Clear-water scour depth prediction in long channel contractions: Application of new hybrid machine learning algorithms

Khabat Khosravi*[1], Mir Jafar Sadegh Safari[2], James R.Cooper[3]

1. Department of Watershed Management Engineering, Ferdowsi University of Mashhad, Mashhad, Iran.
2. Department of Civil Engineering, Yaşar University, Izmir, Turkey. Email: jafar.safari@yasar.edu.tr
3. School of Environmental Sciences, University of Liverpool, Liverpool, UK. Email: james.cooper@liverpool.ac.uk

*Corresponding author: K. Khosravi (Khabat.khosravi@gmail.com and kh.khosravi@um.ac.ir)

Abstract

Scour depth prediction and its prevention is one of the most important issues in channel and waterway design. However the potential for machine learning algorithms to provide models of scour depth has yet to be explored. This study provides the first quantification of the predictive power of a range of standalone and hybrid machine learning models. Using previously collected scour depth data from laboratory flume experiments, the performance of five types of recently developed standalone machine learning techniques - the Isotonic Regression (ISOR), Sequential Minimal Optimization (SMO), Iterative Classifier Optimizer (ICO), Locally Weighted learning (LWL) and Least Median of Squares Regression (LMS) - are assessed, along with their hybrid versions with Dagging (DA) and Random Subspace (RS) algorithms. The main findings are five-fold. First, the DA-ICO model had the highest prediction power. Second, the hybrid models had a higher prediction power than standalone models. Third, all algorithms underestimated the maximum scour depth, except DA-ICO which predicted scour depth almost perfectly. Fourth, scour depth was most sensitive to densimetric particle Froude number followed by the non-dimensionalized contraction width, flow depth within the contraction, sediment geometric standard deviation, approach flow velocity and median grain size. Fifth, most of the algorithms performed best when all the input parameters were involved in the building of the model. An important exception was the best performing model that required only four input parameters: densimetric particle Froude

number, non-dimensionalized contraction width, flow depth within the contraction and sediment geometric standard deviation. Overall the results revealed that hybrid machine learning algorithms provide more accurate predictions of scour depth than empirical equations and traditional AI-algorithms. In particular, the DA-ICO model not only created the most accurate predictions but also used the fewest easily and readily measured input parameters. Thus this type of model could be of real benefit to practicing engineers required to estimate maximum scour depth when designing in-channel structures.

Keywords: Scour depth prediction; data mining; iterative classifier optimizer algorithms; model calibration

1. Introduction

Channel contractions in rivers exist when there is a reduction in the width of the cross-section. The length of the contraction is defined based on the ratio between the length of the contacted area (L) and the width of the channel upstream of the contracted area (b_1, approach width), although the criterion for defining whether a contraction is 'long' varies between researchers. For example Komura (1966) defined contractions as being long if the ratio was above unity, Webby (1984) if it was greater than two, and Raikar (2004) stated the contraction could only be considered long if the flow velocity and turbulence remained constant in the length of the contraction when $L/b_1 \geq 1$. Natural contractions in alluvial rivers, such as debris accumulations, longitudinal bars and confluences, tend to be long contractions, since the flow is nearly uniform in both the undisturbed channel and the contracted reach. Human made examples include gradual contractions created on both sides of the channel to accommodate dams, bridges, weirs and barrages, and partial constrictions on one side of the channel such as spur dikes or cofferdams (Lim, 1993).

Local scour at the contraction occurs because of a local increase in flow velocity and bed shear stress. The severity of this scour varies according to upstream sediment supply conditions, and thus the scour is categorized into two main groups, clear-water and live-bed scour (Dey and Raikar, 2005). Clear-water scour takes place when sediment transport occurs from the scour hole throughout the contraction length and when there is no upstream sediment supply. Live-bed scour takes place when there is upstream sediment supply (Dey, 1997).

To successfully design in-channel structures, the maximum scour depth, which is mainly caused by constriction (Lim, 1993), must be predicted because the depth dictates morphological change around the structure, especially in long contractions, and thus the structure's stability. This prediction is commonly made under the assumption of an equilibrium condition (Lim and Cheng, 1998). The first analytical study of scour depth prediction in this condition in a long contraction was carried out by Straub (1934), and subsequent studies have either proposed different empirical equations or modified Straub's equation (Laursen, 1960; Ashida, 1963; Komura, 1966; Gill, 1981; Webby, 1984; Lim, 1993; Lim and Cheng, 1998). For example, Lim (1993) revealed that previously incorporated effective variables in clear-water scour depth equations were inadequate and therefore proposed a new empirical formula for long contraction based on the approach flow velocity and depth, median grain size of the bed material and the geometry of the constriction. In a following study, Lim and Cheng (1998) revealed Straub's equation underestimates scour depth in both clear-water and live-bed conditions, and proposed a new empirical formula based on only the approach water depth and the ratio between the approach and constricted channel width. Dey and Raikar (2005) used an experimental approach to investigate the controlling variables on scour depth for uniform and graded sediments. They found that Lim's (1993) equation over predicted scour depth and developed their own empirical equation based on the limiting stability of bed sediments in the approaching channel under clear-water scour. These examples illustrate that a range of equations exist, all based on differing controlling variables, and each performing differently according to the channel and flow conditions being studied. No universal equation exists to predict scour depth well in all conditions. One possible reason is the proposed scour depth relationships were developed using conventional regression analysis.

Recently artificial intelligence (AI) algorithms have gained much interest because of their non-linear structure (Maier et al. 2014), superior prediction power - particularly for complex phenomena - shorter computational times, low sensitivity to missing values, and their ability to handle large datasets with different temporal and spatial scales (Melesse et al, 2011; Yaseen et al. 2016). These AI based algorithms have been widely used in the fields of water science and hydraulic engineering. Artificial neural network (ANN) algorithms are one of the most widely used (Abrahart et al. 2012). ANN was found initially to successfully predict scour depth around hydraulic structures (Muzzammil, 2008; Mousa, 2013; Onen, 2014), but recent studies have shown low prediction capability if the training datasets are not carefully selected (Melesse et al, 2011; Kisi et al, 2012; Choubin et al, 2018), especially when the range of testing data is out of range of the training dataset (low generalization power).

To predict scour depth around hydraulic structures, Adaptive neuro-fuzzy inference system (ANFIS) algorithms have been developed by integrating ANN with fuzzy logic(Firat, 2009; Rady, 2020). One of the key challenges in developing an accurate ANFIS models is determining accurately the weights in the membership function of the ANFIS algorithm (Chen et al. 2017; Bui et al. 2016)..

