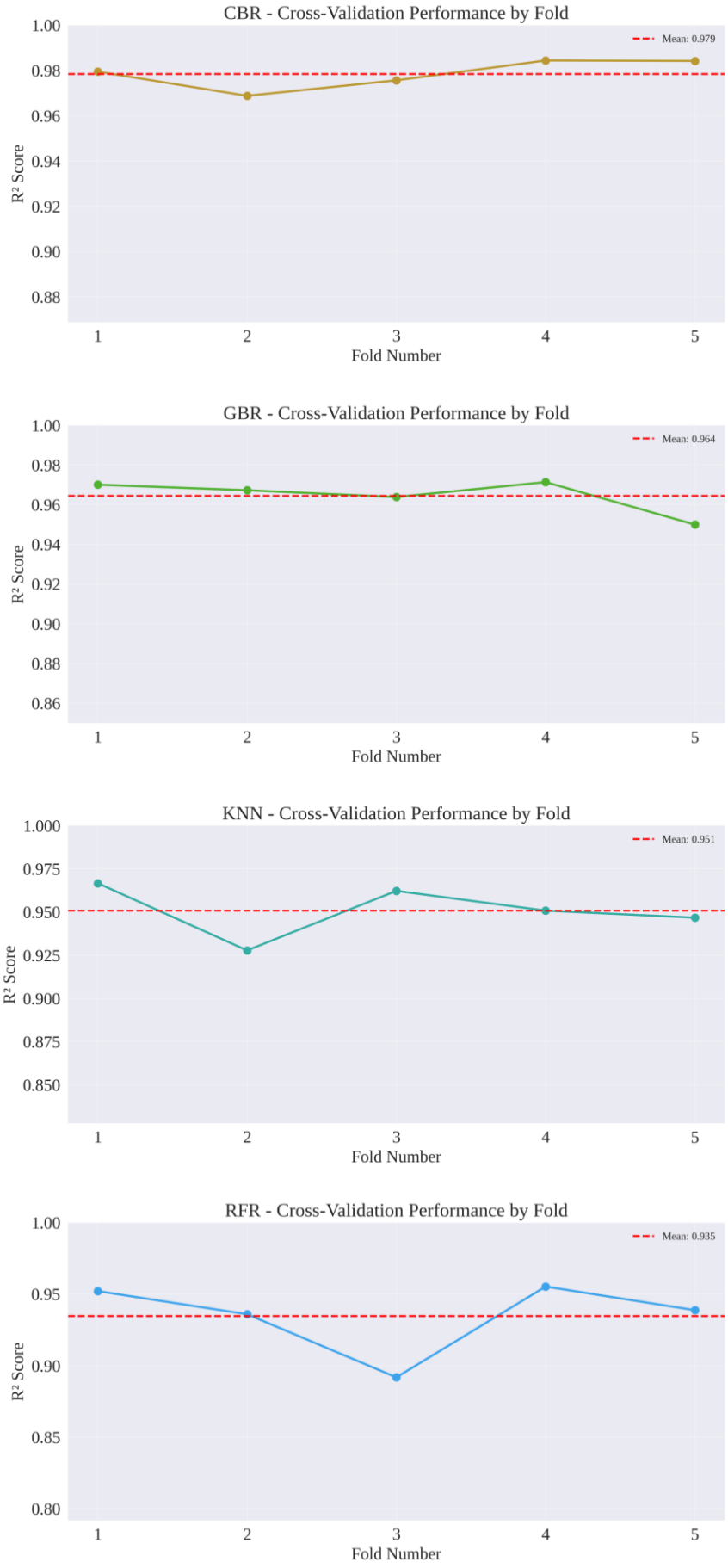


Fig. 7. (continued).

Figure 8 shows the effect of tuning the control parameters of the estimator models. The plots prove that the tuning component improved the prediction performance of the ML predictors. The improvement (degradation) rate of predictive models regarding the R^2 metric is +6.7, +0.2, -0.8, +0.1, and +0.0%, for KNN, RFR, GBR, CBR, and RR models, respectively. Regarding the MAE metric, the improvement (degradation) rate is +27.9, +3.1, +18.3, -41.5, +3.8%, and in terms of RMSE, the improvement rate is +33.9, +2.4, -15.8, +1.4, and +0.2%, for KNN, RFR, GBR, CBR, and RR models, respectively. Figure 9 illustrates the distribution of performance indexes generated by models on training data. The algorithms were executed by a 5-fold cross-validation strategy.

