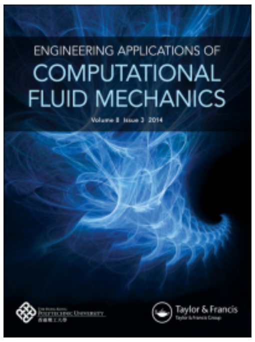


Taylor & Francis
Taylor & Francis Group

**Engineering Applications of Computational Fluid Mechanics**

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tcfm20


# Designing an explainable bio-inspired model for suspended sediment load estimation: eXtreme Gradient Boosting coupled with Marine Predators Algorithm


**Roozbeh Moazenzadeh, Okan Mert Katipoğlu, Ahmadreza Shateri, Hamid Nasiri & Mohammed Abdallah**



**To cite this article:** Roozbeh Moazenzadeh, Okan Mert Katipoğlu, Ahmadreza Shateri, Hamid Nasiri & Mohammed Abdallah (2024) Designing an explainable bio-inspired model for suspended sediment load estimation: eXtreme Gradient Boosting coupled with Marine Predators Algorithm, Engineering Applications of Computational Fluid Mechanics, 18:1, 2391449, DOI: 10.1080/19942060.2024.2391449

**To link to this article:** https://doi.org/10.1080/19942060.2024.2391449





Published online: 20 Aug 2024.

ENGINEERING APPLICATIONS OF COMPUTATIONAL FLUID MECHANICS
2024, VOL. 18, NO. 1, 2391449
https://doi.org/10.1080/19942060.2024.2391449

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

# Designing an explainable bio-inspired model for suspended sediment load estimation: eXtreme Gradient Boosting coupled with Marine Predators Algorithm

Roozbeh Moazenzadeh [a], Okan Mert Katipoğlu [b], Ahmadreza Shateri [c], Hamid Nasiri [d,e] and Mohammed Abdallah [f,g]

[a]Department of Water Engineering, Faculty of Agriculture, Shahrood University of Technology, Shahrood, Iran; [b]Department of Civil Engineering, Erzincan Binali Yıldırım University, Erzincan, Turkey; [c]Electrical and Computer Engineering Department, Semnan University, Semnan, Iran; [d]School of Computing and Communications, Lancaster University, Lancaster, UK; [e]Department of Computer Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran; [f]College of Hydrology and Water Resources, Hohai University, Nanjing, Jiangsu, People's Republic of China; [g]The Hydraulics Research Station, Wad Medani, Sudan

**ABSTRACT**

This study aimed to develop an accurate and reliable model for predicting suspended sediment load (SL) in river systems, which is crucial for water resource management and environmental protection. While Xtreme Gradient Boosting (XGB), a powerful ensemble machine learning (ML) model, has been employed in previous studies, the novelty of this research lies in the introduction of a hybrid approach that synergistically combines XGB with the bio-inspired Marine Predators Algorithm (XGB-MPA) to estimate SL in the Yeşilirmak River (Turkey). To this end, streamflow (Q) and sediment concentration (SC) values as well as their lag times (1 to 3 month lag times) were fed as input variables – under 9 scenarios – into ML models. A time series of datasets from March 1973 to December 2011 and January 2012 to March 2023 were used for training and testing of ML models, respectively. The superiority of the proposed model (XGB-MPA) compared to two other hybrid models, including XGB-PSO (Particle Swarm Optimization) and XGB-GWO (Grey Wolf Optimization) was also investigated. According to the results, the simultaneous application of Q and SC lag time values as inputs has led to the best SL estimates by XGB-MPA, with XGB-MPA9 (RMSE = 103.7 ton/day; NSE = 0.96) exhibiting the lowest error rates. In addition, XGB-MPA has performed better than XGB in all scenarios, with the lowest and highest reduction in RMSE being 19.3% (scenario 5) and 97.4% (scenario 1), respectively. When comparing the performance of hybrid models, the proposed XGB-MPA model has performed best with MAE, RMSE and NSE of 40.94, 103.7 and 0.96, respectively, in comparison with 816.02, 1063.74 and −2.94 for XGB-PSO and 693.16, 981.68 and −2.37 for XGB-GWO. Further research can include the use of time series of efficient variables extracted from satellite images (e.g. land cover, river morphology, etc.) as model inputs.

## Highlights

- Improvement of SL estimation by coupling MPA with XGB model
- Using delayed combinations of streamflow and sediment concentration as model inputs
- Superiority of MPA compare to PSO and GWO in SL estimation
- Greater variations in SHAP caused by sediment concentration compared to streamflow
- Further studies required on the effects of hydrological and topographical features

**ARTICLE HISTORY**
Received 10 April 2024
Accepted 7 August 2024

**KEYWORDS**
Suspended sediment load; Machine learning; Hybrid algorithm; Data-driven models; Sediment concentration; SHAP

## 1. Introduction

Sediment load and sediment concentration (SC) are two of the most important effective parameters in the analysis of hydraulic and hydrodynamic conditions of flow channels (Yaseen et al., 2015). The design of open channels (such as irrigation and drainage channels or urban water networks) and the design, implementation and maintenance of hydraulic structures installed on these channels (such as weirs and channel junctions) are all influenced by proper knowledge of SL (Palazón et al., 2016; Si et al., 2017). A river's SL is significant both from a river engineering perspective (including control of the river's transport capacity, impact on river roughness and flood-carrying capacity of the river, and

**CONTACT** Roozbeh Moazenzadeh romo_sci@shahroodut.ac.ir

Supplemental data for this article can be accessed online at https://doi.org/10.1080/19942060.2024.2391449.

morphological changes) and from an environmental viewpoint (threats posed by sediment accumulation to the life of aquatic organisms) (Buyukyildiz & Kumcu, 2017; Calsamiglia et al., 2018). Another important factor that is strongly influenced by sediment load is water quality (Uca et al., 2018). Sediment accumulation and decreased active capacity of the river affects its role in surface flow drainage and, in the long term, will also have negative effects on watershed hydrology due to reduced streamflow (Rodriguez-Lloveras et al., 2016; Xia et al., 2016).

Traditionally, a variety of experimental and conceptual approaches are utilized for SL prediction. Although experimental methods led to reliable results, these approaches are expensive and need hypotheses in the physical modelling of river flow, which increases the degree of uncertainty of SL prediction by adding errors in parameter values (Ebtehaj et al., 2020; Shamaei & Kaedi, 2016). One widely used empirical technique for calculating sediment is the sediment-rating curve (Tao et al., 2019). For SL computational modelling, mathematical approaches that are based on the physical principles of fluvial sediment movement are frequently employed (Cao & Carling, 2003; Mohammadian et al., 2004). Nevertheless, these approaches frequently demonstrate weak accuracy in estimating since they need a large number of hydro-meteorological and geomorphological data, several of which are unknowable (Hamel et al., 2017; Zhong et al., 2011). Significant progress in soft computing and the field of ML models has improved modelling in such areas as estimation of river SL, which are characterized by stochastic nature, dynamism, and non-linear and non-stationary trends (Allawi et al., 2023; Hanoon et al., 2022; Kisi & Yaseen, 2019; Sharafati et al., 2020).

As a kind of the tree-based ML model, the XGB can compute variable weights and combine multiple weak learners (classifiers or regressors) to construct a robust model. This process significantly diminishes the estimation error, leading to predictions characterized by higher accuracy (Jia et al., 2019). XGB model is recognized as a scalable tree-boosting system with the capability to reduce dimensionality through a cross-validation approach (Chen & Guestrin, 2016). This model has exhibited the highest accuracy in terms of feature selection techniques than common approaches for predicting SL (Bhagat et al., 2022; Piraei et al., 2023). A thorough review of XGB applications in hydrology, hydraulics and hydroclimatology showed that in 74% of the cases, XGB or a hybrid form of this model has outperformed the other ML models (Niazkar et al., 2024). Five classifiers including random forest (RF), XGB, SVM, MLP, and *k*-nearest neighbours were used for SL classification at the Johor River, Malaysia on the daily and weekly time scales (AlDahoul et al., 2022). Their results showed that XGB has exhibited the best classification skills compared to the other classifiers on both daily and weekly scales and in both five- and ten-class scenarios. The capability of three models including XGB, Light Gradient Boosting Machine (LightGBM), and Catboost for SL prediction was compared over 47 gauges across the contiguous United States between 1950 and 2022 using streamflow and climate data as predictors, and these models achieved an average NSE $> 0.66$ during the testing period (Shojaeezadeh et al., 2024).

While XGB has been extolled for its predictive prowess in various hydrological applications, its performance is highly dependent on the selection of optimal hyperparameters. Traditional parameter tuning techniques, such as grid search, exhaustively evaluate all possible parameter combinations, which can be computationally expensive and inefficient, especially in high-dimensional parameter spaces (Bergstra & Bengio, 2012). In this regard, metaheuristic optimization algorithms, inspired by natural phenomena, have emerged as powerful tools for efficiently exploring the parameter space and identifying near-optimal solutions (Abdallah et al., 2023; Samantaray et al., 2022; Sharghi et al., 2019; Talbi, 2009; Zounemat-Kermani et al., 2020). These algorithms leverage various search strategies, such as exploration and exploitation, to effectively navigate the parameter landscape, converging towards global optima while avoiding local minima. Therefore, this study introduced a coupled form of XGB, that is to say, hybridized with a novel bio-inspired optimization algorithm named MPA. By this hybridization, the authors aimed to overcome the limitations of the standalone XGB model by leveraging the strengths of MPA in exploring the parameter space more effectively and efficiently (Burnwal & Jaiswal, 2023). By integrating MPA into the XGB framework, the proposed XGB-MPA model represents a unique contribution to the field, offering a promising alternative for improving the reliability, accuracy, and interpretability of SL predictions. This hybrid approach not only capitalizes on the predictive capabilities of XGB but also harnesses the global optimization potential of the MPA, potentially leading to more robust and generalized models for sediment load estimation. Although this work is the first to investigate the capability of MPA in SL prediction, its effectiveness has been proven in other areas of science. Coupling MPA with a reservoir simulation model, Ngamsert and Kangrang (2022) improved the rule curves, which reduced minimal average water shortage and excess release water deficit by 53 and 19%, respectively. Ikram et al. (2022) showed in their study that the streamflow prediction performance of an ANN model was significantly improved by hybridizing it with MPA.

Studying the effect of weather conditions on the prediction of evaporation in the Dez Dam reservoir, Farzad et al. (2023) evaluated the potential of LSTM coupled with Artificial Bee Colony (ABC), Horse Herd Optimization (HHO) or Marine Predators (MPA) algorithms. Analysis of the results showed that the hybrid LSTM-MPA model is superior to the other models. Zubaidi et al. (2023) compared the performance of an ANN and MPA-ANN in predicting the water demand of Baghdad City, Iraq, and conducting the MPA has improved the performance of the ANN model. The novelty of this study lies in the development and application of the XGB-MPA hybrid model, which, to our knowledge, has not been previously explored for SL prediction. By optimizing XGB parameters with MPA, the authors aimed to overcome the limitations of traditional hyperparameter tuning methods and provide a more robust and accurate predictive tool for water resource management. This innovative approach not only advances the state-of-the-art in SL modelling, but also sets a precedent for future research into the optimization of ML models for complex environmental predictions.

Given the importance of the Yeşilirma River, which serves as a major flood drainage channel in Turkey, this study aimed to contribute to the field by integrating MPA with XGB to enhance the accuracy of SL prediction. In addition, the superiority of the proposed model (XGB-MPA), rather than two other hybrid models (XGB-PSO and XGB-GWO) was also investigated. To this end, streamflow and sediment concentration data and their 1- to 3-month lag times over the 1973–2023 period were used as model inputs. In addition to quantitative analysis of the models using RMSE, MAE, MAPE and NSE indices, the influence of variations in the input parameters on SL was also investigated by the SHAP technique. Limitations of this study and ideas for further research are also discussed.

## 2. Materials and methods

### 2.1. Study area and data used

Yeşilırmak Basin is located between 39° 30′ and 41° 21′ Northern latitudes and 34° 40′ and 39° 48′ Eastern longitudes. It has an area of approximately 38733 km$^2$. The Yeşilırmak River originates from the mountains at an altitude of 2800 m, and its tributaries are Kelkit, Çekerek, Çorum Çat and Tersakan streams. The length of the main river is approximately 519 km (Ozturk et al., 2008). The average annual precipitation is around 443 mm. The precipitation usually falls in May and April. The average yearly temperature is 12°C. The hottest month is July, and the highest temperature is around 40°C. The coldest month is January, with an average temperature of 1.9 and a minimum temperature of −23.4 (DMGM, 2023). The location map of the Yeşilırmak River is shown in Figure 1.

The prominent water structures in the study area are the Almus Dam and Hydroelectric Power Plant, Ataköy Dam, and Hasan Uğurlu Dam. Yeşilırmak River is one of the rivers that carriy the highest amount of sediment to the Black Sea coast in Turkey (Atalay, 1982). Therefore, there is a great deal of development in the river's delta (Kale, 2018). In this respect, estimating sediment transport in the study area is critical.