Alternative types of AI-based algorithms have been successfully applied and proposed for scour depth prediction. For example, Guven and Gunal (2008) and Azamathulla et al. (2010)

reported that genetic programming (GP) outperforms conventional regression and ANN approaches. Ayoubloo et al. (2010) showed the classification and regression trees (CART) algorthim is more accuratefor scour modeling than the ANN method. Furthermore, Etemad-Shahidi and Ghaemi (2011) examined the potential of two other AI-based algorithms, Support vector machine (SVM) and M5 model tree. Their work found the M5 model tree algorithm outperformed ANN and SVM approaches. Ghaleh Nou et al. (2019) also found the same results for self-adaptive extreme learning machine in comparison to SVM and ANN approaches One possible reason is SVM algorithms have many hyperparameters that require tuning and thus finding the best combination is a challenge. Whereas Parsaie et al. (2019) showed that SVM provided more accurate predictions than ANN and ANFIS approaches. However, Najafzadet et al. (2016) reported the superiority of the ANFIS to SVM, reaavling further that the performance of AI-based algorithms are sensitive to the dataset, and thus fluvial conditions, used to build them Najafzadeh et al. (2013, 2014) explored the influence of hybridization by applying Neuro-fuzzy group method of data handling (GMDH) systems based evolutionary algorithms to predict scour pile groups in clear water conditions. Their work revealed the integration of these models with evolutionary algorithms enhanced model performance. In a subsequent study, Najafzadeh et al. (2018) applied three algorithms of gene expression programming (GEP), model tree (MT) and evolutionary polynomial regression (EPR) to reveal that the MT algorithm had a higher prediction power than GEP and EPR in long channel contractions. This work was developed further by revealing the performance of a hybrid model - GMDH hybridized with GEP - was greater than the standalone ANN, GEP and GMDH algorithms (Najafzadeh and Saberi-Movahed, 2019). .In these studies, almost all of the AI-based algorithms provided more accurate predictions of scour depth than empirical equations.

Recently a new branch of AI algorithms, called machine learning, have been developed, providing strong performance in other environmental and engineering fields. For example, in the field of landslide prediction, stochastic gradient descent (SGD), AdaBoost (AB), logistic model tree (LMT), functional tree (FT), Naïve Bayes Tree (NBT), Bayes network (BN), and Naïve Bayes (NB) algorithms have been successfully applied (Bui et al. 2019a; Pham et al. 2019). LMT, REPT and Alternating Decision Trees (ADT) algorithms have provided strong prediction of flood maps (Khosravi et al. 2018a; Chapi et al. 2017), and ADT and AB algorithms have proved successful in the prediction of groundwater potential (Bui et al. 2019b). Further, Khosravi et al. (2018b) and Salih et al. (2019) have applied BA-M5P, attribute selected classifier (AS), M5Rule (M5R), KStar, instance-based learning (IBK), random committee-REPT (RC-REPT) and random subspace-REPT (RS-REPT) to predict suspended sediment loads. Decision trees algorithms have been used to predict reference evaporation (Khosravi et al., 2019), apparent shear stress in compound channels (Khozani et al., 2019) and solar radiation (Sharafati et al., 2019). Khosravi et al. (2020a) predicted fluoride concentration in groundwater through IBK and locally weighted learning (LWL) algorithms. Khosravi et al. (2020b) developed hybrid algorithms of bagging (BA) with decision trees algorithm (BA-M5P), random forest (BA-RF), random tree (BA-RT) and reduced error pruning tree (BA-REPT) for bed load transport rate prediction. Nitrate and strontium concentration has been predicted using Gaussian Process (GP) algorithms (Bui et al, 2020a), and water quality indices have been simulated using a hybrid of BA, CV parameter selection (CVPS) and randomizable filtered classification (RFC) with decision trees algorithms (Bui et al. 2020b). All of these previous studies have shown that hybrid algorithms have a higher prediction power then their standalone counterparts, but they have yet to be applied to the prediction of scour depth.

Thus the aim of this study was to evaluate the ability of hybrid machine learning algorithms to provide accurate predictions of maximum scour depth. The focus was on scour within long contractions within clear-water conditions. Five standalone machine learning algorithms - Isotonic Regression (ISOR), Sequential Minimal Optimization (SMO), Iterative Classifier Optimizer (ICO), Locally Weighted learning (LWL) and Least Median of Squares Regression (LMS) - were hybridized with Dagging (DA) and Random Subspace (RS) algorithms to develop 10 novel hybrid algorithms DA-ISOR, DA-SMOR, DA-LWL, DA-ICO, DA-LMS, -ISOR, RS-SMOR, RS-LWL, RS-ICO, RS-LMS). This study is the first to apply a diverse range of newly developed machine learning models to the prediction of scour depth. The research offers new insight into which machine learning algorithms offer the potential to provide accurate and efficient predictions of scour depth based on readily and easily measured flow and channel variables.

2. Methodology

2.1. Identifying effective parameters

According to the literature, the parameters which have a significant effect on scour depth in a long contraction can be classified into four different types (Straub, 1934; Laursen, 1960; Ashida, 1963; Komura, 1966; Gill, 1981; Webby, 1984; Lim, 1993; Lim and Cheng, 1998; Raikar, 2004): (1) approaching flow conditions (flow velocity U_1, critical flow velocity U_c, flow depth h_1, water density ρ_w, densimetric particle Froude number Fr_0); (2) characteristics of the bed material characteristics (median grain size d_{50}, sediment density ρ_s, sediment geometric standard deviation σ_g); (3) geometry of the un-contracted section (width b_1, h_1); and (4) geometry of the contracted (width b_2, flow depth h_2) section. The functional relationship of scour depth (d_s) with these effective input parameters can be described as follows:

$$d_s = f(U_1, U_c, \rho_w, d_{50}, \rho_s, \sigma_g, b_1, h_1, b_2, g) \quad (1)$$

To enhance the modeling performance of the AI-based algorithms (Azamathulla et al. 2009; Pal et al. 2014) and allow direct comparisons between datasets and the results of previous studies, all parameters were non-dimensionalized. Such an approach is recommended in the application of AI algorithms because it improves model performance (Azamathulla et al. 2009; Pal et al. 2014). Since there are a large number of variables, a non-dimensional approach was also required in the use of the Buckingham π theorem to determine groupings between parameters. Also more generally, finding a functional relationship among non-dimensional parameters can allow a practical model to be developed to mitigate the adverse consequences of the experimental data scale effects.