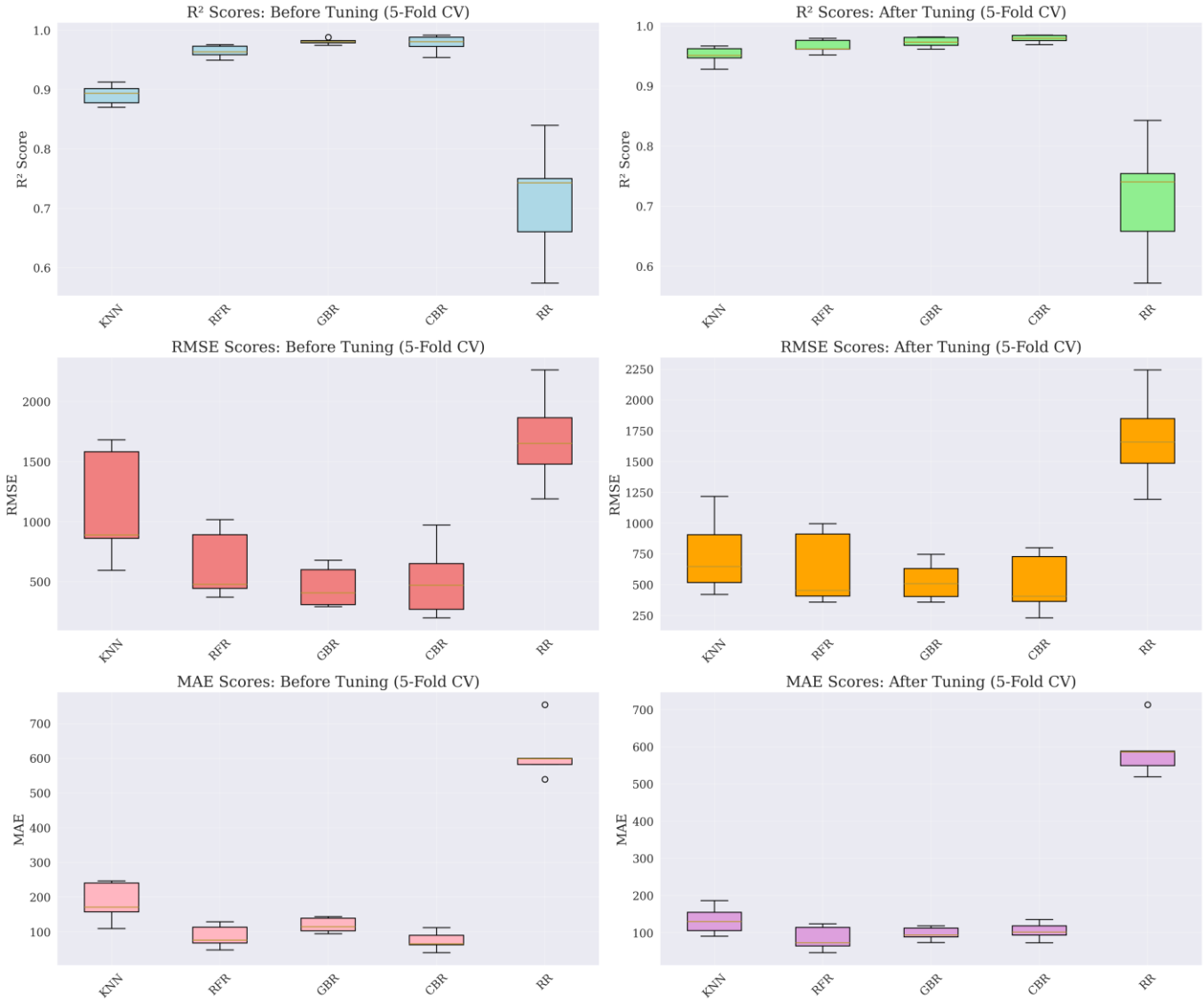


Fig. 8. The impact of tuning the control parameters of the estimators on their estimation performance on the training data.

Figure 10 was added to show a comparative time series of the observed Qs versus the stacking model-predicted Qs for a representative period of the testing dataset (the first 300 records from the testing data). The figure visually shows the model's high accuracy in tracking the temporal dynamics of the observed data, including its performance in predicting the timing and magnitude of peak events.

Fig. 9. Distribution of performance indexes obtained by predictive models in 5-fold cross-validation on the training data.

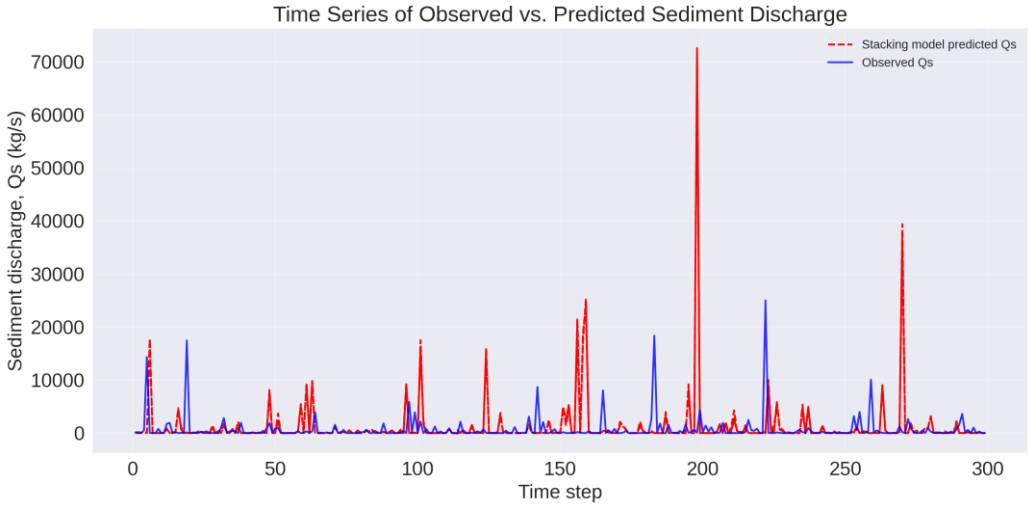
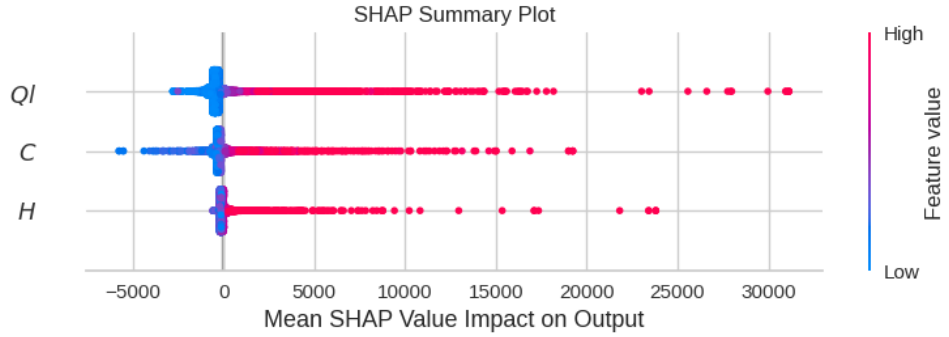


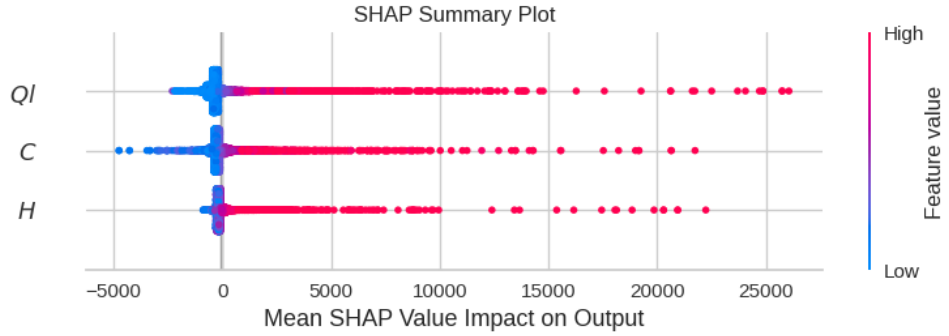
Fig. 10. Comparison of time series of the observed Qs versus the stacking model-predicted Qs for a representative period of the testing dataset.

The SHapley Additive exPlanations (SHAP) shows how much influence each feature has on the model's prediction. It ranks the features in order of importance. Figure 11 shows the results of SHAP analysis on the training and testing datasets. The results confirm that *Ql* is the most important feature in predicting the target Qs. The second most important feature is *C,* followed by *H*. The plots confirm that the ranking of feature importance remains perfectly consistent between the training and testing datasets: Ql > C > H. This stability, despite a slight decrease in absolute SHAP values on the testing set, shows that the reasoning of the proposed model is robust and not a product of overfitting to the training data. It generalizes reliably to unseen testing data and preserves the relative importance of each feature. It should be noted that the SHAP plot shows the relative importance of the features; in other words, the importance of the features compared to other features.

The analysis shows that Ql is the most crucial factor in sediment transport, supported by SHAP analysis and fundamental principles of fluvial geomorphology. Increasing Ql enhances stream power and shear stress on the riverbed, resulting in more sediment mobilization. This is reflected in the strong positive correlation (0.88) between sediment discharge (Qs) and Ql, as well as the continuous increase shown in the PDP plot (Fig. 6). The analysis also confirms that feature C, representing sediment availability, is an important factor, albeit secondary to Ql. High C indicates prior erosion, making sediment readily available. The model reveals that combining high C and high Ql significantly increases sediment discharge, especially during events like floods or heavy rainfall. Parameter H is less significant than Ql in the model, although H and Ql are related. The model likely favors Ql because it is a more direct influence. H only positively affects sediment load after reaching a certain threshold (around 150), probably the bankfull stage, when the river expands onto its floodplain and increases sediment availability, which the model accurately captures. The feature importance ranking (Ql > C > H) aligns with the physical processes controlling sediment transport in the Wadi Mina basin. In summary, Ql is the most important factor, followed by C, with H playing a less direct role.