The dataset used in this study (Table 1) was measured one day in each month in the Yeşilirmak River-Gömeleönü station and it is based on three main variables, including streamflow (m$^3$/s); sediment concentration (ppm); and SL (ton/day). To provide a comprehensive understanding of these variables, we present their statistical characteristics as follows. The mean streamflow is 18.068 m$^3$/s with a standard deviation of 24.287 m$^3$/s, indicating considerable variability in river flow. The median of 8.076 m$^3$/s is notably lower than the mean, reflecting a right-skewed distribution (skewness = 2.781). The kurtosis of 9.889 suggests a peaky distribution with heavy tails. The average sediment concentration is 293.39 ppm, but with a high standard deviation of 735.775 ppm, highlighting significant fluctuations. The data is highly skewed (skewness = 4.934) and has a very high kurtosis of 28.748, indicating extreme values and a very peaky distribution. The mean sediment load is 1462.243 ton/day, with an extremely high standard deviation of 6012.068 ton/day. The median of 33.35 tons/day is much lower than the mean, showing a strong right skew (skewness = 7.542). The kurtosis value of 69.15 indicates a highly peaked distribution with very heavy tails. The statistical characteristics of these variables highlight the complexities and variabilities inherent in hydrological data. The significant differences between the means and medians, along with high skewness and kurtosis values across the variables, underscore the challenges in modelling such data and the necessity of sophisticated machine learning techniques like the XGB-MPA approach for accurate sediment load estimation. Figure 2 illustrates sediment rating curves in a logarithmic scale, which shows the relationship between SL and $Q$ samples and the trend line's slope indicates the nature of the relationship between sediment load and streamflow.

### 2.2. Extreme gradient boosting (XGB)

XGB is a state-of-the-art ML algorithm that has gained popularity for its effectiveness in handling various types of predictive modelling tasks (González et al., 2020; Nguyen et al., 2021). It is an implementation of gradient

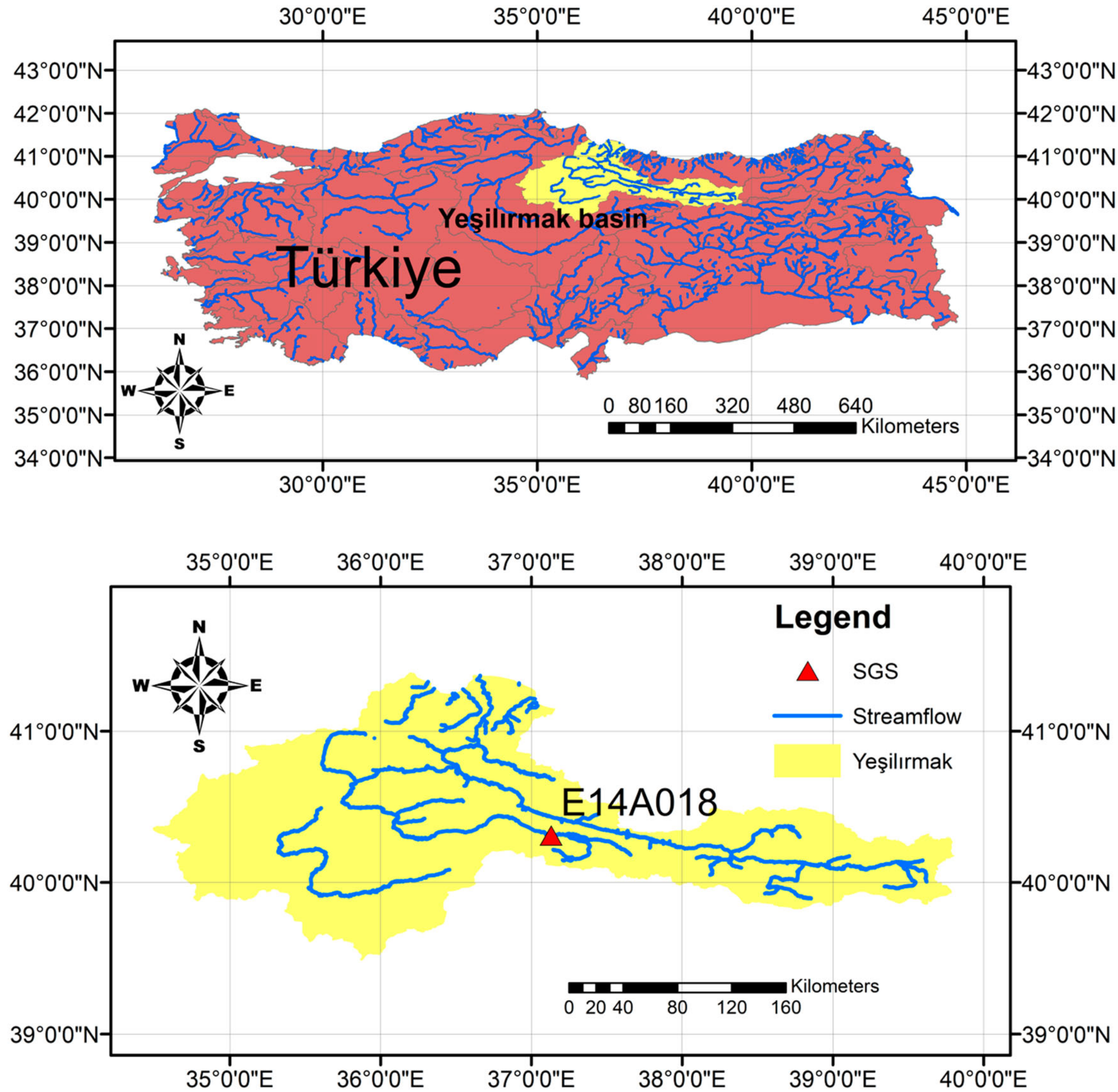


**Figure 1.** Location of the Yeşilırmak River in Turkey and Yeşilırmak River basin map with SGS.

**Table 1.** The statistical characteristics of the measured data (*Q*, SC, and SL).

| Variable (unit) | Mean | Standard deviation | Median | Minimum | Maximum | Skew | Kurtosis |
|---|---|---|---|---|---|---|---|
| *All data* | | | | | | | |
| *Q* ($m^3/s$) | 18.068 | 24.287 | 8.076 | 0.8 | 188.8 | 2.781 | 9.889 |
| SC (ppm) | 293.389 | 735.775 | 51.65 | 2.7 | 6084.8 | 4.934 | 28.748 |
| SL (ton/day) | 1462.243 | 6012.068 | 33.35 | 1.4 | 76673.7 | 7.542 | 69.15 |
| *Training data* | | | | | | | |
| *Q* ($m^3/s$) | 19.561 | 25.602 | 8.711 | 0.8 | 188.8 | 2.652 | 8.922 |
| SC (ppm) | 348.526 | 810.216 | 58.25 | 2.7 | 6084.8 | 4.397 | 22.62 |
| SL (ton/day) | 1777.859 | 6661.766 | 39.5 | 1.4 | 76673.7 | 6.753 | 55.213 |
| *Testing data* | | | | | | | |
| *Q* ($m^3/s$) | 11.936 | 16.634 | 5.396 | 1.197 | 94.506 | 3.011 | 10.242 |
| SC (ppm) | 66.824 | 75.783 | 45.875 | 8.4 | 458.38 | 3.323 | 12.365 |
| SL (ton/day) | 165.348 | 544.389 | 15.61 | 2.44 | 3657.7 | 5.25 | 29.089 |

boosting machines (GBM) and is designed to optimize both speed and model performance (Chen & Guestrin, 2016). The core idea of this model revolves around the concept of boosting, which involves combining multiple weak predictive models to create a strong predictive model (Zhang & Haghani, 2015). The boosting aspect in

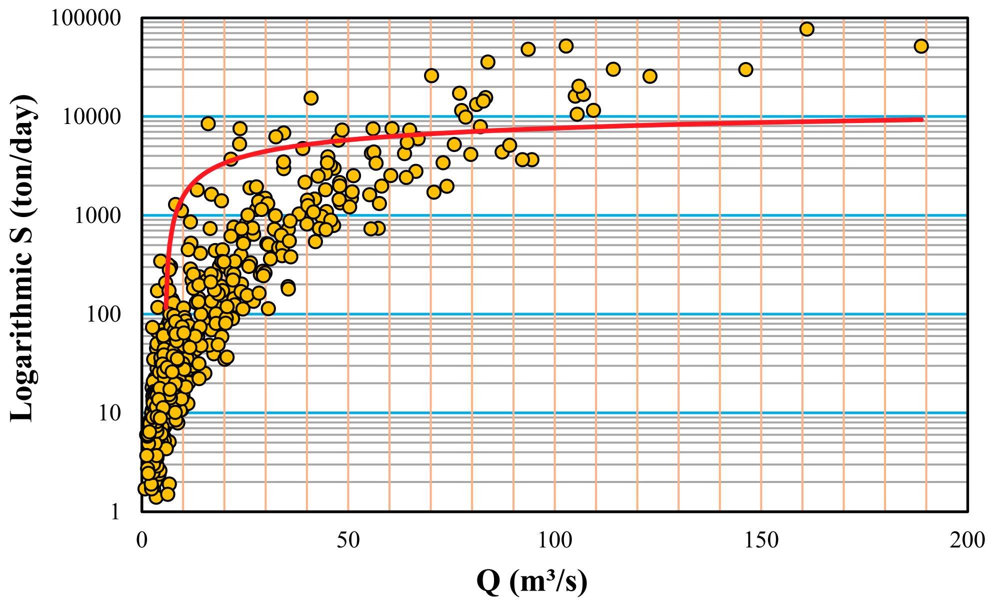


**Figure 2.** Sediment rating curves (logarithmic plot) of measured SL and Q samples with the logarithmic trend line (red line).

XGB involves sequentially adding decision trees to the model. Each new tree is created to correct the errors made by the previous trees. Unlike random forests, which build trees independently, gradient boosting builds them sequentially (Ghaheri et al., 2023). In XGB, the residual errors from the previous trees are used to train the subsequent tree (Ayyadevara, 2018). This model involves calculating the gradient (the first derivative) of the loss function (a measure of how far the model's predictions are from the actual values) to determine the direction in which the model's performance is improving the most rapidly (Patnaik et al., 2021). XGB comes with several hyperparameters that can be tuned to optimize the model's performance. These include the learning rate, the maximum depth of the trees, the minimum child weight, and the subsample ratio of the training instances (Kavzoglu & Teke, 2022). Tuning these hyperparameters is crucial to achieve the best performance from an XGB model (Yagin et al., 2024). In our study, the XGB model serves as the foundation upon which we build our hybrid XGB-MPA approach.

### 2.3. Marine predators algorithm (MPA)

MPA is a novel optimization algorithm inspired by the foraging behaviour of marine predators. This algorithm mimics the strategies employed by marine animals in locating and capturing prey, translating these behaviours into a mathematical framework to solve optimization problems (Tu et al., 2023). It is based on the concept of marine predators' foraging behaviour, which involves searching, encircling, and attacking prey. These behaviours are mathematically modelled to explore and exploit the search space in an optimization problem efficiently (Faramarzi et al., 2020). In the initial phase of the algorithm, marine predators explore the search space. This phase is characterized by random movements to scan different areas, analogous to predators searching for prey in various locations. This exploration allows the MPA to avoid local minima and ensures a diverse search of the solution space. Once potential solutions (prey) are identified, the MPA shifts to the exploitation phase, where the algorithm focuses on refining these solutions. This phase mimics the behaviour of predators narrowing down on their prey and includes strategies such as encircling the prey and spiral bubble-net attacking method (Tu et al., 2023). This approach helps in fine-tuning the solutions to reach the global optimum. MPA dynamically adapts between exploration and exploitation phases based on the current state of the search. This adaptation is crucial for maintaining a balance between exploring new areas and exploiting known good solutions, leading to effective optimization (Sekhar et al., 2023). Finally, The algorithm searches for the optimal set of hyperparameters, such as learning rate, tree depth, and regularization terms, which minimizes the error in SL estimation.

In this study, to implement the XGB-MPA model, we utilized the MEALPY Python library (Van Thieu & Mirjalili, 2023) for the MPA algorithm, known for its extensive collection of population-based meta-heuristic algorithms. Additionally, Python's XGBoost package was employed for the implementation of the XGB algorithm. Figure 3 shows the modelling flowchart for MPA optimization procedure as used in this study.