Applying the Buckingham π theorem and using $\Delta\rho = \rho_s - \rho_w$ instead of ρ_s, and taking U_1, b_1 and ρ as repeating variables, allows the following dimensionless parameters to be obtained.

$$f\left(d_s/b_1\,, U_1/U_c\,, d_{50}/b_1\,, b_2/b_1\,, h_1/b_1\,, \Delta\rho/\rho\,, U_1^2/gb_1\,, \sigma_g\right) = 0 \quad (2)$$

Combining the three π parameters of U_1^2/gb_1, d_{50}/b_1 and $\Delta\rho/\rho$ as $[(U_1^2/gb_1) \times (d_{50}/b_1)^{-1}/(\Delta\rho/\rho)]^{0.5}$ gives

$$Fr_0 = \frac{U_1}{\sqrt{g((\rho_s/\rho_w)-1)d_{50}}} \quad (3)$$

where Fr_0 is the densimetric particle Froude number. Using the Buckingham theorem, six dimensionless parameters were extracted as having the most effect on scour depth. Thus the scour depth was normalized using b_1 and the effective variables were extracted as follows:

$$d_s/b_1 = f(U_1/U_c, Fr_0, d_{50}/b_1, b_2/b_1, h_1/b_1, \sigma_g) \quad (4)$$

Eq. (4) shows all effective variables given in Eq. (1) are appropriately incorporated as dimensionless model parameters. In order to compute scour depth as a dependent variable,

d_s/b_1 is considered as a model output and dimensionless parameters given in the right hand side of Eq. (4) are used as model inputs. The form of Eq. (4) is in agreement with that used for scour depth determination by Dey and Raikar (2005) and Najafzadeh et al. (2018).

2.2. Dataset collection and preparation

In the current study 204 datasets from four laboratory flume studies, collected and compiled by Najafzadeh et al. (2016, 2018) for the testing of standalone learning algorithms, were used (Kamura, 1966; Gill, 1981; Webby, 1984; Lim, 1993; and Dey and Raikar, 2005). All of the data were measured in a long contraction rectangular channel in a clear-water condition. The datasets were divided into three sections randomly. Among the 204 datasets, 70% (140 row-data) was used as a training dataset for model building, 10% for calibration (22 row-data) and the remaining 20% (42 row-data) for model validation. A statistical summary of the datasets is presented in Table 1.

Table 1. Descriptive statistics of utilized data

Parameters	Training					Calibration					Testing				
h_1/b_1	Max	Min	Mean	STD	Skew	Max	Min	Mean	STD	Skew	Max	Min	Mean	STD	Skew
b_2/b_1	0.23	0.04	0.14	0.05	-0.21	0.22	0.07	0.14	0.06	-0.21	0.23	0.06	0.15	0.06	-0.22
Fr_0	0.70	0.25	0.53	0.13	-0.36	0.70	0.25	0.53	0.13	-0.51	0.70	0.25	0.54	0.13	-0.51
U_1/U_c	5.05	0.11	0.60	0.57	4.77	2.34	0.12	0.57	0.34	1.17	1.52	0.12	0.55	0.34	1.07
σ_g	1.00	0.39	0.89	0.12	-2.40	1.00	0.55	0.90	0.10	-2.44	1.00	0.55	0.90	0.10	-2.44
d_s/b_1	23.75	0.88	7.54	7.06	1.01	23.75	0.88	7.52	7.01	1.01	23.75	0.88	7.52	7.01	1.01

2.3. Best input combination and sensitivity analysis

Apart from dataset quality, the correct selection of the input parameters has the largest impact on model performance. As explained in the application of the Buckingham π theorem, six different dimensionless parameters were obtained from the effective variables involved. In order to understand the influence of individual dimensionless parameters on model performance, different input combinations were evaluated. This influence was determined by examining the correlation coefficient (*r*) between the input parameters and output (d_s/b_1) (Table 2) for different input combinations (Table 3). The input combinations were constructed by first using the input parameter with the highest correlation coefficient (Fr_0; combination No. 1), and then creating subsequent combinations by adding each time the parameter with the next highest *r* until the parameter with the lowest *r* was finally added (combination No. 6). This approach is the common way of determining the most effective input parameters (Yaseen et al, 2016; Khozani et al, 2019; Salih et al, 2019). At first, each developed algorithm was tested using all the input combinations and with default model parameter values to determine the best input combination. Once complete, a sensitivity analysis, examining the effect of each parameter on scour depth, was performed) using the testing dataset.

Table 2. Correlation coefficients between input variables and scour depth

Input Variables	h_1/b_1	b_2/b_1	Fr_0	U_1/U_c	d_{50}/b_1	σ_g
r	0.563	-0.578	-0.624	0.330	0.292	0.331

Table 3. Different input combinations used to model scour depth

No.	Different input combination	Output
1	Fr_0	d_s/b_1
2	Fr_0, b_2/b_1	d_s/b_1
3	Fr_0, b_2/b_1, h_1/b_1	d_s/b_1
4	Fr_0, b_2/b_1, h_1/b_1, σ_g	d_s/b_1
5	Fr_0, b_2/b_{11}, h_1/b_1, σ_g , U_1/U_c	d_s/b_1
6	Fr_0, b_2/b_1, h_1/b_1, σ_g , U_1/U_c, d_{50}/b_1	d_s/b_1

2.4. Models parameter optimization

The determination of the optimum values for each model parameter has a great effect on the models predictive power. Optimal parameter values differ from study to study and there is not a global optimum value; hence, in any study, identifying optimal model parameters is an important step in the model building process. In the current study optimal values were determined using the widely accepted trial-and-error approach, and the calibration dataset (Choubin et al, 2018; Sherafati et al, 2019). The root mean square error (*RMSE*) metric was used for determination of the optimum value, as the lower the *RMSE* for the testing phase, the more effective the model performance. Optimum values for each machine learning algorithm, along with some key definitions, are presented in Table A and B in the supplementary materials.

2.5. Model theory background

2.5.1. Isotonic Regression (ISOR)

Isotonic Regression (ISOR), which is also called monotonic regression, is an approach, like any form of regression, of fitting a line through the measured data but a number of rules and restrictions apply. For example, the fitting line (isotonic curve) must be non-decreasing/non-increasing and has to be the closest distance from the measured data. The key advantages of this algorithm are that it minimizes the mean square error in the training dataset and is not restricted in a functional linearity form, such as a linear regression model, as long as the function is monotonic increasing (Barlow et al., 1972). Figure 5 illustrates the difference between a linear regression and isotonic regression model. More information about this algorithm was given in Kruskal (1964) and de Leeuw et al. (2009).