(a) Training



(b) Testing

Fig. 11. SHAP analysis results on the training and testing data to show the importance of features.

Figure 12 shows the SHAP dependency plots with interaction effects on the testind data. The results indicate that water flow (Ql) and sediment concentration (C) have a strong synergistic effect on sediment transport predictions. Also, the water height variable (H) shows a secondary but positive effect, especially in high flow conditions. As shown in Fig. 13, SHAP analysis across flow regimes proves demonstrates that Ql is the principal driver of sediment transport predictions. Its relative importance increases substantially during high-flow conditions, while the effect of flow H inversely decreases with greater discharge.

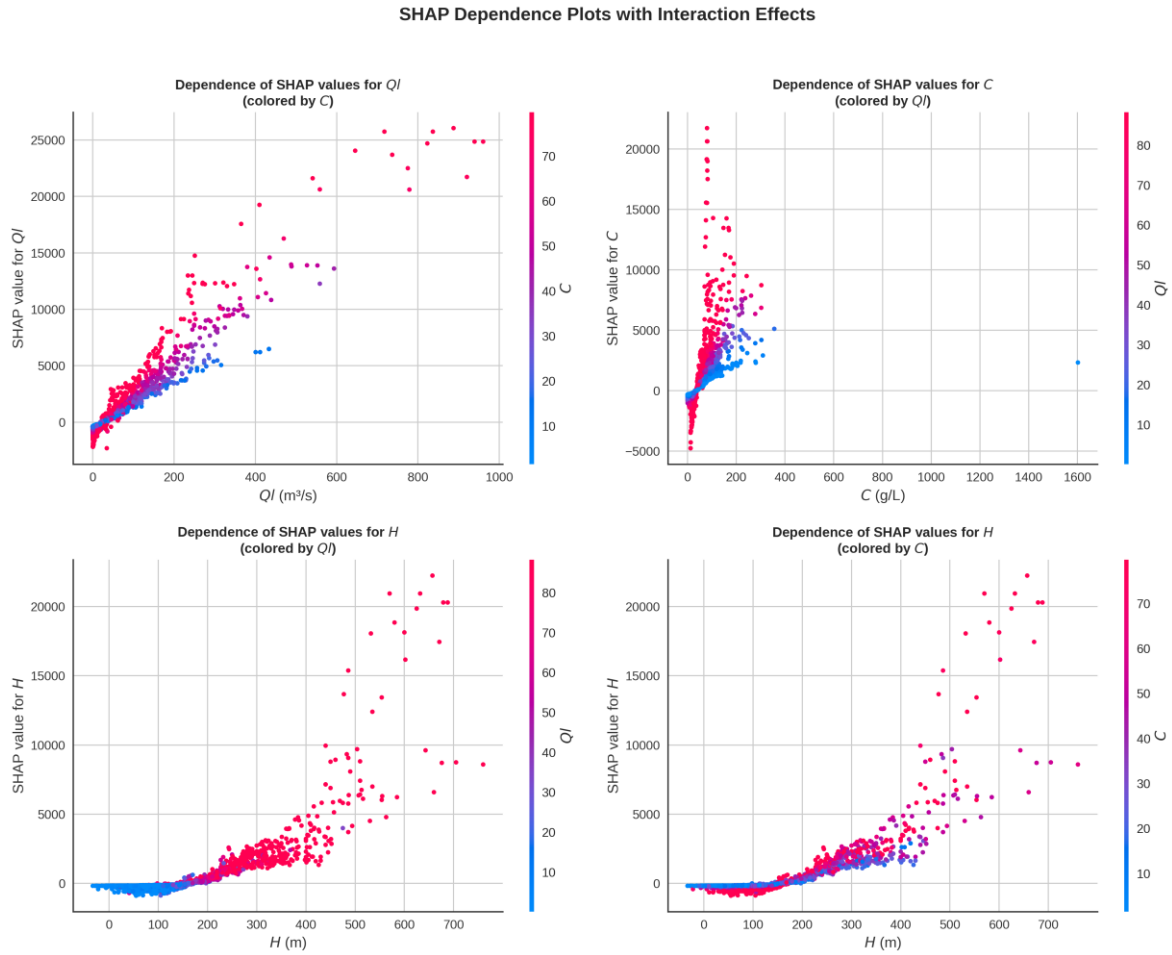


Fig. 12. SHAP dependency plots with interation effects.

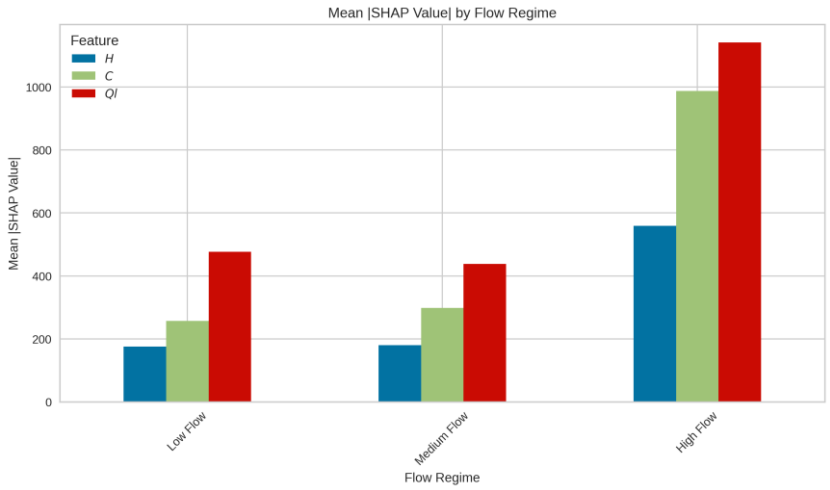


Fig. 13. SHAP feature importance by flow regime.

Figure 14 shows the Taylor diagram for the estimators on the benchmark sets. The graph shows that all models have a strong correlation close to the observation point for the training set. Regarding the test set, the proposed model generated better results compared to other models and is close to the observation point. The other algorithms are slightly away from the observation point. The angular axis (θ) shows the arc-cosine of the Pearson correlation coefficient between the observations and model predictions. Therefore, the angle in radians (converted to degrees on the plot labels) directly shows the correlation value. The value of 0° shows a perfect positive correlation (R = 1.0), and 90° shows no correlation (R = cos(90°) = 0).