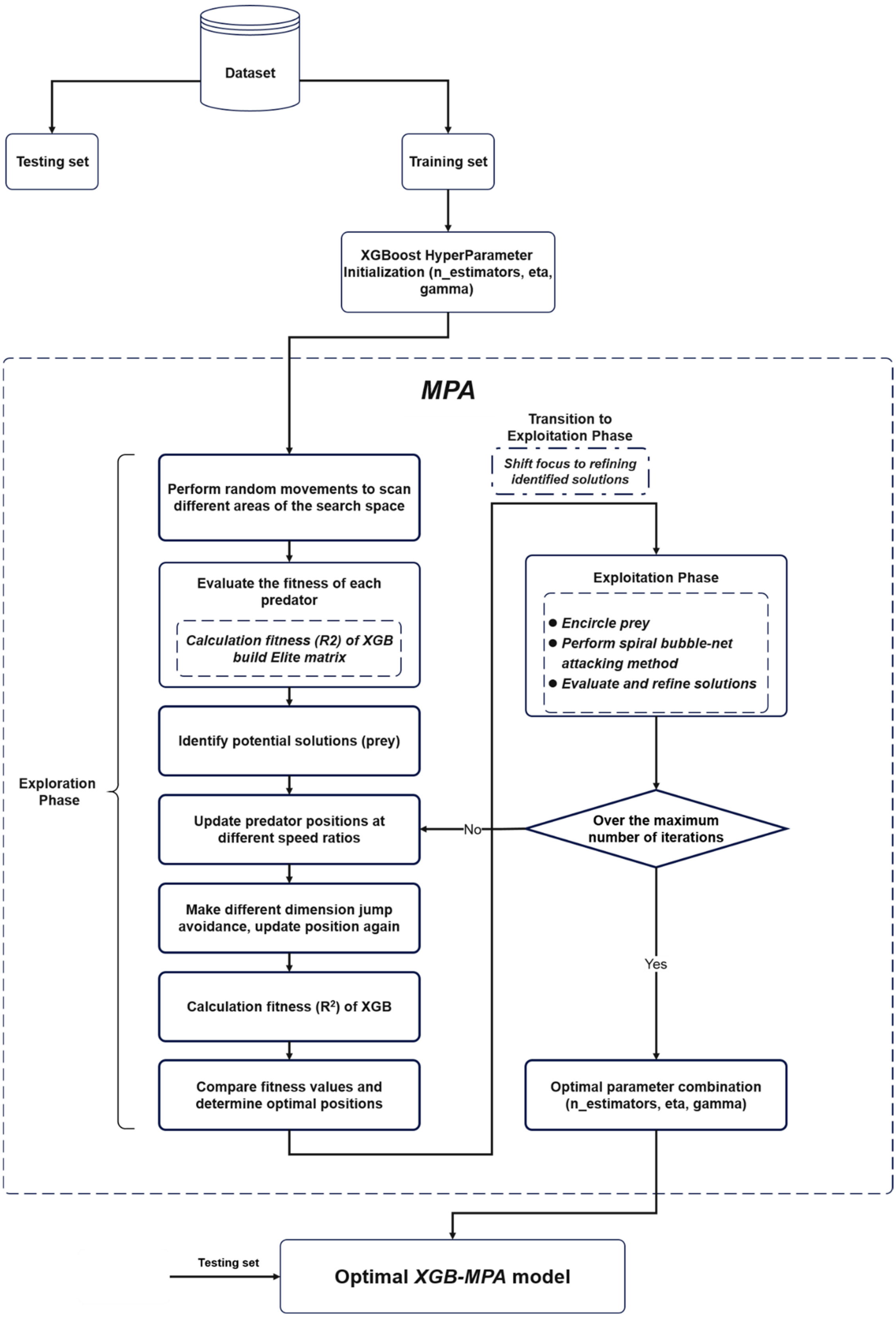


**Figure 3.** Flowchart of Marine Predators Algorithm.

**Table 2.** Model input combinations based on different lag times of *Q* and SC.

| No. | Input(s) | Output | Models | |
|---|---|---|---|---|
| 1 | $Qt$ | $SLt$ | XGB1 | XGB-MPA1 |
| 2 | $Qt, Qt-1$ | $SLt$ | XGB2 | XGB-MPA2 |
| 3 | $Qt, Qt-1, Qt-2$ | $SLt$ | XGB3 | XGB-MPA3 |
| 4 | $Qt, Qt-1, Qt-2, Qt-3$ | $SLt$ | XGB4 | XGB-MPA4 |
| 5 | $SCt$ | $SLt$ | XGB5 | XGB-MPA5 |
| 6 | $SCt, SCt-1$ | $SLt$ | XGB6 | XGB-MPA6 |
| 7 | $SCt, SCt-1, SCt-2$ | $SLt$ | XGB7 | XGB-MPA7 |
| 8 | $SCt, SCt-1, SCt-2, SCt-3$ | $SLt$ | XGB8 | XGB-MPA8 |
| 9 | $Qt, SCt-1, SCt-3$ | $SLt$ | XGB9 | XGB-MPA9 |

*Q*: Streamflow ($m^3$/s), SC: Sediment concentration (ppm), SL: Sediment load (ton/day).

**Table 3.** Statistical indices of XGB model under nine scenarios in testing phase.

| Model | $R^2$ | MAE (ton/day) | RMSE (ton/day) | MAPE | NSE |
|---|---|---|---|---|---|
| XGB1 | 0.85 | 934.5 | 5793.3 | 1.69 | −113.47 |
| XGB2 | 0.81 | 791.9 | 4398.3 | 3.11 | −65.39 |
| XGB3 | 0.77 | 759.2 | 3866.1 | 4.36 | −50.79 |
| XGB4 | 0.79 | 747.9 | 3545.5 | 5.03 | −42.33 |
| XGB5 | 0.87 | 102.46 | 382.2 | 1.11 | 0.50 |
| XGB6 | 0.82 | 110.9 | 413.85 | 1.02 | 0.41 |
| XGB7 | 0.84 | 111.4 | 404.01 | 1.13 | 0.44 |
| XGB8 | 0.85 | 111.9 | 399.33 | 1.16 | 0.45 |
| XGB9 | 0.92 | 365.2 | 1247.68 | 3.04 | −4.3 |

### 2.4. Selection of model inputs and definition of scenarios

The input scenarios and data size used in hydrological predictions with ML models are of great importance for the learning of the model and revealing its high prediction potential. In this study, the performance of 9 scenarios for SL prediction was investigated (Table 2). These scenarios were created with current and historical streamflow and sediment concentration values. Spearman's correlation coefficients were calculated to be about 0.93, 0.55, 0.28, −0.26 and −0.17 between SL and SC data and its lag times (1-,2-,3-,4-, and 5-month lag times), respectively. Corresponding values for Q data and its lag times were 0.89, 0.63, 0.3, −0.35 and −0.22. Therefore, Q and SC and their lag times of up to three months were chosen as model inputs. To this end, predictions were first made only with combinations of Q with scenarios 1–4 and of SC with scenarios 5–8, and finally, the prediction inputs were evaluated with the best combinations of these two variables with scenario 9. Here, $t$ shows real-time, while $t-1$ shows one-month lag time. The dataset consisted of a total of 562 data points (March 1973 to March 2023), of which 452 data points (approximately 80%, March 1973 until December 2011) were used to train the ML models and the remaining 110 (approximately 20%, Jan 2012 until March 2023) were reserved for model testing.

### 2.5. Evaluation criteria

In this study, five statistical indices including coefficient of determination ($R^2$), mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE) and Nash-Sutcliffe efficiency (NSE) were used (Equations 1–5). $R^2$ represents the proportion of variance for the dependent variable (SL in this case) that is explained by the independent variables in the model ($0 \leq R^2 \leq 1$, optimum = 1). MAE calculates the average magnitude of the errors in a set of predictions, without considering their direction ($0 \leq \text{MAE} \leq \infty$, optimum = 0). RMSE compares the measured and predicted values ($0 \leq \text{MAE} \leq \infty$, optimum = 0). MAPE represents the mean absolute percentage deviation between measured and predicted values ($0 \leq \text{MAPE} \leq \infty$, optimum = 0). NSE varies between $-\infty$ and 1 (perfect fit).

$$R^2 = \frac{\left(\sum_{i=1}^{n}(\text{SLi}_{\text{obs}} - \overline{\text{SLi}_{\text{obs}}})(\text{SLi}_{\text{est}} - \overline{\text{SLi}_{\text{est}}})\right)^2}{\sum_{i=1}^{n}(\text{SLi}_{\text{obs}} - \overline{\text{SLi}_{\text{obs}}})^2 \sum_{i=1}^{n}(\text{SLi}_{\text{est}} - \overline{\text{SLi}_{\text{est}}})^2} \tag{1}$$

$$\text{MAE} = \frac{1}{n}|\text{SLi}_{\text{obs}} - \text{SLi}_{\text{est}}| \tag{2}$$

$$\text{RMSE} = \sqrt{\frac{(\text{SLi}_{\text{obs}} - \text{SLi}_{\text{est}})^2}{n}} \tag{3}$$

$$\text{MAPE} = \frac{\sum_{i=1}^{n}|\text{SLi}_{\text{est}} - \text{SLi}_{\text{obs}}/\text{SLi}_{\text{obs}}|}{n} \tag{4}$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{n}(\text{SLi}_{\text{obs}} - \text{SLi}_{\text{est}})^2}{\sum_{i=1}^{n}(\text{SLi}_{\text{obs}} - \overline{\text{SLi}_{\text{obs}}})^2} \tag{5}$$

where, $n$ represents the number of data, $i$ indicates counter variable, $\text{SL}_{\text{iobs}}$ is observational data, $\overline{\text{SL}_{\text{iobs}}}$ is average of the observational data, $\text{SL}_{\text{iest}}$ represents computational data, and $\overline{\text{SL}_{\text{iest}}}$ is average predicted data.

## 3. Results and discussion

### 3.1. XGB and XGB-MPA performance

The average prediction performance of the XGB model and its scatter plot, both in the testing phase, are listed and shown in Table 3 and Figure 4, respectively. According to the results of Table 3, it was concluded that the XGB5 model with assigning values of102.46, 382.23, 1.11 and 0.5 for the MAE, RMSE, MAPE and NSE, respectively, outperformed the other models, reflecting the importance of SC values in SL estimation.

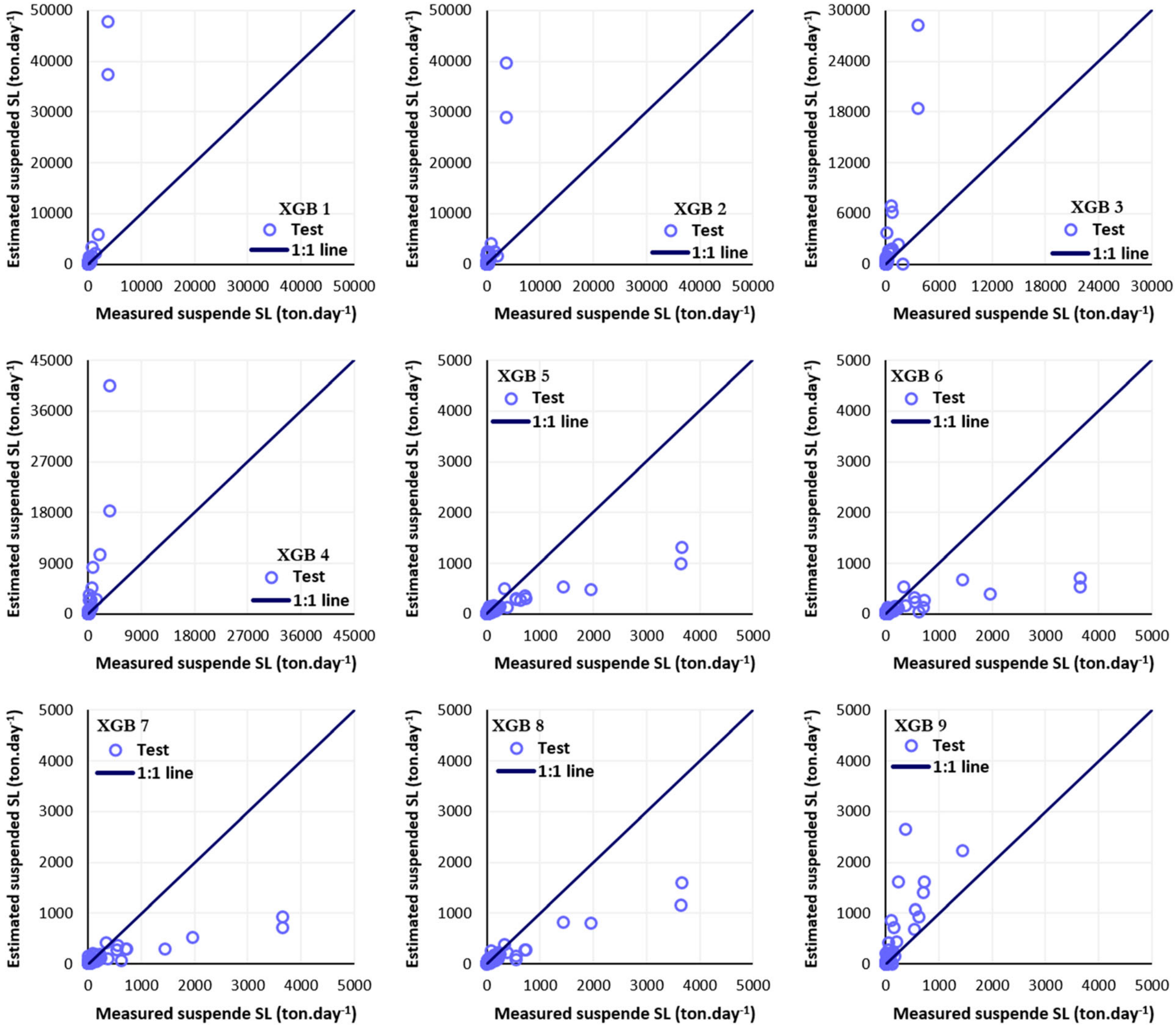


**Figure 4.** Measured SL values versus the values predicted by the XGB model under nine defined scenarios in testing phase.