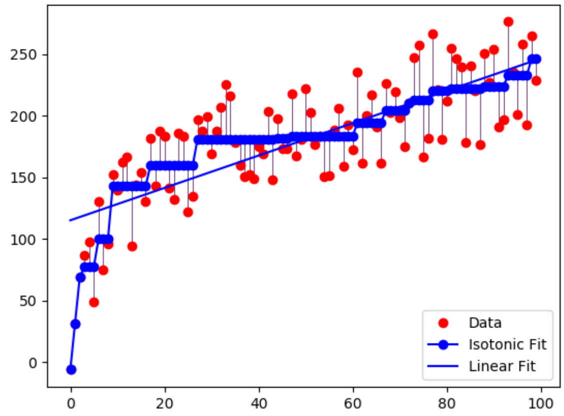


Fig.1. A comparison between an Isotonic Regression and a linear regression

2.5.2. Sequential Minimal Optimization Regression (SMOR)

Sequential Minimal Optimization Regression (SMOR) was first introduced by Platt (1999) and later improved by Smola and Schölkopf (1998) and Shevade et al. (2000) for solving very large quadratic programming issues which can occur during the training of a SVM algorithm. These programming issues are divided into a smaller series of optimization quadratic programming sub-issues based on Osuna's theorem (Osuna et al., 1997). Next, an objective function is decreased at each step until a feasible point that satisfies all of the constraints is retained (Yang et al., 2007). The SMOR computes the maximum error deviation (*MED*) between measured and predicted values; if predicted values are higher than *MED*, performance of the system components was fully satisfactory, and if predicted values are lesser than MED the model is often-overlooked (Gao et al. 2019). More information about this algorithm structure can be found in Platt (1999), Yang et al. (2007), Cheng and Qu (2013) and Yang et al. (2014).

2.5.3. Locally Weighted Learning (LWL)

Locally weighted learning (LWL) is a class of function approximation techniques, in which a prediction is made by using an approximated local model around a point of interest. This approximation is achieved by using an instance-based algorithm to assign instance weights which are then used by a specified Weighted Instances Handler. LWL can perform classification (e.g. using naive Bayes) or regression (e.g. using linear regression), and is one of the most widely used lazy leaner algorithms. LWL is a simple but appealing tool, both instinctively and statistically, for learning process dynamics of non-linear problems, due to its high flexibility. The main drawback of this algorithm is the longer time required in the modeling process, from model building to making predictions.

The algorithm is based on the following equation: if $y(r) = z(x(r), u(r))$ is considered as a non-linear event, the optimum output $u_d(r)$ can be calculated using the inverse approach of the event as follows (Arif et al., 2001):

$$u_d(r) = z^{-1}(x_d(r), y_d(r)) \tag{5}$$

where $z(.)$ is considered as a non-linear function, $x_d(r)$ are the states, and $y_d(r)$ is the optimum output. More information about this algorithm can be found in Atkeson et al. (1997).

2.5.4. Iterative Classifier Optimizer (ICO)

The Iterative Classifier Optimizer (ICO) algorithm uses a cross-validation or percentage split approach to optimize the number of iterations of the given iterative classifier. The algorithm has the ability to handle missing, nominal, binary classes and attributes like numeric, nominal, binary and empty nominal (Saad, 2018). A two iteration process is used: models are run and the results compared with measured values and then feedback is submitted to the model to further learn and fine tune the results.

2.5.5. Least Median of Squares Regression (LMS)

The Least Median of Squares Regression (LMS) algorithm implements a least median squared linear regression using the existing (weka) linear regression class to form predictions.

These LMS functions are generated from random subsamples of the data. The LMS with the lowest median squared error is chosen as the final model. More information about this algorithm is presented in Rousseeuw and Leroy (1987) and Giloni (2002).

2.5.6. Disjoint Aggregating (Dagging)

The Disjoint Aggregating (DA) algorithm is a type of meta-classifier that creates a number of disjoint, stratified folds out of the data and feeds each chunk of data to a copy of the supplied base classifier. Predictions are made via averaging, since all the generated base classifiers are put into the Vote Meta classifier. The algorithm is useful for base classifiers that are quadratic or worse in time behavior, in regard to the number of instances in the training data. The strong capabilities of this algorithm include handling missing class values, binary class, nominal class, nominal attributes, empty nominal attributes and unary attributes. More information about this algorithm can be found in Ting and Witten (1997).

2.5.7. Random Subspace

Random Subspace (RS) constructs a decision tree based classifier that maintains the highest accuracy on the training data and improves generalization accuracy as it grows in complexity (Ho, 1998). This algorithm enhances the prediction power of the weak classifier algorithms. The classifier in this algorithm consists of multiple trees constructed systematically by pseudo randomly selecting subsets of components of the feature vector, that is, trees constructed in randomly chosen subspaces. More information about this algorithm can be found in Ho (1998).

2.6. Model evaluation

After the determination of the most effective input variable combination and the optimum operator values, each algorithm was trained by a training dataset and evaluated by a testing dataset. Since the models were built by a training dataset, this evaluation can only show how well the constructed model fits the testing dataset, and cannot be used for model validation (Chen et al., 2019). For a visual analysis and assessment of the applied models, scatter plots, Taylor diagrams and box-plots were used. One distinct advantage of the Taylor diagram is that it benefits from the use of the two most common correlation statistics: correlation and standard deviation (*SD*) (Taylor, 2001). Points are depicted on the diagram to compare the performance of different developed models. The measured data point in the Taylor diagram is considered as the reference point. The closer the predicted value to this reference value, in terms of *r* and *SD*, the higher the prediction capability. The advantage of a box-plot is that it can show how well a model predicts extreme, median and quartile values.

In addition, *RMSE*, Mean Absolute Error (*MAE*), the Nash-Sutcliffe efficiency (*NSE*) were used to quantify model performance. These criteria were calculated as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[(d_{s\,m}/b_1)_i - (d_{s\,p}/b_1)_i]^2} \tag{6}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|(d_{s\,m}/b_1)_i - (d_{s\,p}/b_1)_i\right| \tag{7}$$

$$NSE = 1 - \frac{\sum_{i=1}^{N}[(d_{s\,m}/b_1)_i - (d_{s\,p}/b_1)_i]^2}{\sum_{i=1}^{N}[(d_{s\,m}/b_1)_i - \overline{d}_{s\,m}/b_1]^2} \tag{8}$$

where d_s/b_1 is the scour depth, N is the number of datasets, $\overline{d_s}/b_1$ is the mean scour depth, and *m* and *p* subscripts denote the measured and predicted values, respectively.

The lower the *RMSE* and *MAE*, the better the model performance. Model performance can be classified using the *NSE* values (between $-\infty$ and 1; Moriasi et al., 2007): (i) unsatisfactory:

NSE≤0.4; (ii) acceptable: 0.40<NSE≤0.50; (iii) satisfactory: 0.50<NSE≤0.65; (iv) good: 0.65< NSE≤0.75; (v) very good: 0.75< NSE ≤1.00.

A reliability analysis was also applied to reveal the consistency of applied models or permissible level of model performance, as follows:

$$\text{Reliability} = [(1/N) * \sum_{i=1}^{N} k_i] * 100 \tag{9}$$

The variable k_i was estimated based on the relative average error (*RAE*). If *RAE* is ≤ 0.2, then k_i =1, else k_i = 0. This threshold of 0.2 was determined based on the Chinese Standard value (Saberi-Movahed et al. 2020). RAE was calculated as follows:

$$RAE = \left| \frac{(d_{s_m}/b_1) - (d_{s_p}/b_1)}{(d_{s_m}/b_1)} \right| \tag{10}$$

3. Results and analysis

3.1. Most/least effective variables

Table 2 reveals which input parameters had the most effect on local scour depth. According to the correlation coefficients, Fr_0 had the most impact (r = -0.62) followed by b_2/b_1 (r = -0.58), h_1/b_1(r = -0.56), σ_g (r = 0.34), U_1/U_c (r = 0.33), and d_{50}/b_1 (r = 0.29).