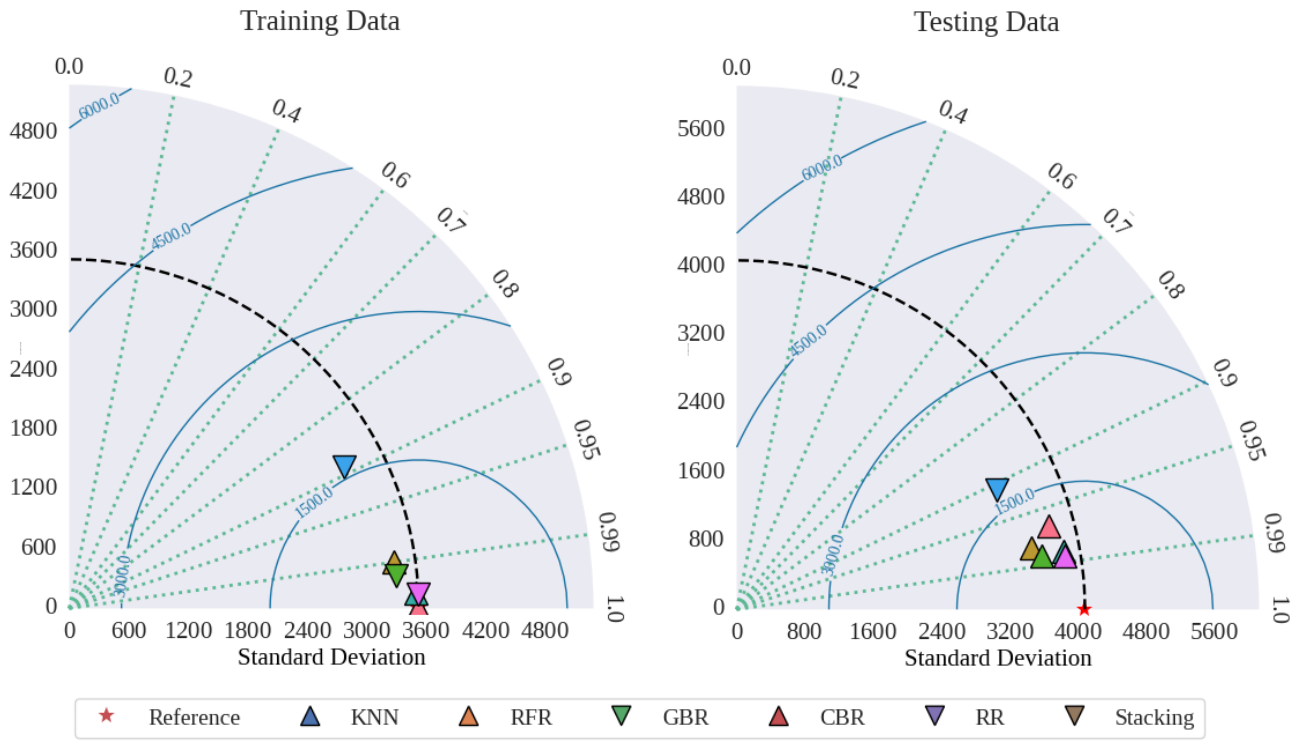


Fig. 14. Taylor diagram for the estimators on the training and testing data.

Figure 15 shows the residual plots of the predictive models on the benchmark data. Residuals measure the differences between predicted and actual values. A residual plot examines the homoscedasticity and linearity of ML models. Residual plots confirm the success of ML in covering the diverse relations between datapoints in the dataset. The analysis of plots proves the superiority of the stacking model (with the minimum RMSE value) in terms of reliability and robustness compared with other models.

Figure 16 shows the Quantile-Quantile (Q-Q) plot of residuals generated by the predictive models on the testing data to test if the residuals of a model are normally distributed. In a perfectly normal distribution of residuals, all points lie around the straight diagonal line. The Q-Q plot for the stacking model shows that the residuals in the middle of the distribution are largely close to the reference line, indicating a better fit than some other models, such as RR, RFR, GBR, and KNN; however, there is a significant deviation from the normal line in the tails, especially on the right tail. Overall, the stacking model performs better in normalizing errors than other models, but the problem of outliers and heavy tails still persists and can negatively affect the quality of the fit.

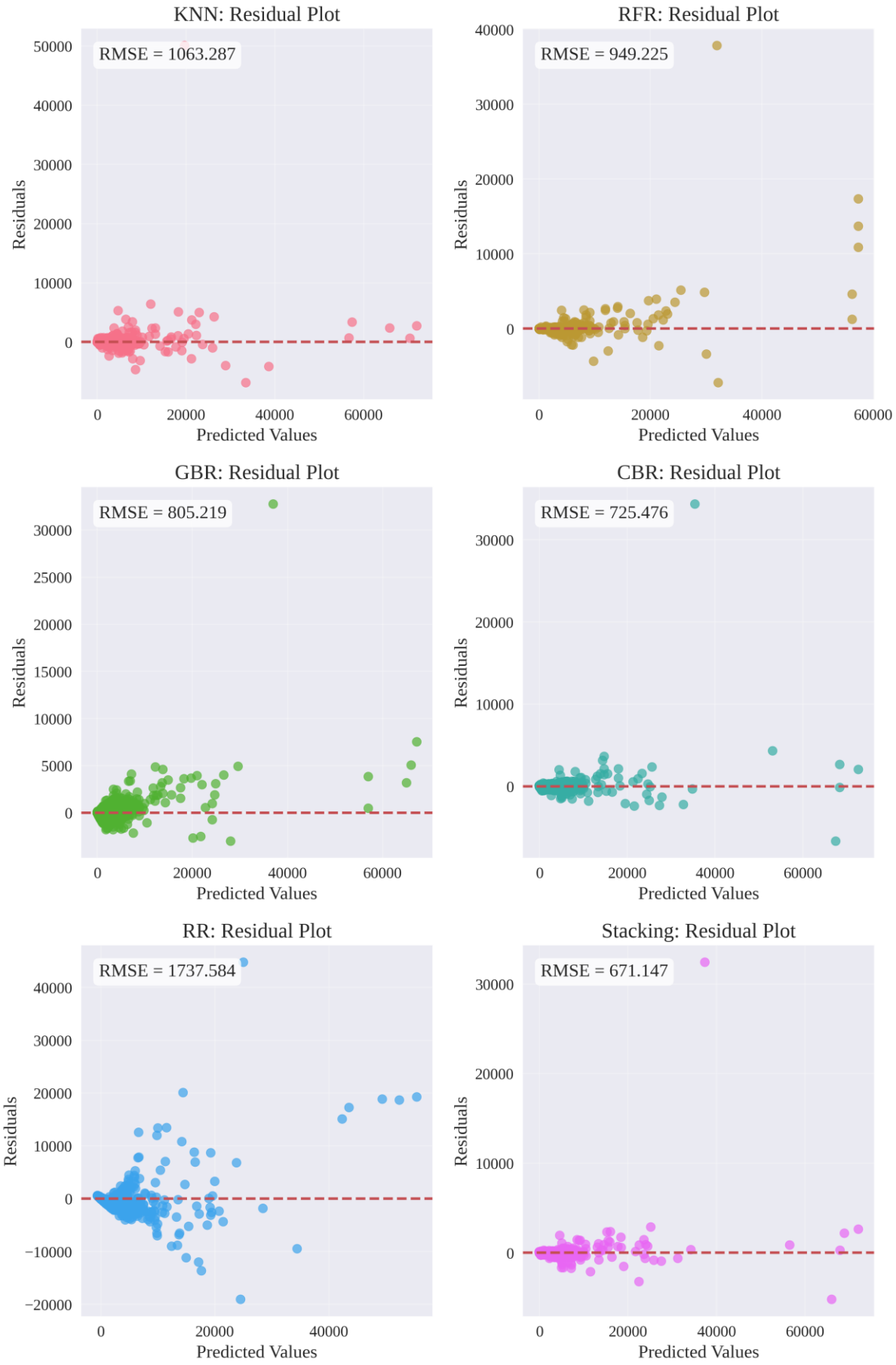


Fig. 15. Residual plot generated by the proposed stacking model and its constituent models on the testing data.