In the training phase of the XGB model, the addition of lag times (either of streamflow discharge or sediment concentration) has improved the model's performance in estimating sediment load, with 1, 2 and 3-lag times of streamflow discharge reducing error rates by approximately 38.5, 17.5 and 13%, and those of sediment concentration reducing error rates by 49, 17 and 18%. In the training phase of this model, scenario 8 exhibited the best performance in estimation of sediment load (MAE = 27.32 ton/day). The testing phase of the XGB model responded differently to the use of streamflow discharge and sediment concentration, in terms of sediment load estimation. The results showed that the addition of 1, 2 and 3-lag times of streamflow discharge to the model inputs has reduced estimation error by about 15, 4 and 1.5%, respectively. However, the error rates of these scenarios are relatively high (748–934 ton/day), which is mainly due to the poor performance of the model in estimating two particular values (Figure 4). Another important point is that the performance of the model in estimating sediment load for the two aforesaid values improves by increasing the number of model inputs (from scenario 1 to scenario 4), so that, for example, the overall model error is 20% lower under scenario 4 compared to scenario 1. With the same model and when sediment concentration is added to model inputs with 1, 2 and 3-lag times, although error rates increase by about 8.2, 0.48 and 0.44%, respectively (meaning that lag times of sediment concentration have not affected model performance), but the overall error rates of the model for different scenarios vary between 102.4 and 111.9 ton/day which is considered to be appropriate.

For XGB model in the testing phase, the scenarios including SC have outperformed the scenarios including Q in evaluating SL – as was the case in training phase – and the XGB5 model (MAE = 102.46; NSE = 0.5) has

**Table 4.** Statistical indices of XGB-MPA model under nine scenarios in testing phase.

| Model | $R^2$ | MAE (ton/day) | RMSE (ton/day) | MAPE | NSE |
|---|---|---|---|---|---|
| XGB-MPA1 | 0.95 | 53.9 | 150.45 | 0.56 | 0.9 |
| XGB-MPA2 | 0.93 | 58.3 | 173.43 | 0.58 | 0.84 |
| XGB-MPA3 | 0.92 | 55.46 | 152.79 | 0.59 | 0.92 |
| XGB-MPA4 | 0.91 | 57.92 | 162.82 | 0.56 | 0.91 |
| XGB-MPA5 | 0.93 | 91.47 | 308.26 | 1.43 | 0.67 |
| XGB-MPA6 | 0.90 | 96.75 | 305.05 | 1.55 | 0.68 |
| XGB-MPA7 | 0.89 | 93.73 | 272.2 | 1.66 | 0.74 |
| XGB-MPA8 | 0.89 | 99.79 | 279.59 | 2.22 | 0.73 |
| XGB-MPA9 | 0.97 | 40.94 | 103.7 | 0.47 | 0.96 |

exhibited the lowest error rate. Also, the hybrid scenario (the scenario which used Q and two lag times of SC as inputs; XGB9 model) has outperformed the scenarios, which included Q and its lag times (models XGB1 through XGB4), whereas it has led to higher estimation errors in comparison with models XGB5 through XGB8. Ahmadi et al. (2022) used gene expression programming (GEP), extreme learning machine (ELM), and XGB algorithms with empirical mode decomposition (EMD) and ensemble EMD to predict SL at two stations in the Tajan Basin in northern Iran. According to the analysis, the most superior model is the EEMD-XGB hybrid model with $R^2$ higher than 0.9. AlDahoul et al. (2021) compared the Long Short Term Memory (LSTM) model, ElasticNet, Linear Regression, Multilayer Perceptron neural network (MLP) and XGB models for prediction of SC in Johor River, Malaysia. According to the analysis, LSTM has the highest accuracy with regression values of 92.01%, 96.56%, 96.71% and 99.45% in daily, weekly, 10-day and monthly scenarios, respectively.

The test prediction results of the XGB-MPA model are shown in Table 4. Measured versus predicted SL values for all nine models in testing phases are depicted in Figure 5. According to XGB-MPA results, the scenario which simultaneously used $Q$ and two lag times of SC as inputs (XGB-MPA9 model) has outperformed the other scenarios in SL estimation, with $R^2$, MAE, RMSE, MAPE and NSE of 0.97, 40.94, 103.7, 0.47 and 0.96, respectively.

In the training phase of the XGB-MPA model, addition of lag-timed streamflow discharge values has not exerted much effect on the improvement of model performance, which stands in contrast to the addition of sediment concentration lag times, especially in scenarios 6 and 7 where estimation error is reduced by 49 and 37%, respectively. For the same model and in the testing phase, model performance has exhibited the same behaviour when lag times of streamflow discharge and sediment concentration are added, with 1 and 3-lag times of streamflow discharge increasing the error rate by about 8 and 4.5% and corresponding lag times of sediment concentration increasing estimation error by about 5.7 and 6.4%; and adding the 2-lag time of discharge and sediment concentration has reduced SL estimation error by about 5 and 3%, respectively. A point to consider regarding XGB-MPA results in the testing phase is the lower error rates (53.9, 58.26, 55.46 and 57.9 ton/day for scenarios 1 to 4, respectively) when streamflow discharge is used as input, compared to corresponding scenarios which have used sediment concentration as input and estimated the sediment load with error rates of 91.47, 96.75, 93.7 and 99.8 ton/day, respectively. In this phase, scenario 9, which includes both Q and SC lag times as model inputs, has performed best with MAE $= 40.94$ ton/day. In the XGB-MPA model, evaluation of different scenarios that include Q and its time lags as model inputs showed that XGB-MPA1 has performed best with MAE, RMSE, MAPE and NSE of 53.9, 150.45, 0.56 and 0.9, respectively. However, performance comparison among the XGB-MPA2 and XGB-MPA4 models revealed that XGB-MPA3 has had the lowest error rate (RMSE $= 152.79$), indicating that addition of the 3-month time lag of $Q$ ($Qt-3$) to the inputs of the model has not contributed to the improvement of SL estimation. For the XGB-MPA model and scenarios that used SC as input, XGB-MPA7 has delivered an overall finer performance with RMSE, MAE and NSE of 272.2, 93.7 and0.74, respectively.

To provide a clear comparative analysis of the models, two bar plots are prepared as shown in Figure 6. These graphs illustrate the performance metrics ($R^2$ and MAE values in the testing set) of all models (XGB1 to XGB9 and XGB-MPA1 to XGB-MPA9) in a more visually accessible format. These visualizations allow for an immediate understanding of which model outperforms the others, enhancing the clarity of our comparative analysis. Apart from the better performance of scenarios 5 and 9 relative to the other scenarios (for XGB and XGB-MPA, respectively), Figure 6 shows that SC and its time lags – as inputs – for XGB, and Q and its time lags for XGB-MPA have played a more influential role in estimating SL. However, the dependence of SL estimation accuracy on the input type is clearer for XGB. In other words, the performance of XGB-MPA does not show considerable variations when the type of input (SC or Q) is changed, and its error rates remain relatively low, which is regarded as an advantage for this hybrid model.

For a more precise analysis of the results – based on the distribution of points in Figures 4 and 5 – performance of XGB4 and XGB5 models, and performance of XGB-MPA1, XGB-MPA7 and XGB-MPA9 (both in testing set) were investigated over various ranges of measured SL values (Figure 7). Accordingly, it can be concluded that in general, higher measured SL values have led to higher estimation errors by both the base and the hybrid models.

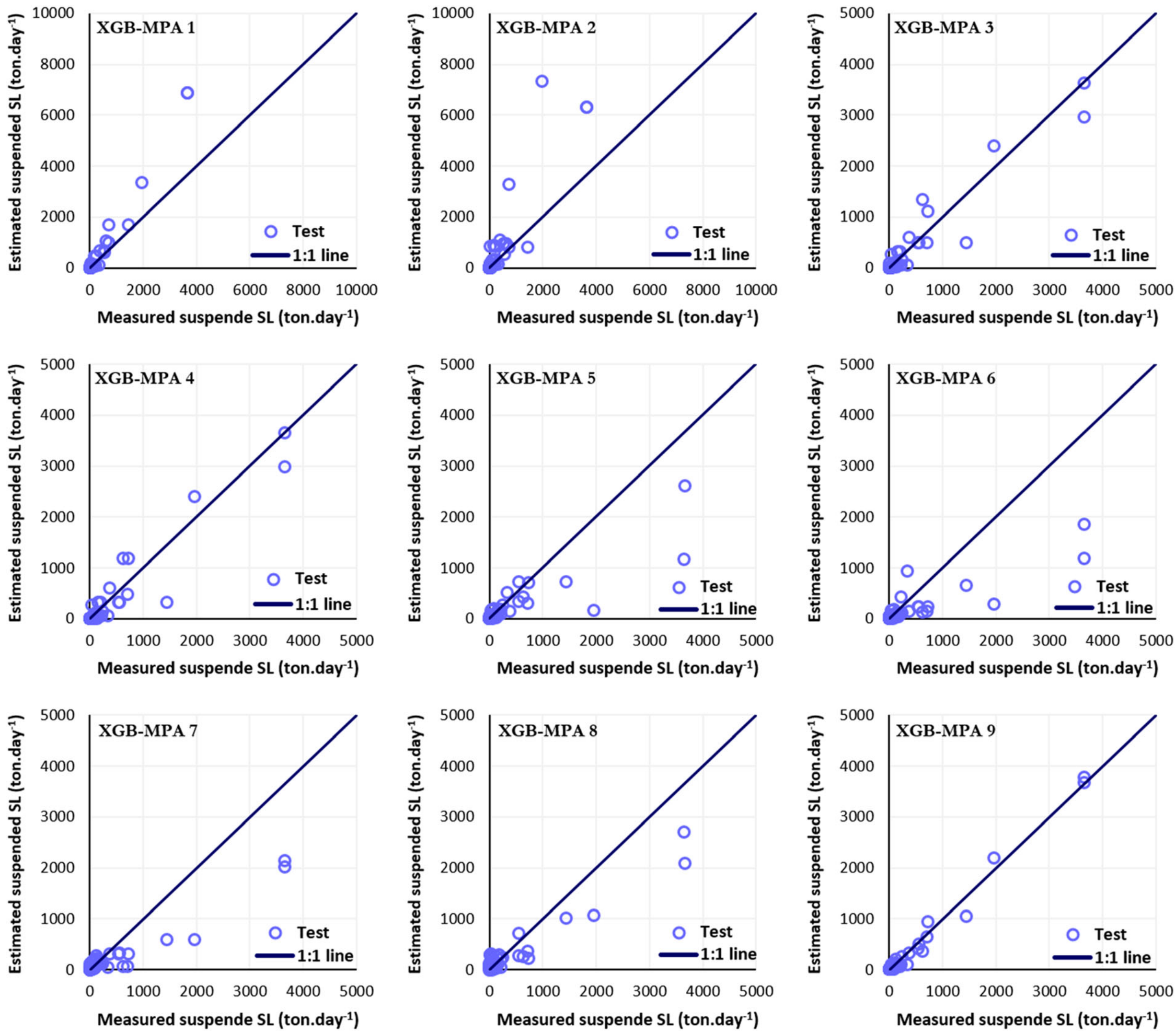


**Figure 5.** Measured SL values versus the values predicted by the XGB-MPA model under nine defined scenarios in testing phase.

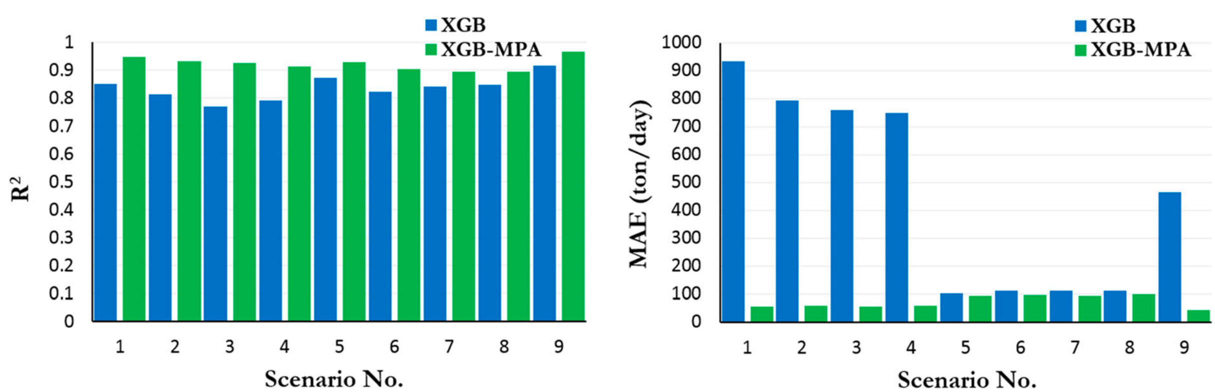


**Figure 6.** Bar plot of $R^2$ and MAE values for XGB and XGB-MPA models in all 9 scenarios in testing phase.

For the base model, a comparison between the best scenario based on Q and its lag times (XGB4) and the best scenario based on SC and its lag times (XGB5) indicates the noticeably better performance of the XGB5 model, such that it has led to an approximately 90% reduction of RMSE in both SL ranges of less than and more than 1000 ton/day compared with the XGB4 model. In the case of the hybrid model, comparison between the best scenario based on Q and its lag times (XGB-MPA1), the best scenario based on SC and its lag times (XGB-MPA7)