3.2. Best input combination

Most of the algorithms performed best when all the input parameters were involved in the building of the model (input No. 6). In six scenarios, this was not the case, and input combinations 3, 4 and 5 gave the lowest *RMSE* values correlation coefficients between observed and predicted scour depth. This contrast reflects the different structures of the algorithms.

A comparison in r values between combinations 1 and 2 shows that adding the b_2/b_1 parameter caused the prediction accuracy to significantly decrease for the majority of algorithms. In contrast, adding the h_1/b_1 parameter enhanced the prediction performance of all algorithms significantly (comparison between input No. 2 and 3). The effect of adding σ_g was more mixed, causing an increase in some cases, and little change in others (see the comparison between No. 3 and No. 4). Adding U_1/U_c and d_{50}/b_1 caused the model performance to improve in the majority of cases.

The models with just one or two input parameters failed to provide accurate predictions of scour depth. This poor performance occurred because some parameters had a poor linear correlation with d_s/b_1, since the correlation was non-linear. However incorporating these parameters into the models that can handle non-linear relationships enhanced the model accuracy significantly.

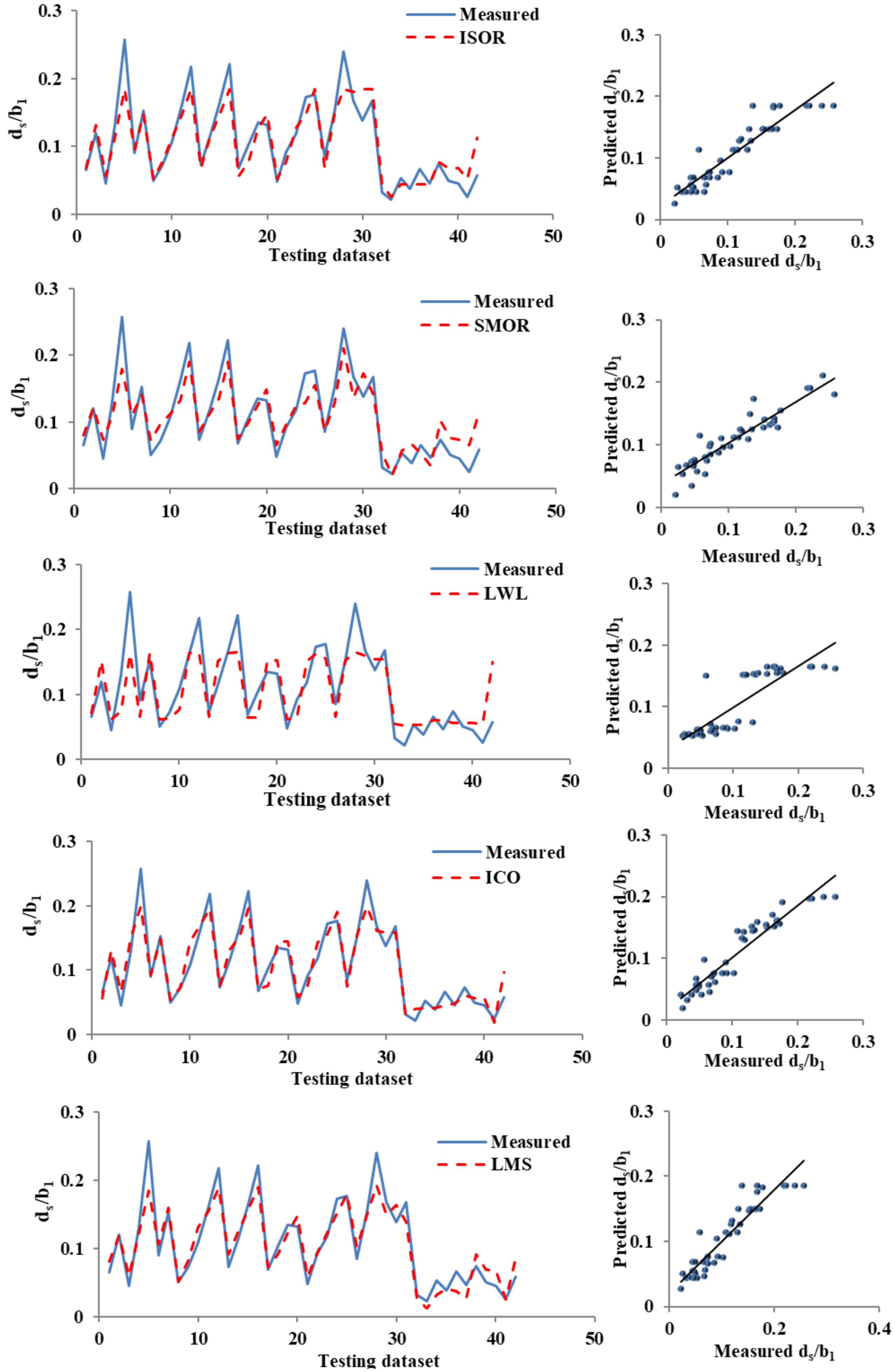
3.3. Models performance evaluation

A visual comparison of the prediction power of the machine learning models is shown in Figure 2. These plots reveal that all developed algorithms could predict scour depth reasonably well, but all algorithms underestimated the maximum scour depth, except DA-ICO which predicted scour depth almost perfectly. The LWL algorithm by contrast had the weakest performance. Among the standalone algorithms, the ICO model provided slightly better performance than others.