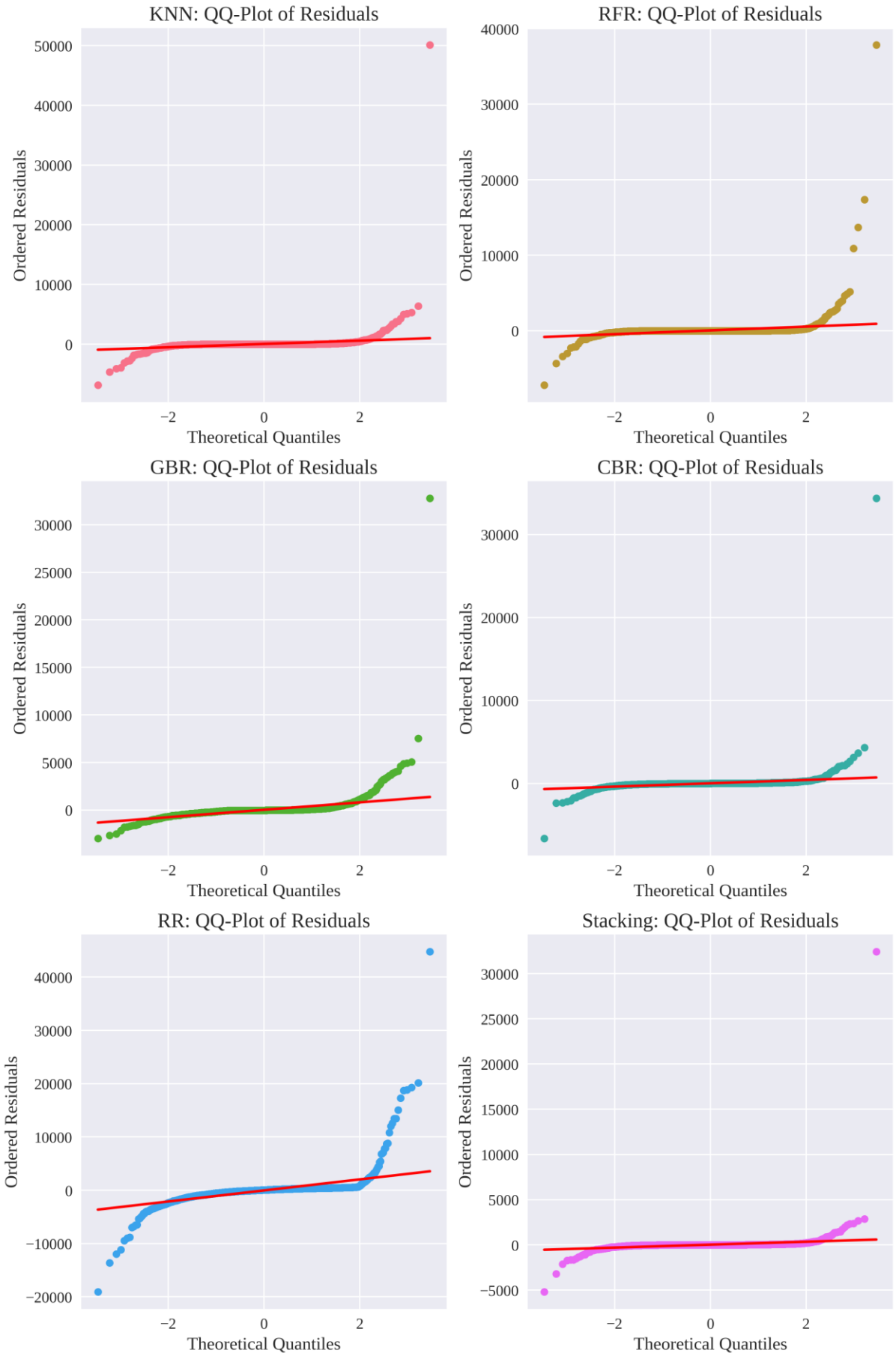


Fig. 16. Q-Q plot of residuals generated by predictive models on the testing data.

Figure 17 shows the residuals versus input variables (Q1 and C) for the models. The objective of these plots is to examine systematic patterns in the residuals and identify overfitting or underfitting. Analysis of the graphs shows that the residuals in the stacking model are random and the model has been able to learn the overall pattern of the data without overfitting. The graphs clearly show that there are large outliers that increase the error variance. Therefore, the overfitting problem has been largely solved, but for further improvement, work should be done to remove outliers or use data transformation, which is considered part of the future work of this research.