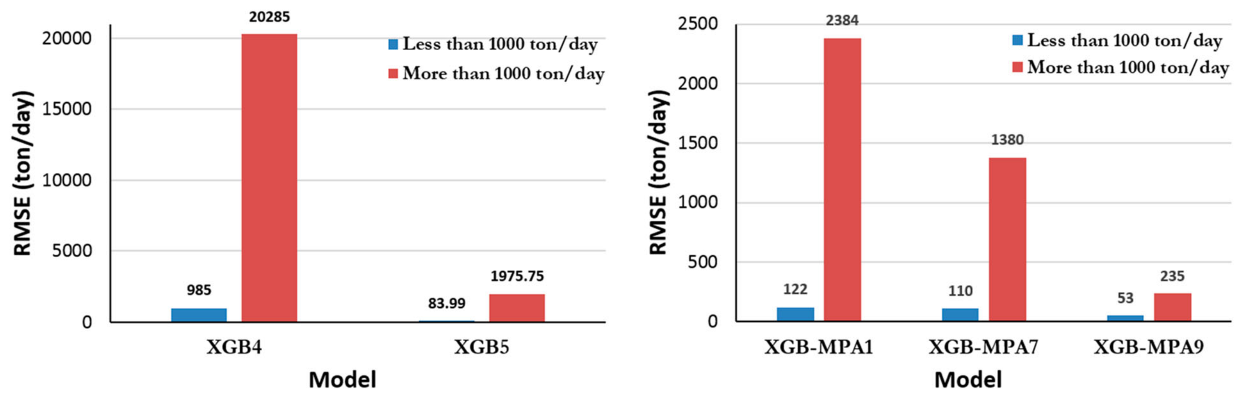


**Figure 7.** RMSEs of different models over various intervals of measured SL values in testing phases.

and the best hybrid scenario (XGB-MPA9) shows the noticeable superiority of the XGB-MPA9 model which has reduced estimation error by 56.5 and 52% for SL values lower than 1000 tons/day and by 90 and 83% for SL values higher than 1000 tons/day compared to XGB-MPA1 and XGB-MPA7 models, respectively. The best hybrid model (XGB-MPA9) has led to increased error rates for measured SL values greater than 1000 tons/day relative to those below 1000 tons/day; however, the increment is much less compared with XGB-MPA1 and XGB-MPA7. Poor performance in estimating higher measured SL values is thus one of the limitations of these models; although it could be circumvented to a large extent by changing the scenario (for either base models or hybrid models) and changing the model (from the base to the hybrid model).

A new SL forecasting framework was introduced by integrating seasonal adjustment and Bayesian optimization into ML models (Li & Yang, 2022). According to their results, a significant improvement occurred in Boosting model where RMSE, MAE and NSE values varied between (0.44–0.59) and (0.28–0.36) ton/s and (0.91–0.95), respectively. Six ML models including iterative absolute error regression coupled with alternating model tree (IAER-AMT) and dual perturb and combine tree (IAER-DPCT), AMT, DPCT, long short-term memory coupled with grey wolf optimizer (LSTM-GWO) and recurrent neural network coupled with GWO (RNN-GWO) as well as two empirical models, Einstein (1950) and Recking (2013), were examined for prediction of fluvial bedload transport (Khosravi et al., 2024). RMSEs (g/m/s) were 4.48, 4.93, 5.23, 5.3, 5.58, 5.78, 81.37 and 83.3, whereas NSEs were 0.8, 0.76, 0.73, 0.72, 0.69, 0.67, −63.87 and −67 for the aforementioned models, respectively.

A hybridized and integrated form of MLP with PSO and differential evolution algorithm (MLP-PSODE) was implemented for SL prediction in Mahabad River, northwest of Iran (Mohammadi et al., 2021). With four streamflow lag times as inputs, RMSE (ton/day), MAE (ton/day) and MAPE varied between (1794.407–2118.393), (1124.608–1354.879) and (0.415–0.56) in testing phase, respectively.

**Table 5.** Parameters setting and their values for the base (XGB) and hybrid (XGB-MPA) models.

| XGB | | |
|---|---|---|
| Parameter | Value | |
| Base learner | dart | |
| Subsample | | 0.5 |
| Learning rate ($\eta$) | | 0.3 |
| Lagrange multiplier ($\gamma$) | | 0 |
| Maximum depth of trees | | 6 |
| **XGB-MPA** | | |
| Hyper-parameter | Domain range | Optimal value |
| Base learner | ('gblinear', 'gbtree') | Gradient boosted tree |
| Learning rate | [0, … , 1] | $2.14 \times 10^{-3}$ |
| n_estimator | [2, … , 300] | 142 |
| gamma ($\gamma$) | [0, … , 50] | 49.79 |

The XGB-MPA ensemble invites a good deal of hyperparameter tuning. To address this, we focused on the most sensitive hyperparameters. To set up the model, the Gradient-boosted tree (gbtree) was applied as the base learner due to its superior performance. Only three sensitive parameters, namely the learning rate the number of estimators and Lagrange multiplier ($\gamma$) had to be optimized. The optimized values of the XGB-MPA9 and their domain range are shown in Table 5.

To justify the selection of models for this study (XGB and XGB-MPA as the base and the hybrid model, respectively), a performance comparison was made between them and 4 other models (2 base and 2 hybrid models). To this end, we conducted additional experiments using Artificial Neural Networks (ANN) and Random Forests (RF) as baseline ML models for SL estimation. The performance of these models was assessed using the same dataset and metrics, enabling a direct comparison of their predictive accuracy. The results in Table 6 indicate that the XGB5 model outperforms both the ANN

**Table 6.** Statistical indices of proposed XGB and two other base models under scenario 5 in testing phase.

| Models | $R^2$ | MAE (ton/day) | RMSE (ton/day) | NSE |
|---|---|---|---|---|
| ANN5 | 0.85 | 193.21 | 269.55 | 0.75 |
| RF5 | 0.85 | 104.6 | 392 | 0.47 |
| XGB5 | 0.87 | 102.46 | 382.2 | 0.5 |

**Table 7.** Statistical indices of proposed XGB-MPA and two other hybrid models under scenario 9 in testing phase.

| Models | $R^2$ | MAE (ton/day) | RMSE (ton/day) | MAPE | NSE |
|---|---|---|---|---|---|
| XGB-PSO9 | 0.85 | 816.02 | 1063.74 | 57.25 | −2.94 |
| XGB-GWO9 | 0.86 | 693.16 | 981.68 | 45.94 | −2.37 |
| XGB-MPA9 | 0.97 | 40.94 | 103.7 | 0.47 | 0.96 |

and RF models, demonstrating its superior predictive accuracy for SL estimation. In addition, to justify the use of MPA compared to other methods such as GWO and PSO, additional experiments have been conducted to evaluate the performance of the XGB model when optimized with these algorithms. Table 7 summarizes the statistical indices values for the three hybrid models: XGB-MPA, XGB-PSO, and XGB-GWO. The results given in Table 7 indicate that the proposed XGB-MPA9 model has achieved a markedly superior accuracy in predictions compared with the two other hybrid models – which have completely identical inputs – such that its MAE and RMSE have been, respectively, about 95 and 90% lower relative to the XGB-PSO9 model and about 94 and 89% lower in comparison with the XGB-GWO9 model. These findings underscore the efficacy of the MPA in enhancing the predictive performance of the XGB model.

According to the performance of a hybrid model under different scenarios, residual values of SL estimation are calculated in scenarios 1 (the best scenario when Q data are used as inputs), 7 (the best scenario when SC data are used as inputs) and 9 (the hybrid scenario), with the results depicted in Figure 8. Given that the majority of measured values are below 1000 ton/day, the figures on the right only show the distributions in that range. As measured SL values increase, higher SL estimation errors and a tendency towards underestimation are observed for all three models. In case of measured SL values between 0 and 200 tons/day, although the number of data points in the underestimation set (n = 81) is much more than the overestimation set (n = 15), XGB-MPA1 model has exhibited a lower error rate in the underestimation set (RMSE = 24.4 tons/day) compared to the overestimation set (RMSE = 88.3 tons/day), one of the main reasons of which being the accumulation of a large number of points with residual values below −50 mm/day in underestimation set.

Consistent with the results given in Table 4, the overall performance of XGB-MPA7 is slightly poorer relative to XGB-MPA1, which is caused not by the performance of the two models across the 0–200 ton/day range, but by the poorer performance of XGB-MPA7 across the 400–800 tons/day range where it has had greater residual values. One of the advantages of the XGB-MPA9 model (hybrid scenario) across the 0–200 tons/day range of SL values, compared with the other two models, is its relatively consistent performance in underestimation set (n = 71; RMSE = 24 ton/day) and overestimation set (n = 25; RMSE = 25 ton/day). For measured SL in the range 400–800 tons/day and in underestimation set, XGB-MPA1 (RMSE = 145 ton/day) and XGB-MPA9 (RMSE = 153 tons/day) have shown a relatively similar and much better performance compared to XGB-MPA7 (RMSE = 400 tons/day). Therefore, one of the main reasons for superiority of the XGB-MPA9 model over XGB-MPA1 is the difference in the performance of the two models in the 0–200 tons/day range (overestimation set).

Taylor diagram is a graphical comparison to determine, which model produces the most realistic result (Taylor, 2001). It condenses statistical comparison between models and observations into a single diagram, with the observed data represented by a point on the x-axis. The radial distance from the origin indicates the standard deviation of each model, while the azimuthal position gives the correlation coefficient between the model and observations. Figure 9 shows the Taylor diagram of the test data and the performance comparison of the XGB and XGB-MPA models. Accordingly, the best model is XGB-MPA9 because it has the lowest mean square error and a standard deviation close to the real data. Additionally, it can be inferred that each hybrid model produces more accurate outputs than the single XGB model in the same combination.

Comparing the normalized standard deviation (NSD) between the real data and models provides another assessment. When the NSD values of the real data and prediction models are compared, the XGB-MPA1, XGB-MPA3 and XGB-MPA9 algorithms are promising to show an NSD similar to the real data. When correlation coefficients are examined, XGB-MPA9 is the best model, with a value equal to 0.97, while correlation coefficients for the other models are generally between 0.95 and 0.99.

### 3.2. Sensitivity analysis

This section provides a sensitivity analysis of the proposed XGB-MPA9 model to evaluate the sensitivity of the model to the population size parameter in the MPA. We conducted experiments to observe how varying the