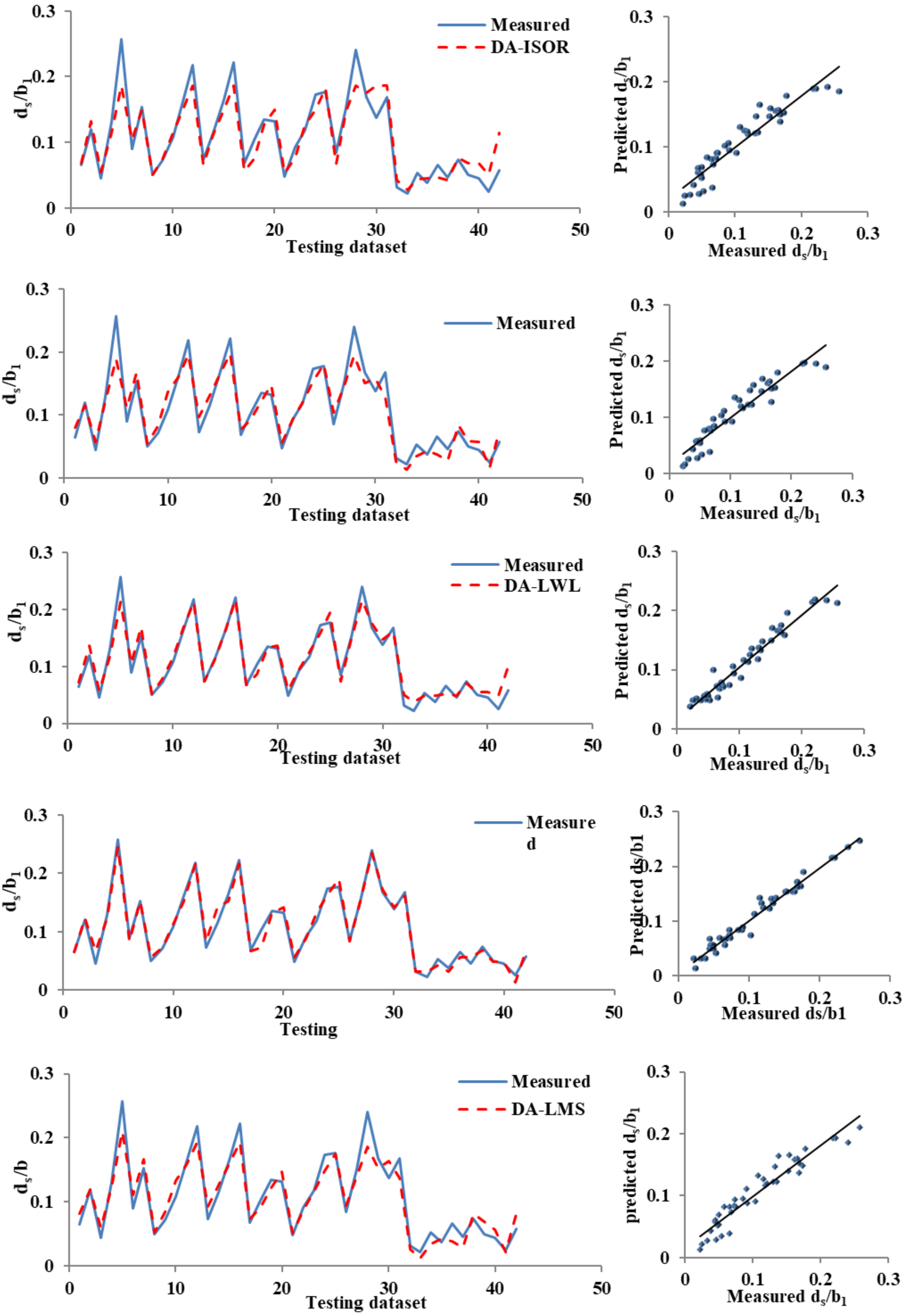
Table 4. Correlation coefficient values of model performance for different input combinations

Models	Input number.											
	1		2		3		4		5		6	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing
ISO	0.92	0.928	0.92	0.929	0.92	0.929	0.902	0.93	0.92	0.933	0.902	0.935
SMO	0.6	0.8	0.68	0.74	0.85	0.905	0.83	0.88	0.912	0.944	0.913	0.945
LWL	0.82	0.81	0.83	0.82	0.86	0.849	0.856	0.843	0.858	0.844	0.862	0.84
ICO	0.9	0.9	0.92	0.91	0.934	0.93	0.943	0.946	0.938	0.927	0.936	0.92
LMS	0.6	0.8	0.684	0.742	0.81	0.865	0.81	0.865	0.815	0.865	0.836	0.881
DA-ISO	0.92	0.931	0.916	0.93	0.916	0.93	0.916	0.934	0.916	0.939	0.916	0.942
DA-SMO	0.61	0.81	0.68	0.75	0.84	0.898	0.834	0.889	0.913	0.948	0.914	0.951
DA-LWL	0.86	0.87	0.87	0.87	0.88	0.87	0.887	0.875	0.891	0.886	0.892	0.896
DA-ICO	0.91	0.928	0.93	0.94	0.942	0.946	0.95	0.95	0.949	0.948	0.949	0.943
DA-LMS	0.6	0.81	0.672	0.777	0.81	0.863	0.81	0.879	0.861	0.918	0.879	0.917
RS-ISO	0.92	0.93	0.844	0.843	0.91	0.93	0.919	0.93	0.866	0.866	0.934	0.94
RS-SMO	0.6	0.8	0.68	0.707	0.76	0.88	0.752	0.89	0.838	0.935	0.736	0.867
RS-LWL	0.82	0.81	0.83	0.826	0.87	0.87	0.884	0.88	0.876	0.862	0.877	0.874
RS-ICO	0.902	0.904	0.927	0.913	0.941	0.935	0.939	0.931	0.946	0.945	0.949	0.951
RS-LMS	0.600	0.800	0.684	0.746	0.810	0.865	0.810	0.865	0.815	0.865	0.836	0.881

* Shadow cells show optimum input number



Measured
ISOR
SMOR
LWL
ICO
LMS
d_s/b_1
Testing dataset
Predicted d_s/b_1
Measured d_s/b_1



Measured
DA-ISOR
d_s/b_1
Testing dataset
Predicted d_s/b_1
Measured d_s/b_1
Measured
d_s/b_1
Testing dataset
Predicted d_s/b_1
Measured d_s/b_1
Measured
DA-LWL
d_s/b_1
Testing dataset
Predicted d_s/b_1
Measured d_s/b_1
Measure
d
d_s/b_1
Testing
Predicted ds/b1
Measured ds/b1
Measured
DA-LMS
d_s/b
Testing dataset
predicted d_s/b_1
Measured d_s/b_1