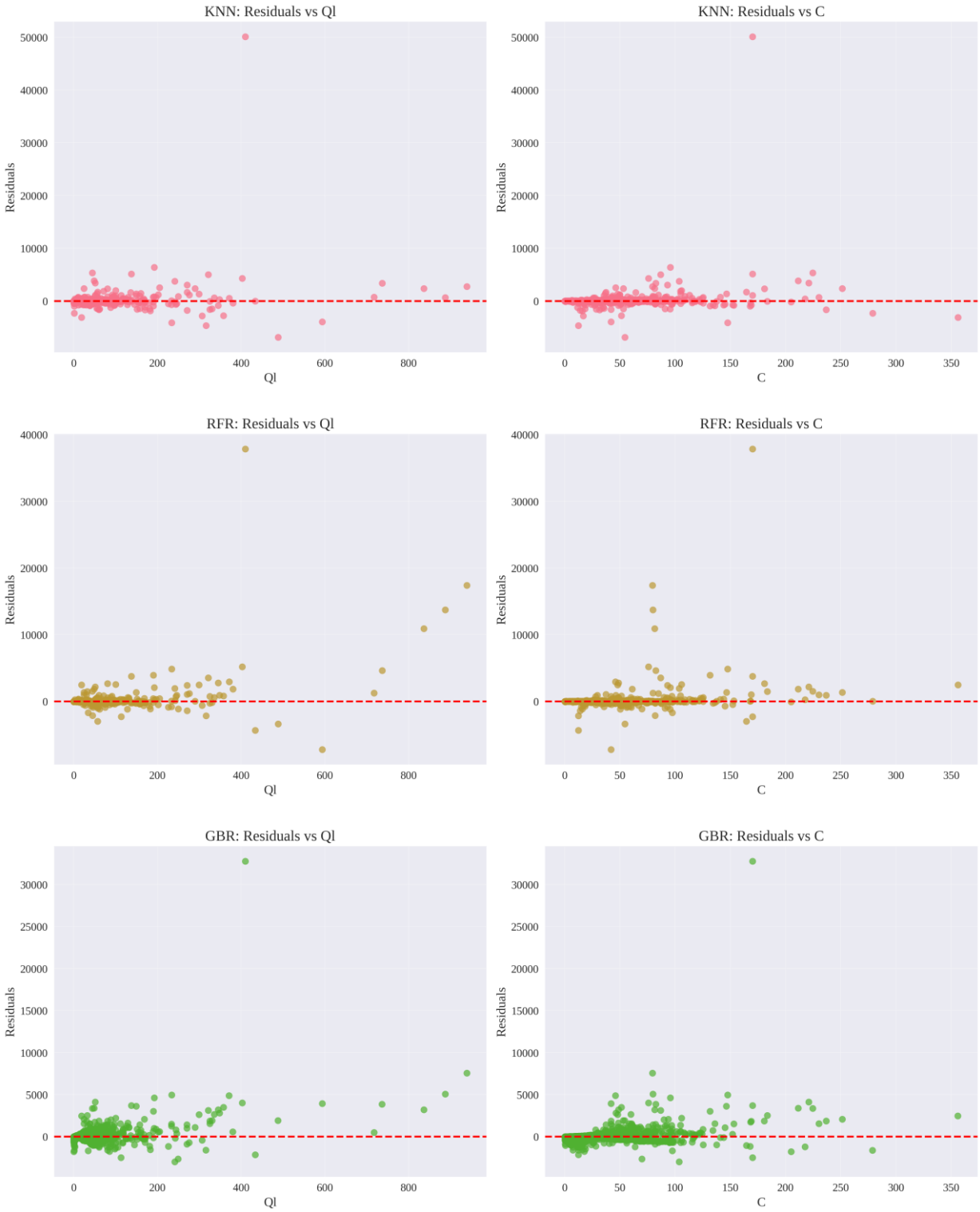


Fig. 17. Residuals-vs-predictors plots for Ql and C variables.

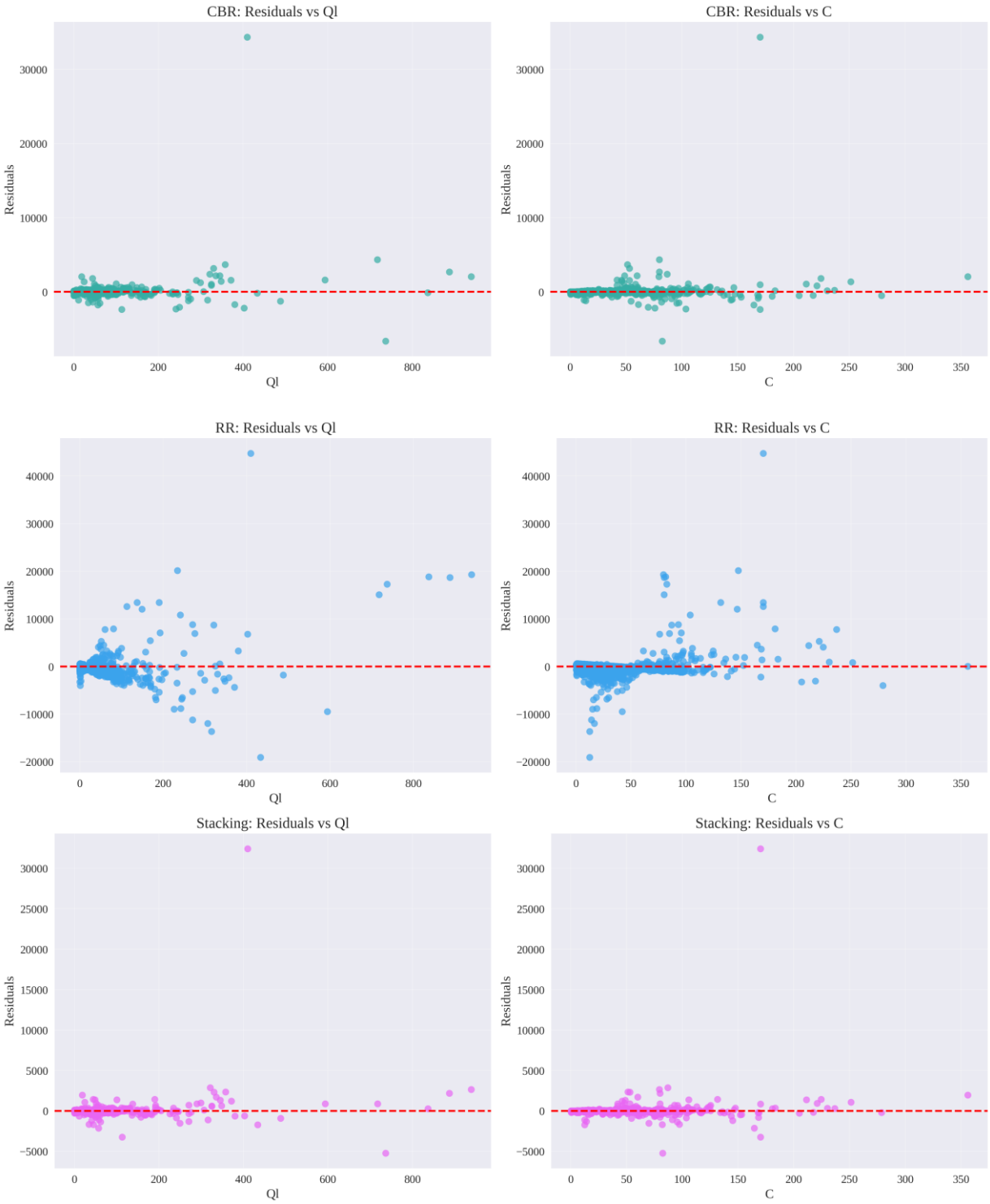


Fig. 17. (continued).

Table 6 lists the residual analysis for predictive models on the testing dataset. The analysis of results confirmed that none of the models fully satisfied the assumption of normality and homogeneity of error variances. Despite the very high error, the RR model is closer to normal in terms of error distribution (less skewness and kurtosis). The CBR model performed best regarding accuracy and the lowest standard deviation of error, but due to the very high skewness and kurtosis, its error distribution is unbalanced. The two GBR and Stacking models have provided a good balance between accuracy and normality of errors, and can be considered as more balanced options for predicting in conditions of heterogeneous data with outliers.

Table 6
Results of residual analysis for predictive models on the testing dataset.