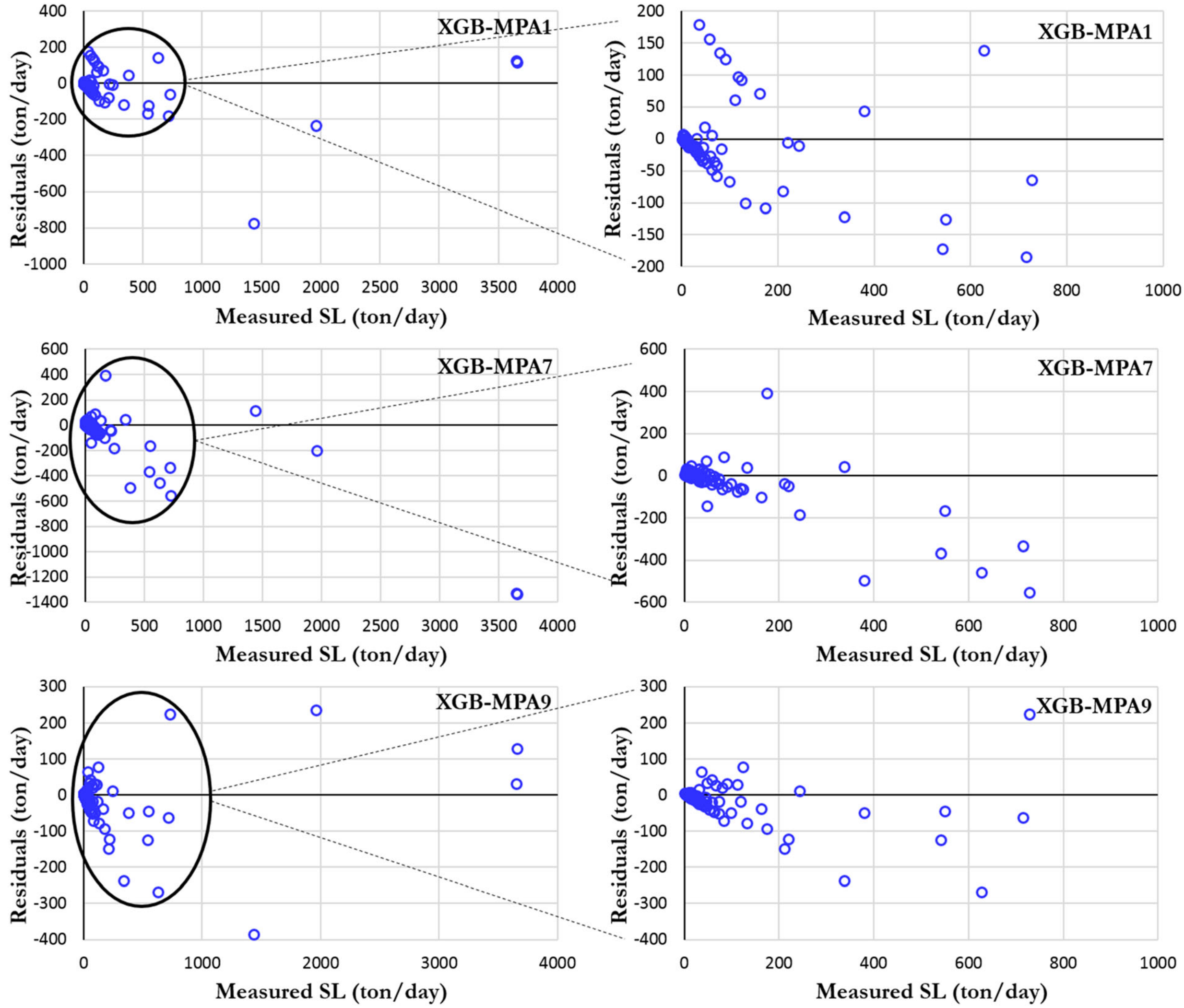


**Figure 8.** Residual plot of XBG-MPA1, XGB-MPA7, and XGB-MPA9 models in testing phase.

population size impacts the model's performance, specifically focusing on the $R^2$ values. The population size ranged from 10 to 100, with other parameters kept constant. The relationship between the $R^2$ values and the population size is depicted in Figure 10. Accepting 0.9 as the ideal value for $R^2$, the greatest variations and the least average $R^2$ have occurred in $< 50$ MPA population size interval; such that in 45 and 18% of cases, $R^2$ is less than 0.9 and 0.7, respectively. For MPA population sizes above 50, the highest average $R^2$ (0.93) was obtained in 50–60 and 80–90 intervals, and both of them can thus be candidates for the best MPA population size interval in the hybrid model. Overall, although the number of $R^2$ local minimums is higher in the 50–60 interval (3 minimums) compared to the 80–90 interval (1 minimum), the lowest $R^2$ local minimums are 0.8 and 0.68 in 50–60 and 80–90 intervals, respectively, and this can be considered an advantage of the former interval. In other words, the narrower range of $R^2$ variations in the 50–60 interval (0.8–0.96) compared with the 80–90 interval (0.68–0.96) – bearing in mind that average $R^2$ values are equal in the two intervals – and the lower computational complexity and burden due to the smaller MPA population size in the former interval can be taken into account as advantages in the choice of the best MPA population size interval. This sensitivity analysis underscores the importance of choosing an appropriate population size for the MPA to achieve the best trade-off between model accuracy and efficiency.

### *3.3. Explainability of proposed models by SHapley Additive exPlanations (SHAP)*

Calculation of SHAP values for the XGB-MPA4 model showed that the streamflow discharge has had a noticeably greater impact on sediment load estimation in comparison with 1, 2 and 3-lag times of this same parameter, in such a way that replacement of $Q_t$ with each of 1, 2

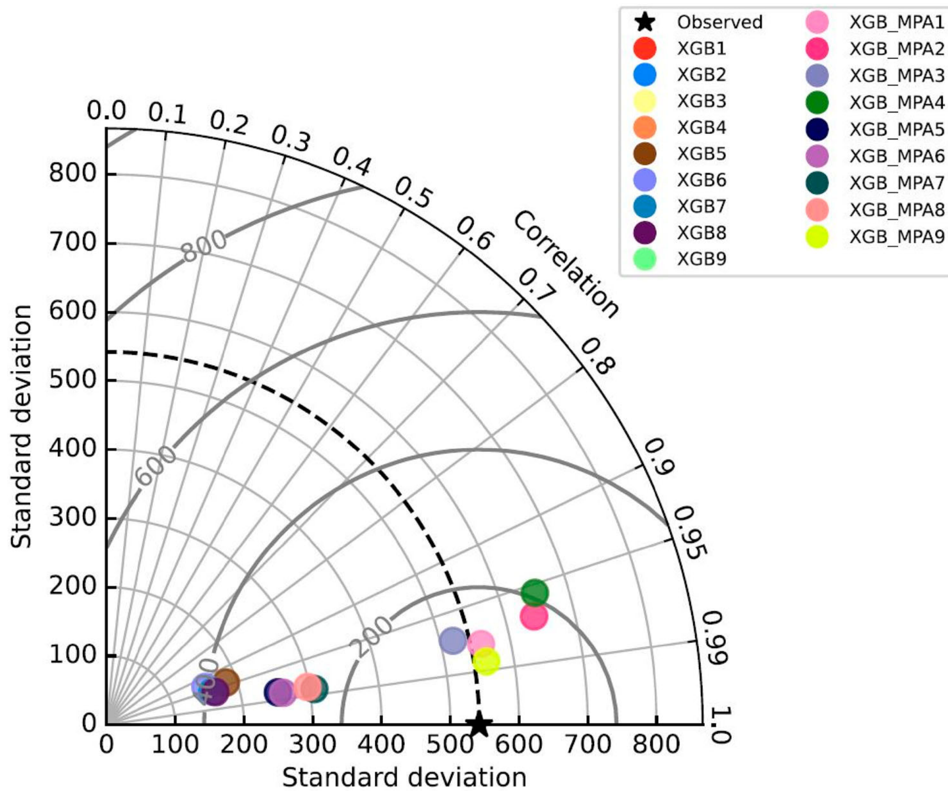


**Figure 9.** Taylor diagram of all scenarios for SL prediction by XGB and XGB-MPA models in testing phase.

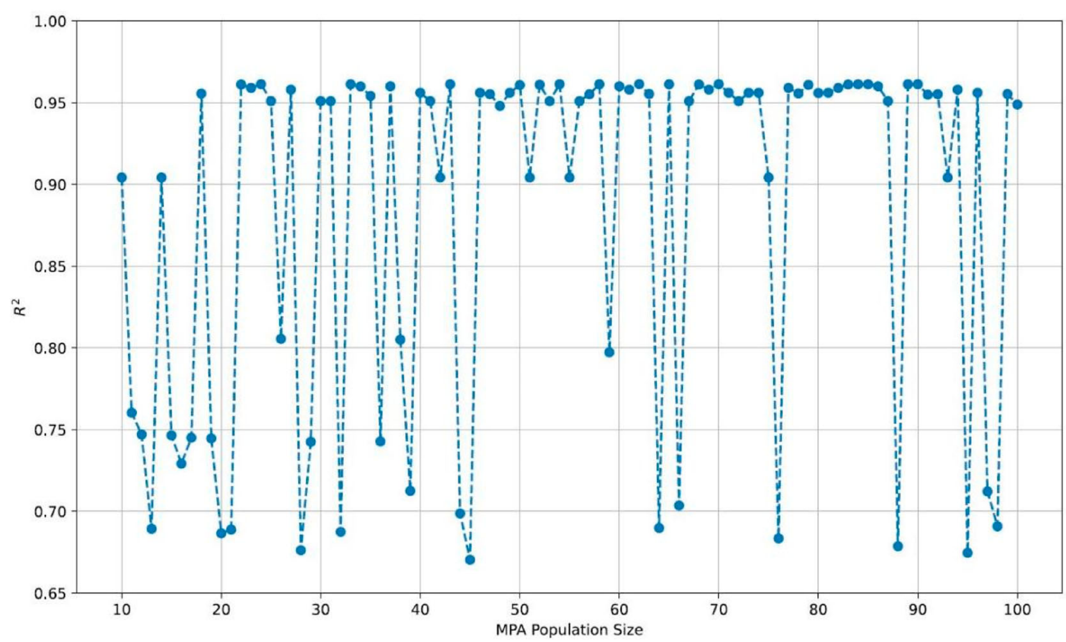


**Figure 10.** The impact of population size on $R^2$ values for the XGB-MPA9 model.

and 3-lag times of streamflow discharge (as input) has led to a 95% decrease in average SHAP (Figure 11). In addition, the results showed that different values of the streamflow discharge across its range of variation have more distinct and separate effects on sediment load estimation compared to $Q_{t-1}$, $Q_{t-2}$ and $Q_{t-3}$ parameters. In other words, the impact on the estimation of sediment load by XGB-MPA4 model has been clearer for alterations of the $Q_t$ parameter relative to those of the other three parameters, indicating that the model is more sensitive to variations of this parameter (especially when its values are relatively high).

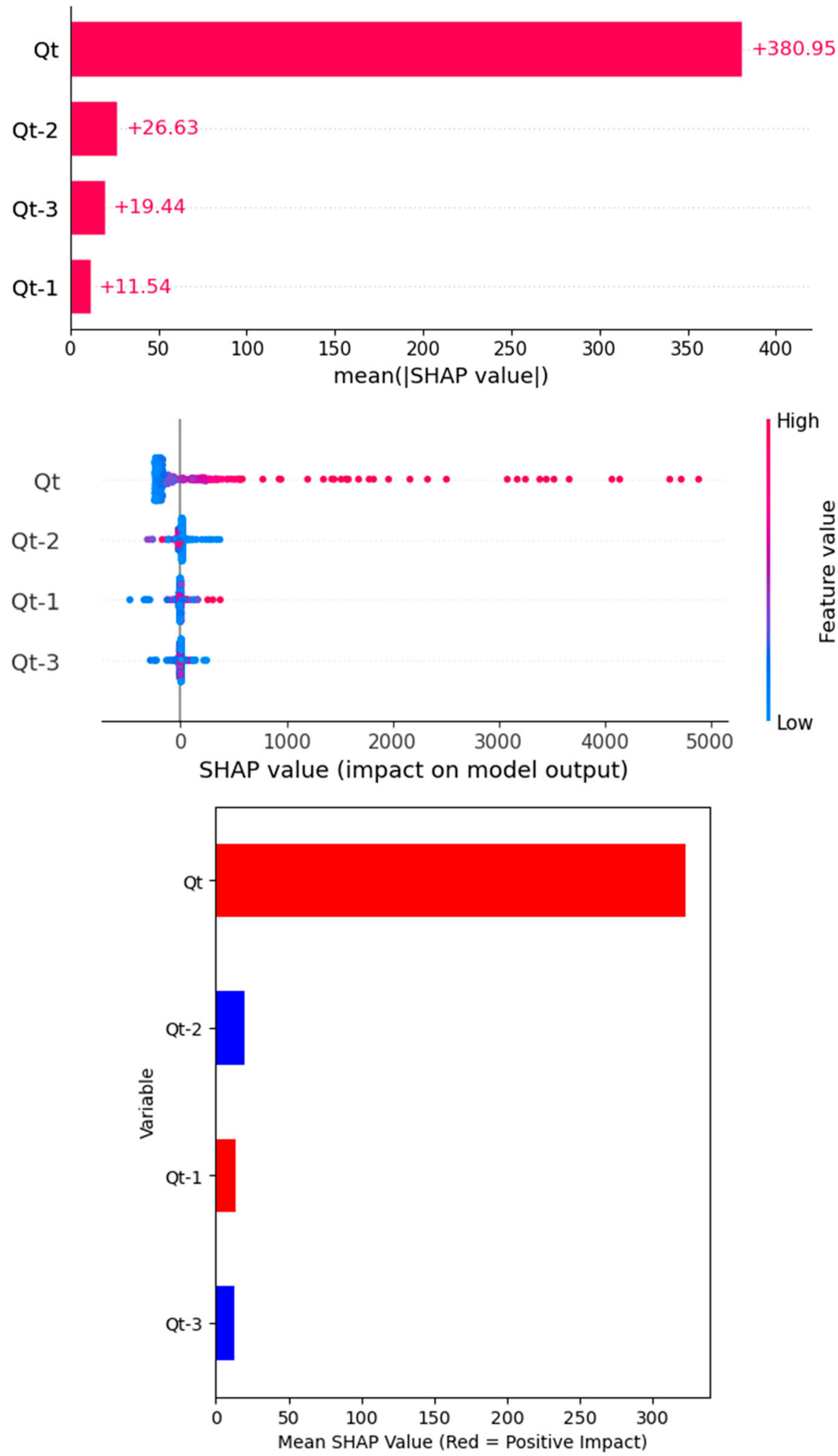


**Figure 11.** Bar plot, beeswarm plot, and SHAP simplified plot for XGB-MPA4 model.

Examining SHAP values regarding the effect of different sediment concentrations on the estimation of sediment load by XGB-MPA8 model revealed that the sediment concentration parameter has exerted the greatest impact on sediment load estimation in comparison with 1, 2 and 3-lag times of this same parameter, which have led to reduction of average SHAP by 81, 89 and 90%, respectively (Figure 12). Similar to the streamflow, variations of the sediment concentration (especially in higher values) have been much more effective on sediment load