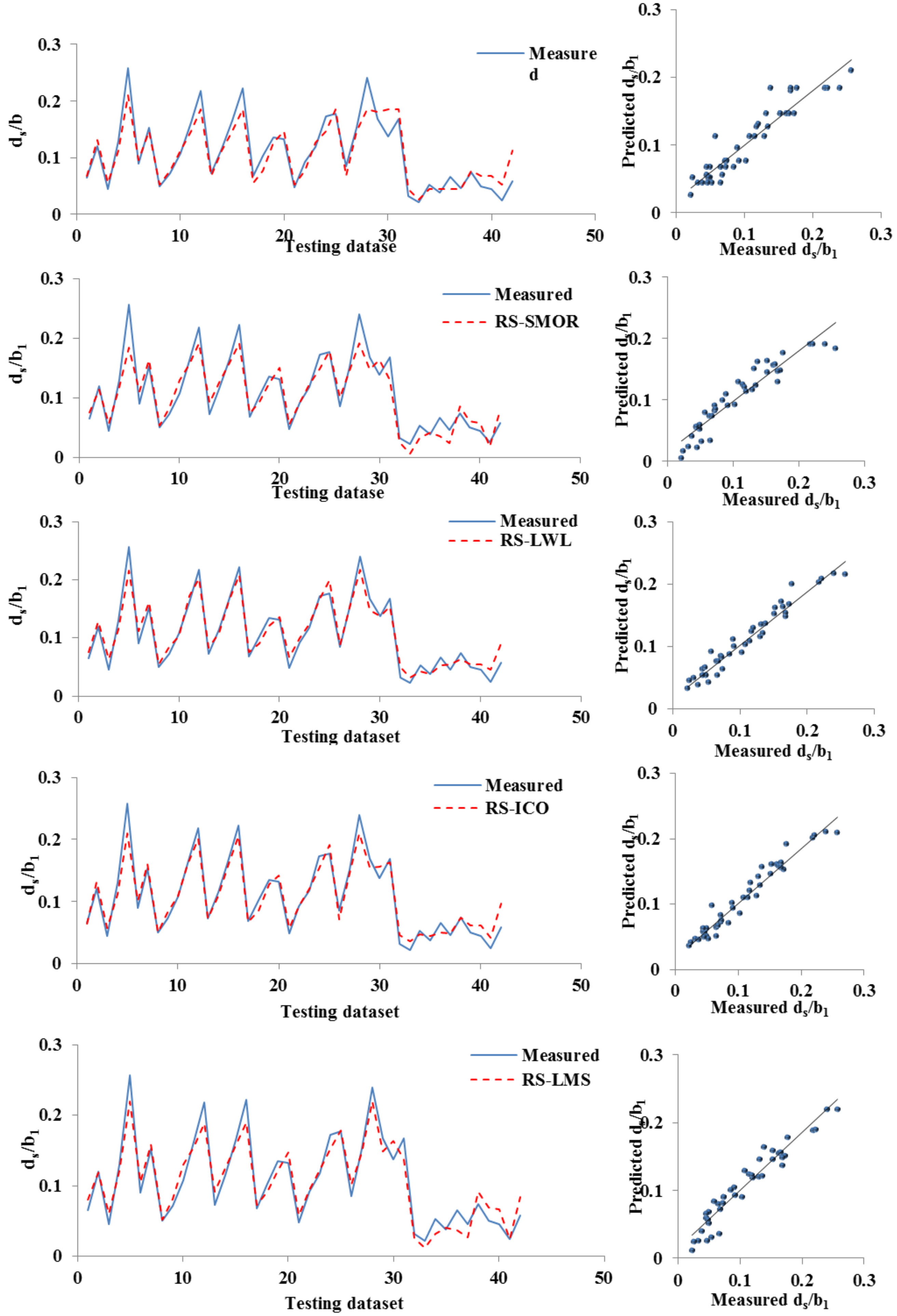


Fig 2. Comparison between predicted and measured d_s/b_1 values during the testing phase.

To further compare the performance of the standalone and hybridized machine learning models, box-plots are shown in Figure 3. Only the DA-ICO model predicted maximum scour depth well (but not perfectly) and all other algorithms underestimated maximum values. The SMOR and ICO models estimated minimum scour depth very well. The first quartile value was very well replicated by the DA-LWL model, and ICO, DA-ICO and RS-ICO models predicted the third quartile most accurately. The RS-ICO model estimated the median scour depth almost perfectly.

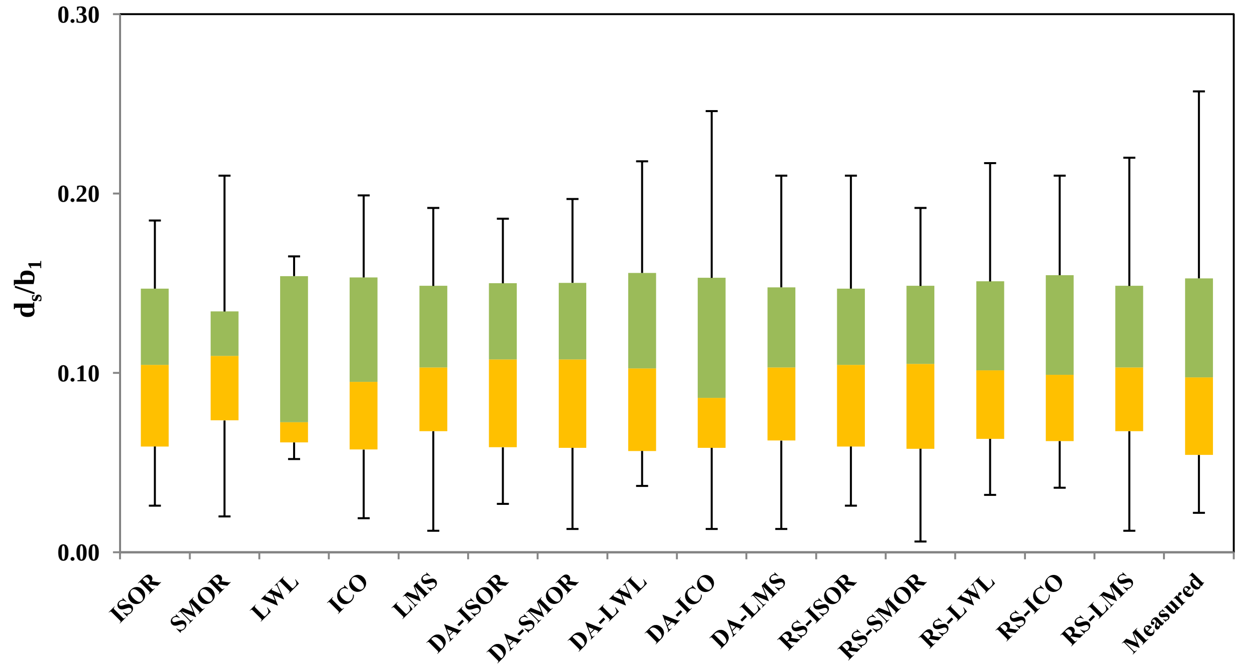


Fig. 3. Box plots of observed and predicted d_s/b_1.

The DA-ICO model was closest to the observed reference point in the Taylor plot (Figure 4). This plot showed that the DA-ICO model had the highest performance because the predicted standard deviation of scour depth was closest to the standard deviation of the observed data, and the correlation was also the highest. The DA-LWL, RS-LWL and RS-ICO points almost overlapped on each other on the Taylor plot, indicating similar model performance.