Model	Breusch-Pagan		Residuals			
	LM	p-value	Mean	Std	Skew	Kurtosis
KNN	74.788	4.02E-16	32.709	1077.634	38.144	1766.563
RFR	200.454	3.37E-43	30.534	885.089	26.877	1014.162
GBR	118.364	1.74E-25	10.296	767.510	24.474	938.407
CBR	**87.697**	6.84E-19	**16.777**	**634.897**	35.762	1616.419
RR	505.594	2.9E-109	-61.545	1724.792	**8.122**	**202.580**
Stacking	120.243	6.84E-26	70.441	722.787	32.753	1367.546

Table 7 compares the achievements of the proposed stacking model on the testing dataset with relevant and recent models that used advanced ML methods. Analyzing the results, we find that the stacking model outperforms most of the other competitors. It achieves low errors (RMSE, MAE), high accuracy (R^2, NSE), and maintains a low PBIAS. The DNN model [20] generated the lowest RMSE and the highest R^2. Our model obtained an R^2 of 0.973, indicating its ability to explain 97.3% of the variance in the sediment transport data. These results confirm the applicability and high accuracy of the proposed stacking approach. Regarding the DNN model, which generated a perfect R^2 of 1.00 and a remarkably low RMSE of 243.72 kg/s on the testing dataset, it is important to note the potential for overfitting indicated by its higher MAE (102.175 kg/s) compared to the MAE generated by our proposed model (68.887 kg/s). A perfect R^2 score on real-world hydrological data can sometimes be an indicator of overfitting, showing the model may have memorized the training data rather than learning generalizable patterns within the data. The stacking model's strong performance across all metrics confirms its balanced and robust predictive ability. Further investigation reveals that our stacking model provides a highly competitive and reliable alternative for sediment transport estimation.

Table 7
Comparison of performance indicators generated by the proposed model and its peers on the testing data.

Model	Model Name	R^2	RMSE (kg/s)	MAE (kg/s)	NSE	PBIAS (%)
Emami et al. [1]	WOA-ANFIS	0.954	983.211	98.715	0.9538	4.717
Kundu et al. [18]	ETR	0.970	792.351	**56.896**	0.970	3.483
Achite et al. [20]	DNN	**1.00**	**243.72**	102.175	**0.990**	6.81
Achite et al. [20]	GRU	0.91	1597.56	269.942	0.870	−14.78
Achite et al. [20]	SRC	0.89	2597.51	470.890	0.660	−40.59
Tamrabet et al. [19]	ANFIS	0.964	871.551	90.791	0.963	4.732
Current work	Stacking	0.973	671.147	68.887	0.968	**3.109**

4. Conclusion

This paper presents an enhanced ensemble machine learning model to enhance the estimation performance of the sediment transport. The proposed ensemble improves estimation performance by integrating diverse machine learning models. Each model discovers specific data patterns, and their intelligent combination enhances overall performance, robustness, and stability. With a test on a dataset collected from the Wadi Mina basin at Oued Abtal hydrometric station, the proposed model (with R^2:

0.973, MAE: 68.887 kg/s, RMSE: 671.147 kg/s) is proven to have significant improvement over base machine learning models and several other comparison algorithms. This study faces several limitations that motivate future research directions. First, the proposed model was developed and evaluated on a specific dataset from a semi-arid basin, which questions its transferability to other hydrological basins. Second, the accuracy of the model relies on the availability of three input variables, especially the instantaneous concentration (C), which is often scarce and costly to compute in data-scarce regions. Third, the superior performance reported by the staking model needs a trade-off of higher computational complexity compared to individual learners. Future work will therefore focus on embedding new learners into the stacking model to improve the sediment estimation performance, examining the model's generalizability across different datasets, developing strategies to control missing data, and using further climatic features to improve applicability and robustness in data-scarce regions.

CRediT authorship contribution statement

Hojjat Emami: Conceptualization, Investigation, Methodology, Writing, Project administration, Software, Visualization, Writing – original draft; Writing – review & editing.

Somayeh Emami: Conceptualization, Data curation, Validation, Writing – review & editing.

Mohammed Achite: Formal analysis, Supervision, Writing – review & editing.

Aseel Smerat: Investigation, Writing – review & editing.

Funding

This research received no external funding.

Conflicts of interest

The authors declare no conflict of interest.

References

[1] Emami H, Emami S. Application of Whale Optimization Algorithm Combined with Adaptive Neuro-Fuzzy Inference System for Estimating Suspended Sediment Load. J Soft Comput Civ Eng 2021;5:1–14.

[2] E D, İ Y, Ö K. Estimation of total sedimentload concentration obtained by experimentalstudyusingartificial neural networks. Env FluidMech 2007;7:271–288.

[3] Fakharian P, Rezazadeh Eidgahee D, Akbari M, Jahangir H, Ali Taeb A. Compressive strength prediction of hollow concrete masonry blocks using artificial intelligence algorithms. Structures 2023;47:1790–802. https://doi.org/10.1016/j.istruc.2022.12.007.

[4] Ghanizadeh AR, Ziaie A, Khatami SMH, Fakharian P. Predicting resilient modulus of clayey subgrade soils by means of cone penetration test results and back-propagation artificial neural network. J Rehabil Civ Eng 2022;10:146–62. https://doi.org/10.22075/JRCE.2022.25013.1568.

[5] Jain SK. Development of integrated sediment rating curves using ANNs. J Hydraul Eng 2001;127:30–7.

[6] Ciğizoğlu HK. Suspended sediment estimation for rivers using artificial neural networks and sediment rating curves. Turkish J Eng Environ Sci 2002;26:27–36.

[7] Tayfur G. Artificial neural networks for sheet sediment transport. Hydrol Sci J 2002;47:879–92. https://doi.org/10.1080/02626660209492997.

[8] Cigizoglu HK, Alp M. Generalized regression neural network in modelling river sediment yield. Adv Eng Softw 2006;37:63–8. https://doi.org/10.1016/j.advengsoft.2005.05.002.

[9] Tayfur G, Guldal V. Artificial neural networks for estimating daily total suspended sediment in natural streams. Nord Hydrol 2006;37:69–79. https://doi.org/10.2166/nh.2005.031.

[10] Raghuwanshi NS, Singh R, Reddy LS. Runoff and Sediment Yield Modeling Using Artificial Neural Networks: Upper Siwane River, India. J Hydrol Eng 2006;11:71–9. https://doi.org/10.1061/(asce)1084-0699(2006)11:1(71).

[11] Alp M, Cigizoglu HK. Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. Environ Model Softw 2007;22:2–13. https://doi.org/10.1016/j.envsoft.2005.09.009.

[12] Melesse AM, Ahmad S, McClain ME, Wang X, Lim YH. Suspended sediment load prediction of river systems: An artificial neural network approach. Agric Water Manag 2011;98:855–66. https://doi.org/10.1016/j.agwat.2010.12.012.

[13] Mustafa MR, Rezaur RB, Saiedi S, Isa MH. River suspended sediment prediction using various multilayer perceptron neural network training algorithms-A case study in Malaysia. Water Resour Manag 2012;26:1879–97. https://doi.org/10.1007/s11269-012-9992-5.

[14] Mustafa MR, Isa MH, Rezaur RB. Artificial Neural Networks Modeling in Water. Nternational J Civ Environ Eng 2012;6:128–36.