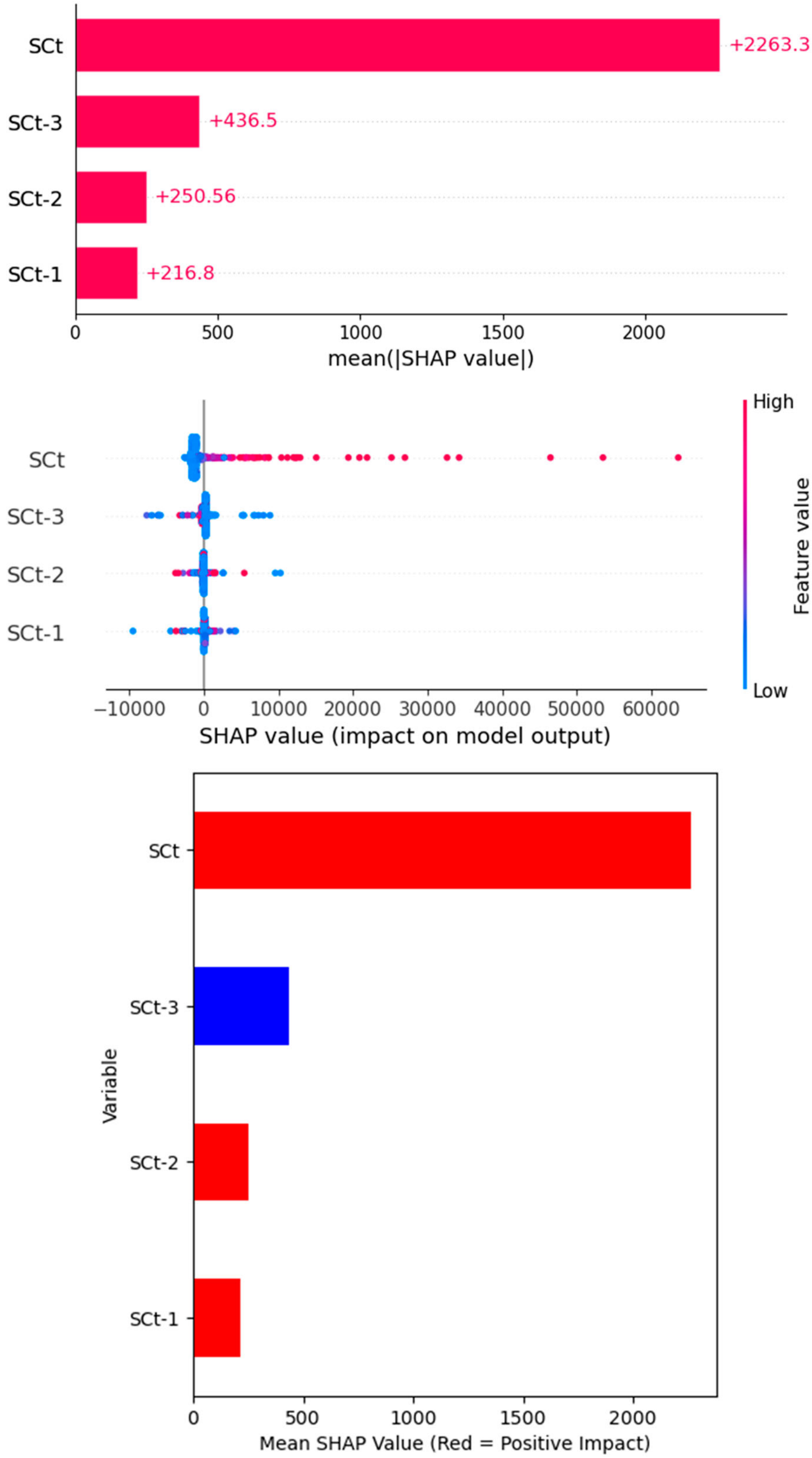


**Figure 12.** Bar plot, beeswarm plot, and SHAP simplified plot for XGB-MPA8 model.

estimation compared with the variations in 1, 2 and 3-lag times of sediment concentration. However, and somewhat contrary to the relatively continuous trend of SHAP variations for different streamflows, variations in higher sediment concentration values have mainly led to SHAP values in the 0–27000 range, which covers almost half of the total range of SHAP values, i.e. 0–60000. In other words, variations in the streamflow (as the most effective discharge parameter in estimation of sediment load, compared with 1 to 3-lag times of this same parameter) have led to a relatively continuous increase in SHAP values, whereas variations in the sediment concentration (as the most effective sediment parameter in estimation of sediment load, compared with 1 to 3-lag times of the same parameter) has led to an intermittent increase in the upper half of SHAP values. This finding shows that variations in sediment concentration have led to greater variations in SHAP values – and, ultimately, in sediment load estimation – compared to those of streamflow discharge.

### 3.4. Limitation of this study and recommendations for future studies

While the proposed XGB- MPA model demonstrates significant advancements in SL estimation, it is important to acknowledge certain limitations inherent in this study. The performance of the XGB- MPA model heavily relies on the quality and quantity of the available data. In cases where the dataset is limited, lacks variability, or contains measurement errors, the model's accuracy and generalizability can be impacted. This study is constrained to the data from the Yeşilirmak River- Gömeleönü station, and the findings may not directly translate to other watersheds with different hydrological and geological characteristics. This study focuses on dataset that measured one day per month for SL prediction, which may not capture short-term sediment dynamics effectively. Events such as flash floods or rapid land-use changes, which can significantly influence sediment load on a shorter timescale, might not be adequately reed in this study modelling framework. The model's applicability to different environmental conditions or regions with varying climatic, hydrological, and topographical features remains a question. Another limitation of the study is that changes in rainfall, vegetation, and other factors cannot be included in the model. The adaptability of the model to these diverse conditions without significant re-tuning is not fully explored in this study. Recognizing these limitations is crucial for guiding future potential research and for the prudent application of the model in practical scenarios.

Building upon the findings and acknowledging the limitations of this study, several recommendations for future research emerge. These suggestions aim to enhance the understanding, applicability, and effectiveness of bio-inspired models like XGB-MPA in SL estimation and similar environmental modelling tasks. Future research can consider incorporating data from a variety of geographic locations, river types, and climatic conditions. This approach will test the model's robustness and adaptability to different environmental settings, potentially increasing its generalizability. Investigating the model's performance with high-frequency data, such as daily or hourly sediment load measurements, could provide useful information into its capability to capture short-term dynamics and extreme events like flash floods. Conducting comparative studies with other machine learning and statistical models would provide a broader perspective on the strengths and weaknesses of the XGB-MPA approach in SL estimation. Future research could experiment with additional or alternative input features, such as land use data, rainfall intensity, and temperature, to understand their impact on sediment load predictions and to capture the influence of external factors more comprehensively. Implementing more advanced interpretability techniques could help in better understanding the model's decision-making process, making it more accessible to a wider range of users, including those with limited expertise in machine learning. Integrating the model into real-time prediction and monitoring systems could be explored. This application could provide timely information for flood risk management, reservoir operation, and environmental protection. Experimenting with hybrid models that combine XGB with other optimization algorithms could uncover new understanding and potentially lead to more effective modelling approaches. Moreover, the prediction performance of the established model can be evaluated by reducing the calculation complexity and removing noise with dimensionality reduction techniques such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Independent Component Analysis (ICA) and Stochastic Neighbor Embedding (SNE). To increase the capability and accuracy, some techniques can be integrated with XGB-MPA model such as variational model decomposition, empirical wavelet transforms, Cosine-Sine decomposition, and singular spectrum analysis. Also, the Box-Cox transformation (BC) and Relative Strength Index (RSI) techniques can be used for improving the data normalization, prediction accuracy, stability, noise reduction, and the correlation between predictors and SL. By considering these recommendations, future studies can contribute to the advancement of SL modelling, offering more accurate, efficient, and versatile tools for environmental monitoring and management.

## 4. Conclusion

This study presents a novel interpretable approach to estimating SL in river systems by coupling a bio-inspired optimization algorithm with XGB (XGB-MPA). Streamflow and sediment concentration data and their 1- to 3-month lag times over the 1973–2023 period were used as model inputs. According to the results, simultaneous application of Q as well as 1 and 3-lag times of SC values as inputs have led to the best SL estimates by the proposed hybrid model, with XGB-MPA9 (MAE = 40.94 ton/day; RMSE = 103.7 ton/day; NSE = 0.96) exhibiting the lowest error rates.

Overall, the results of this work showed that for the base model, the scenarios using SC and its lag times had the lowest SL estimation errors, followed by the hybrid scenario and the scenarios using Q and its lag times. For the hybrid model and when the MPA optimizer algorithm was coupled with the base model, the behaviour patterns of the models changed, such that the hybrid scenario (XGB-MPA9) had the best performance in SL estimation, followed by the scenarios using Q and its lag times and the scenarios using SC and its lag times. Each scenario exhibited a better performance in the hybrid model relative to its corresponding scenario in the base model; and the differences between corresponding scenarios (between the base model and the hybrid model) were much greater when Q and its lag times were used compared to when SC and its lag times were used as model inputs. Key findings from this study include:

- Enhanced Predictive Performance: The integration of MPA with XGB significantly improves the prediction accuracy of SL, making it a valuable tool for water resource management and environmental monitoring.
- Optimization Efficiency: The use of MPA for parameter optimization is validated against other optimization algorithms such as PSO and GWO, demonstrating superior performance and justifying its selection.
- Practical Implications: The robust performance of the XGB-MPA model underscores its potential for practical applications in river sediment management, contributing to more effective and informed decision-making processes.
- Limitations of the present study and ideas for further research: Generalizability of the proposed hybrid model in different hydrological conditions and with various datasets was the main limitation. Further research can include the use of time series of efficient variables extracted from satellite images (e.g. land cover, river morphology, etc.) as model inputs.

These findings highlight the importance of using advanced hybrid models for complex hydrological predictions, providing a foundation for future research and applications in similar environmental and water resource management contexts.

## CRediT authorship contribution statement

1. Roozbeh Moazenzadeh: Methodology, Conceptualization, Formal analysis, Writing – review & editing, Software
2. Okan Mert Katipoglu: Data curation, Methodology, Writing original draft
3. Ahmadreza Shateri: Writing original draft, Formal analysis, Software
4. Hamid Nasiri: Methodology, Writing original draft, Formal analysis, Software
5. Mohammed Abdallah: Writing original draft, Formal analysis

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Availability of data and materials

The datasets used and/or analyzed during this study are available from the corresponding author on reasonable request.

## ORCID

*Roozbeh Moazenzadeh* http://orcid.org/0000-0002-1057-3801

## References

Abdallah, M., Mohammadi, B., Nasiri, H., Katipoğlu, O. M., Abdalla, M. A. A., & Ebadzadeh, M. M. (2023). Daily global solar radiation time series prediction using variational mode decomposition combined with multi-functional recurrent fuzzy neural network and quantile regression forests algorithm. *Energy Reports*, *10*, 4198–4217. https://doi.org/10.1016/j.egyr.2023.10.070

Ahmadi, H., Mastouri, R., & Jafarinia, R. (2022). On the capability of preprocessing techniques for suspended sediment load prediction using artificial intelligence methods. *Hydrological Sciences Journal*, *67*(3), 369–384. https://doi.org/10.1080/02626667.2021.2016768

AlDahoul, N., Ahmed, A. N., Allawi, M. F., Sherif, M., Sefelnasr, A., Chau, K. W., & El-Shafie, A. (2022). A comparison of machine learning models for suspended sediment load classification. *Engineering Applications of Computational Fluid Mechanics*, *16*(1), 1211–1232. https://doi.org/10.1080/19942060.2022.2073565

AlDahoul, N., Essam, Y., Kumar, P., Ahmed, A. N., Sherif, M., Sefelnasr, A., & Elshafie, A. (2021). Suspended sediment load prediction using long short-term memory neural network.

*Scientific Reports*, *11*(1), 7826. https://doi.org/10.1038/s41598-021-87415-4

Allawi, M. F., Sulaiman, S. O., Sayl, K. N., Sherif, M., & El-Shafie, A. (2023). Suspended sediment load prediction modelling based on artificial intelligence methods: The tropical region as a case study. *Heliyon*, *9*(8), e18506. https://doi.org/10.1016/j.heliyon.2023.e18506

Atalay, I. (1982). *Introduction to geomorphology of Turkey*. Ege University Ege University Faculty of Social Sciences Publications.

Ayyadevara, V. (2018). Gradient boosting machine. In [place unknown]; 117–134. https://doi.org/10.1007/978-1-4842-3564-5_6

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

Bhagat, S. K., Tiyasha, T., Kumar, A., Malik, T., Jawad, A. H., Khedher, K. M., Deo, R. C., & Yaseen, Z. M. (2022). Integrative artificial intelligence models for Australian coastal sediment lead prediction: An investigation of in-situ measurements and meteorological parameters effects. *Journal of Environmental Management*, *309*, 114711. https://doi.org/10.1016/j.jenvman.2022.114711

Burnwal, Y., & Jaiswal, R. C. (2023). A comprehensive survey on prediction models and the impact of XGBoost. *International Journal for Research in Applied Science and Engineering Technology*, *11*(12), 1552–1556. https://doi.org/10.22214/ijraset.2023.57625

Buyukyildiz, M., & Kumcu, S. Y. (2017). An estimation of the suspended sediment load using adaptive network based fuzzy inference system, support vector machine and artificial neural network models. *Water Resources Management*, *31*(4), 1343–1359. https://doi.org/10.1007/s11269-017-1581-1

Calsamiglia, A., Fortesa, J., García-Comendador, J., Lucas-Borja, M. E., Calvo-Cases, A., & Estrany, J. (2018). Spatial patterns of sediment connectivity in terraced lands: Anthropogenic controls of catchment sensitivity. *Land Degradation & Development*, *29*(4), 1198–1210. https://doi.org/10.1002/ldr.2840

Cao, Z., & Carling, P. A. (2003). On evolution of bed material waves in alluvial rivers. *Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group*, *28*(4), 437–441. https://doi.org/10.1002/esp.493

Chen, T., & Guestrin, C. (2016, August 13-17). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785–794). https://doi.org/10.1145/2939672.2939785.