[15] Panahi F, Ehteram M, Emami M. Suspended sediment load prediction based on soft computing models and Black Widow Optimization Algorithm using an enhanced gamma test. Env Sci Pollut Res 2021;21:48253–48273.

[16] Samantaray S, Sahoo A, Satapathy DP, Oudah AY, Yaseen ZM. Suspended sediment load prediction using sparrow search algorithm-based support vector machine model. vol. 14. Nature Publishing Group UK; 2024. https://doi.org/10.1038/s41598-024-63490-1.

[17] Bezak N, Lebar K, Bai Y, Rusjan S. Using Machine Learning to Predict Suspended Sediment Transport under Climate Change. Water Resour Manag 2025;39:3311–26. https://doi.org/10.1007/s11269-025-04108-7.

[18] Kundu S, Swarnkar S, Agarwal A. Bayesian-optimized recursive machine learning for predicting human-induced changes in suspended sediment transport. Environ Monit Assess 2025;197:603. https://doi.org/10.1007/s10661-025-14039-w.

[19] Zeyneb T, Nadir M, Boualem R. Modeling of suspended sediment concentrations by artificial neural network and adaptive neuro fuzzy interference system method–study of five largest basins in Eastern Algeria. Water Pract Technol 2022;17:1058–1081.

[20] Achite M, Katipoğlu OM, Elshaboury N, Tuğrul T, Pandey K. Intercomparison of sediment transport curve and novel deep learning techniques in simulating sediment transport in the Wadi Mina Basin, Algeria. Environ Earth Sci 2025;84. https://doi.org/10.1007/s12665-024-12051-w.

[21] Anaraki MV, Kadkhodazadeh M, Morshed-Bozorgdel A, Farzin S. Predicting rainfall response to climate change and uncertainty analysis: Introducing a novel downscaling CMIP6 models technique based on the stacking ensemble machine learning. J Water Clim Chang 2023;14:3671–91. https://doi.org/10.2166/wcc.2023.477.

[22] Kadkhodazadeh M, Anaraki MV, Kachoueiyan F, Farzin S. Introducing a Novel Double Hybrid Algorithm (DHA) and Developing Its Application for Predicting Air Temperature Under Climate Change Conditions (A Case Study of Iran Country). Eng J 2024;28:31–67. https://doi.org/10.4186/ej.2024.28.11.31.

[23] Farzin S, Valikhan Anaraki M, Kadkhodazadeh M, Morshed-Bozorgdel A. Novel methodology for prediction of missing values in river flow based on convolution neural networks: Principles and application in Iran country. Phys Chem Earth 2025;138. https://doi.org/10.1016/j.pce.2025.103875.

[24] Hancock JT, Khoshgoftaar TM. CatBoost for big data: an interdisciplinary review. J Big Data 2020;7:94. https://doi.org/10.1186/s40537-020-00369-8.

[25] Tahmouresi MS, Niksokhan MH, Ehsani AH. Enhancing spatial resolution of satellite soil moisture data through stacking ensemble learning techniques. Sci Rep 2024;14:1–25. https://doi.org/10.1038/s41598-024-77050-0.

[26] Ileri K. Comparative analysis of CatBoost, LightGBM, XGBoost, RF, and DT methods optimised with PSO to estimate the number of k-barriers for intrusion detection in wireless sensor networks. Int J Mach Learn Cybern 2025:1–20. https://doi.org/10.1007/s13042-025-02654-5.

[27] Srisuradetchai P, Suksrikran K. Random kernel k-nearest neighbors regression. Front Big Data 2024;7:1402384. https://doi.org/10.3389/fdata.2024.1402384.

[28] Fernández-Delgado M, Sirsat MS, Cernadas E, Alawadi S, Barro S, Febrero-Bande M. An extensive experimental survey of regression methods. Neural Networks 2019;111:11–34. https://doi.org/10.1016/j.neunet.2018.12.010.

[29] Aldin Shojaeezadeh S, Al-Wardy M, Reza Nikoo M. Suspended sediment load modeling using Hydro-Climate variables and Machine learning. J Hydrol 2024;633:130948. https://doi.org/10.1016/j.jhydrol.2024.130948.

[30] Stull T, Ahmari H. Estimation of Suspended Sediment Concentration along the Lower Brazos River Using Satellite Imagery and Machine Learning. Remote Sens 2024;16:1727. https://doi.org/10.3390/rs16101727.

[31] Chen L, Nouri Y, Allahyarsharahi N, Naderpour H, Rezazadeh Eidgahee D, Fakharian P. Optimizing compressive strength prediction in eco-friendly recycled concrete via artificial intelligence models. Multiscale Multidiscip Model Exp Des 2025;8:24. https://doi.org/10.1007/s41939-024-00641-x.

[32] Chen L, Fakharian P, Rezazadeh Eidgahee D, Haji M, Mohammad Alizadeh Arab A, Nouri Y. Axial compressive strength predictive models for recycled aggregate concrete filled circular steel tube columns using ANN, GEP, and MLR. J Build Eng 2023;77:107439. https://doi.org/10.1016/j.jobe.2023.107439.

[33] Dadrasajirlou Y, Ghazvinian H, Heddam S, Ganji M. Reference Evapotranspiration Estimation Using ANN, LSSVM, and M5 Tree Models (Case Study: of Babolsar and Ramsar Regions, Iran). J Soft Comput Civ Eng 2022;6:101–18. https://doi.org/10.22115/scce.2022.342290.1434.

[34] Ziari M, Karami H, Ostadi A, Ghazvinian H. Simulation and prediction of hydraulic jump characteristics over expanding rough beds using FLOW-3D and soft computing techniques. J Hydroinformatics 2025;27:88–106. https://doi.org/10.2166/hydro.2025.270.

[35] Mir AA, Patel M. A Comprehensive Review on Sediment Transport, Flow Dynamics, and Hazards in Steep Channels. J Water Manag Model 2024;32. https://doi.org/10.14796/JWMM.C517.

[36] Komasi M, Alizadefard A, Ahmadi M. Examination of players ' strategies in determining the optimal groundwater exploitation by game theory. Hydrogeol J 2024;32:691–704. https://doi.org/10.1007/s10040-024-02770-6.

[37] Sharghi A, Komasi M, Ahmadi M. Variable sensitivity analysis in groundwater level projections under climate change adopting a hybrid machine learning algorithm. Environ Model Softw 2025;183:106264. https://doi.org/10.1016/j.envsoft.2024.106264.

[38] Shariati M, Saeed M, Behzad M, Fazel G, Masoud A. A novel hybrid extreme learning machine – grey wolf optimizer (ELM - GWO) model to predict compressive strength of concrete with partial replacements for cement. Eng Comput 2020. https://doi.org/10.1007/s00366-020-01081-0.

[39] Delcey M, Cheny Y, Keck JB, Gans A, de Richter SK. Identification of settling velocity with physics informed neural networks for sediment Laden flows. Comput Methods Appl Mech Eng 2024;432:117389. https://doi.org/10.1016/j.cma.2024.117389.