DMGM. (2023). Recieved:23.12.2023. https://www.mgm.gov.tr/veridegerlendirme/il-ve-ilceler-istatistik.aspx?m = TOKAT.

Ebtehaj, I., Bonakdari, H., Safari, M. J. S., Gharabaghi, B., Zaji, A. H., Madavar, H. R., Khozani, Z. S., Es-haghi, M. S., Shishegaran, A., & Mehr, A. D. (2020). Combination of sensitivity and uncertainty analyses for sediment transport modeling in sewer pipes. *International Journal of Sediment Research*, *35*(2), 157–170. https://doi.org/10.1016/j.ijsrc.2019.08.005

Faramarzi, A., Heidarinejad, M., Mirjalili, S., & Gandomi, A. H. (2020). Marine Predators Algorithm: A nature-inspired metaheuristic. *Expert Systems with Applications*, *152*, 113377. https://doi.org/10.1016/j.eswa.2020.113377

Farzad, R., Ahmadi, F., Sharafati, A., & Hosseini, S. A. (2023). Investigating hybrid deep learning models and meta-heuristic algorithms in predicting evaporation from a reservoir: a case study of Dez dam. *Earth Science Informatics*, *16*(4), 3597–3618. https://doi.org/10.1007/s12145-023-01110-8

Ghaheri, P., Shateri, A., & Nasiri, H. (2023). PD-ADSV: An automated diagnosing system using voice signals and hard voting ensemble method for Parkinson's disease. *Software Impacts*, *16*, 100504. https://doi.org/10.1016/j.simpa.2023.100504

González, S., García, S., Del Ser, J., Rokach, L., & Herrera, F. (2020). A practical tutorial on bagging and boosting based ensembles for machine learning: Algorithms, software tools, performance study, practical perspectives and opportunities. *Information Fusion*, *64*, 205–237. https://doi.org/10.1016/j.inffus.2020.07.007

Hamel, P., Falinski, K., Sharp, R., Auerbach, D. A., Sánchez-Canales, M., & Dennedy-Frank, P. J. (2017). Sediment delivery modeling in practice: Comparing the effects of watershed characteristics and data resolution across hydroclimatic regions. *Science of the Total Environment*, *580*, 1381–1388. https://doi.org/10.1016/j.scitotenv.2016.12.103

Hanoon, M. S., Abdullatif B, A. A., Ahmed, A. N., Razzaq, A., Birima, A. H., & El-Shafie, A. (2022). A comparison of various machine learning approaches performance for prediction suspended sediment load of river systems: a case study in Malaysia. *Earth Science Informatics*, *15*(8), 1–14. https://doi.org/10.1007/s12145-021-00689-0

Ikram, R. M. A., Ewees, A. A., Parmar, K. S., Yaseen, Z. M., Shahid, S., & Kisi, O. (2022). The viability of extended marine predators algorithm-based artificial neural networks for streamflow prediction. *Applied Soft Computing*, *131*, 109739. https://doi.org/10.1016/j.asoc.2022.109739

Jia, Y., Jin, S., Savi, P., Gao, Y., Tang, J., Chen, Y., & Li, W. (2019). GNSS-R soil moisture retrieval based on a XGboost machine learning aided method: Performance and validation. *Remote sensing*, *11*(14), 1655. https://doi.org/10.3390/rs11141655

Kale, M. M. (2018). *Determination of changing spatial distribution of precipitation with deterministic and stochastic methods in Yeşilırmak Basin*. Bulletin of the Earth Sciences Application and Research Centre of Hacettepe University. https://doi.org/10.17824/yerbilimleri.503952.

Kavzoglu, T., & Teke, A. (2022). Advanced hyperparameter optimization for improved spatial prediction of shallow landslides using extreme gradient boosting (XGBoost). *Bulletin of Engineering Geology and the Environment*, *81*(5), 201. https://doi.org/10.1007/s10064-022-02708-w

Khosravi, K., Farooque, A. A., Bateni, S. M., Jun, C., Mohammadi, D., Kalantari, Z., & Cooper, J. R. (2024). Fluvial bedload transport modelling: advanced ensemble tree-based models or optimized deep learning algorithms? *Engineering Applications of Computational Fluid Mechanics*, *18*(1), 2346221. https://doi.org/10.1080/19942060.2024.2346221

Kisi, O., & Yaseen, Z. M. (2019). The potential of hybrid evolutionary fuzzy intelligence model for suspended sediment concentration prediction. *Catena*, *174*, 11–23. https://doi.org/10.1016/j.catena.2018.10.047

Li, S., & Yang, J. (2022). Modelling of suspended sediment load by Bayesian optimized machine learning methods with seasonal adjustment. *Engineering Applications of Computational Fluid Mechanics*, *16*(1), 1883–1901. https://doi.org/10.1080/19942060.2022.2121944

Mohammadi, B., Guan, Y., Moazenzadeh, R., & Safari, M. J. S. (2021). Implementation of hybrid particle swarm optimization-differential evolution algorithms coupled with multi-layer perceptron for suspended sediment load estimation. *Catena*, *198*, 105024. https://doi.org/10.1016/j.catena.2020.105024

Mohammadian, A., Tajrishi, M., & Lotfiazad, F. (2004). Two dimensional numerical simulation of flow and geomorphological processes near headlands by using unstructured grid. *National Research: English Edition*, *19*(4), 258–277.

Ngamsert, R., & Kangrang, A. (2022). Applying of marine predators algorithm linked with reservoir simulation model considering sedimentation for reservoir operation. *Advances in Civil Engineering*, *2022*, 1631914. https://doi.org/10.1155/2022/1631914

Nguyen, H., Vu, T., Vo, T. P., & Thai, H. (2021). Efficient machine learning models for prediction of concrete strengths. *Construction and Building Materials*, *266*, 120950. https://doi.org/10.1016/j.conbuildmat.2020.120950

Niazkar, M., Menapace, A., Brentan, B., Piraei, R., Jimenez, D., Dhawan, P., & Righetti, M. (2024). Applications of XGBoost in water resources engineering: A systematic literature review (Dec 2018–May 2023). *Environmental Modelling & Software*, *174*, 105971. https://doi.org/10.1016/j.envsoft.2024.105971

Ozturk, F., Yurekli, K., & Apaydin, H. (2008). Stochastic modeling of suspended sediment from Yesilirmak Basin, Turkey. *International Journal of Natural and Engineering Sciences*, *2*(1), 21–27.

Palazón, L., Latorre, B., Gaspar, L., Blake, W. H., Smith, H. G., & Navas, A. (2016). Combining catchment modelling and sediment fingerprinting to assess sediment dynamics in a Spanish Pyrenean River system. *Science of the Total Environment*, *569-570*, 1136–1148. https://doi.org/10.1016/j.scitotenv.2016.06.189

Patnaik, B., Mishra, M., Bansal, R. C., & Jena, R. K. (2021). MODWT-XGBoost based smart energy solution for fault detection and classification in a smart microgrid. *Applied Energy*, *285*, 116457. https://doi.org/10.1016/j.apenergy.2021.116457

Piraei, R., Afzali, S. H., & Niazkar, M. (2023). Assessment of XGBoost to estimate total sediment loads in rivers. *Water Resources Management*, *37*(13), 5289–5306. https://doi.org/10.1007/s11269-023-03606-w

Rodriguez-Lloveras, X., Buytaert, W., & Benito, G. (2016). Land use can offset climate change induced increases in erosion in Mediterranean watersheds. *Catena*, *143*, 244–255. https://doi.org/10.1016/j.catena.2016.04.012

Samantaray, S., Sahoo, A., & Satapathy, D. P. (2022). Improving accuracy of SVM for monthly sediment load prediction using Harris hawks optimization. *Materials Today: Proceedings*, *61*, 604–617. https://doi.org/10.1016/j.matpr.2022.04.594

Sekhar, J. C., Rajyalakshmi, C., Nagaraj, S., Sankar, S., Saturi, R., & Harshavardhan, A. (2023). Deep generative adversarial networks with marine predators algorithm for classification of Alzheimer's disease using electroencephalogram. *Journal of King Saud University – Computer and Information Sciences*, *35*(10), 101848. https://doi.org/10.1016/j.jksuci.2023.101848

Shamaei, E., & Kaedi, M. (2016). Suspended sediment concentration estimation by stacking the genetic programming and neuro-fuzzy predictions. *Applied Soft Computing*, *45*, 187–196. https://doi.org/10.1016/j.asoc.2016.03.009

Sharafati, A., Haji Seyed Asadollah, S. B., Motta, D., & Yaseen, Z. M. (2020). Application of newly developed ensemble machine learning models for daily suspended sediment load prediction and related uncertainty analysis. *Hydrological Sciences Journal*, *65*(12), 2022–2042. https://doi.org/10.1080/02626667.2020.1786571

Sharghi, E., Nourani, V., Najafi, H., & Gokcekus, H. (2019). Conjunction of a newly proposed emotional ANN (EANN) and wavelet transform for suspended sediment load modeling. *Water Supply*, *19*(6), 1726–1734. https://doi:10.2166/ws.2019.044

Shojaeezadeh, S. A., Al-Wardy, M., & Nikoo, M. R. (2024). Suspended sediment load modeling using Hydro-Climate variables and Machine learning. *Journal of Hydrology*, *633*, 130948. https://doi.org/10.1016/j.jhydrol.2024.130948

Si, W., Bao, W., Jiang, P., Zhao, L., & Qu, S. (2017). A semi-physical sediment yield model for estimation of suspended sediment in loess region. *International Journal of Sediment Research*, *32*(1), 12–19. https://doi.org/10.1016/j.ijsrc.2015.10.002

Talbi, E. (2009). *Metaheuristics*. [place unknown]: Wiley. https://doi.org/10.1002/9780470496916

Tao, H., Keshtegar, B., & Yaseen, Z. M. (2019). The feasibility of integrative radial basis M5Tree predictive model for river suspended sediment load simulation. *Water Resources Management*, *33*(13), 4471–4490. https://doi.org/10.1007/s11269-019-02378-6

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, *106*(D7), 7183–7192. https://doi:10.1029/2000JD900719

Tu, Z., Gao, L., Wu, X., Liu, Y., & Zhao, Z. (2023). Rotate Vector Reducer Fault Diagnosis Model Based on EEMD-MPA-KELM. *Applied Sciences*, *13*(7), 4476. https://doi.org/10.3390/app13074476

Uca, T. E., Jaafar, O., Maru, R., Arfan, A., & Ahmar, A. S. (2018). Daily suspended sediment discharge prediction using multiple linear regression and artificial neural network. *Journal of Physics: Conference Series*, *954*, 012030. IOP Publishing. https://doi.org/10.1088/1742-6596/954/1/012030.

Van Thieu, N., & Mirjalili, S. (2023). MEALPY: An open-source library for latest meta-heuristic algorithms in Python. *Journal of Systems Architecture*, *139*, 102871. https://doi.org/10.1016/j.sysarc.2023.102871

Xia, X., Dong, J., Wang, M., Xie, H., Xia, N., Li, H., Zhang, X., Mou, X., Wen, J., & Bao, Y. (2016). Effect of water-sediment regulation of the Xiaolangdi reservoir on the concentrations, characteristics, and fluxes of suspended sediment and organic carbon in the Yellow River. *Science of the Total Environment*, *571*, 487–497. https://doi.org/10.1016/j.scitotenv.2016.07.015

Yagin, F. H., Shateri, A., Nasiri, H., Yagin, B., Colak, C., & Alghannam, A. F. (2024). Development of an expert system for the classification of myalgic encephalomyelitis/chronic fatigue syndrome. *PeerJ Computer Science*, *10*, e1857. https://doi.org/10.7717/peerj-cs.1857

Yaseen, Z. M., El-Shafie, A., Jaafar, O., Afan, H. A., & Sayl, K. N. (2015). Artificial intelligence based models for stream-flow forecasting: 2000–2015. *Journal of Hydrology*, *530*, 829–844. https://doi.org/10.1016/j.jhydrol.2015.10.038

Zhang, Y., & Haghani, A. (2015). A gradient boosting method to improve travel time prediction. *Transportation Research Part C: Emerging Technologies*, *58*, 308–324. https://doi.org/10.1016/j.trc.2015.02.019

Zhong, D., Wang, G., & Sun, Q. (2011). Transport equation for suspended sediment based on two-fluid model of solid/liquid two-phase flows. *Journal of Hydraulic Engineering*, *137*(5), 530–542. https://doi.org/10.1061/(ASCE)HY.1943-7900.0000331

Zounemat-Kermani, M., Mahdavi-Meymand, A., Alizamir, M., Adarsh, S., & Yaseen, Z. M. (2020). On the complexities of sediment load modeling using integrative machine learning: Application of the great river of Loíza in Puerto Rico. *Journal of Hydrology*, *585*, 124759. https://doi.org/10.1016/j.jhydrol.2020.124759

Zubaidi, S. L., Al-Bdairi, N. S. S., Ortega-Martorell, S., Ridha, H. M., Al-Ansari, N., Al-Bugharbee, H., & Gharghan, S. K. (2023). Assessing the benefits of nature-inspired algorithms for the parameterization of ANN in the prediction of water demand. *Journal of Water Resources Planning and Management*, *149*(1), 1–10. https://doi.org/10.1061/(ASCE)WR.1943-5452.000160