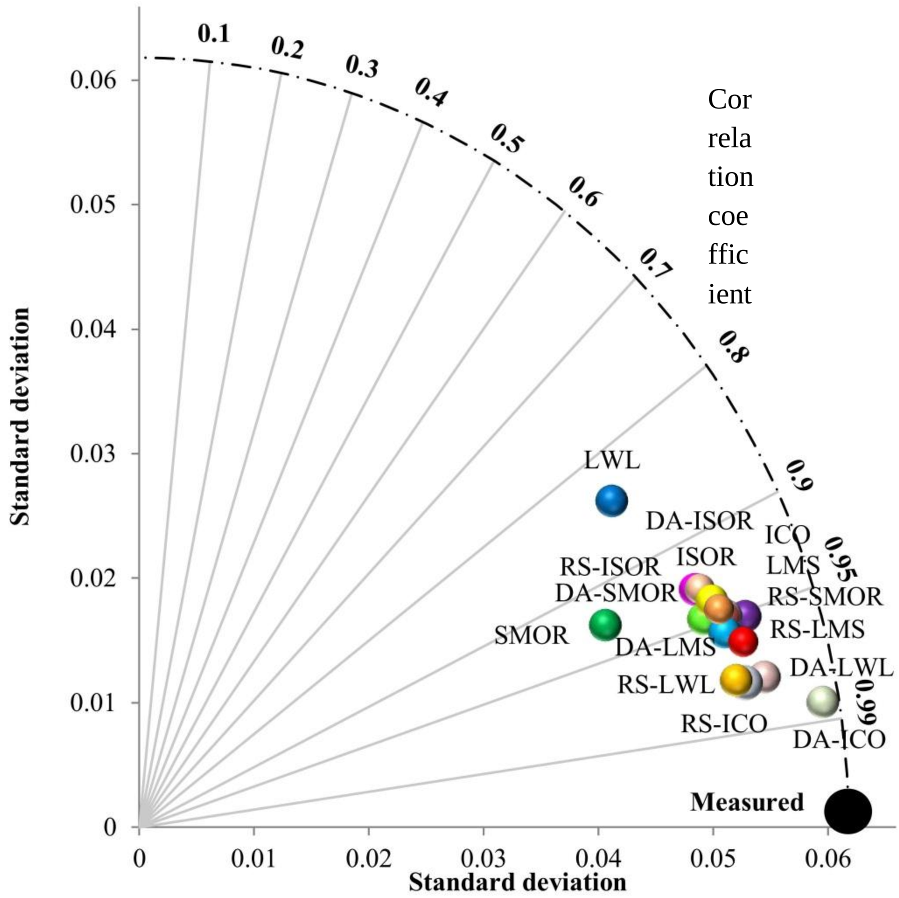
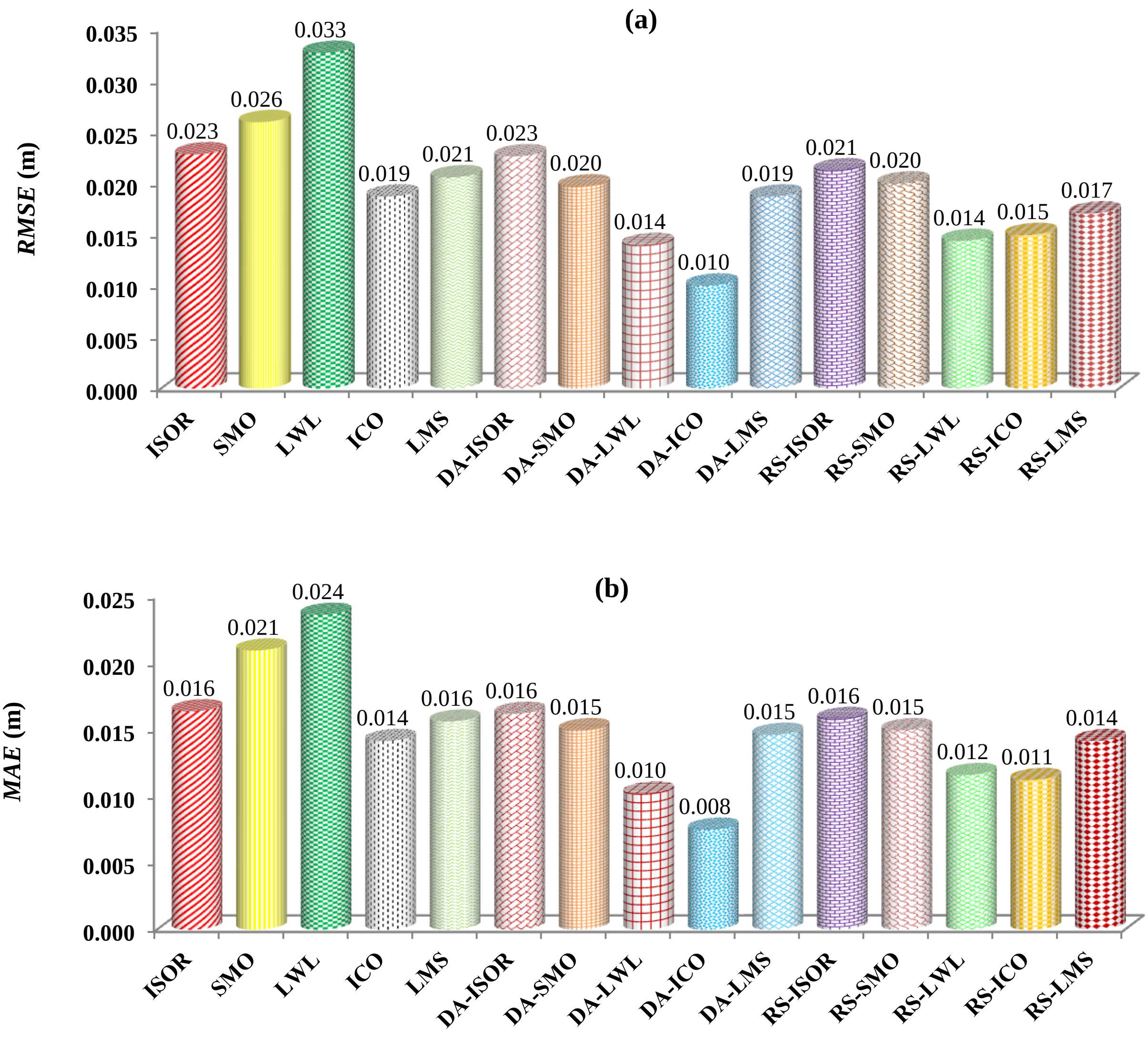


Fig.4. Taylor plot of model performance

A quantitative assessment of the predictive capability of the developed models is shown in Figure 5. Looking across all the evaluation metrics, the DA-ICO model provided the strongest performance with *RMSE*, *MAE* and *NSE* values of 0.010 m, 0.008 m and 0.972, respectively. The next strongest performing model was DA-LWL, followed by RS-LWL, RS-ICO, RS-LMS, DA-LMS=ICO, DA-SMOR=RS-SMOR, RS-ISOR=LMS, DA-ISOR=ISOR, SMOR and LWL. According to the *NSE* metric, all the developed algorithms had a ‘very good’ prediction power except the LWL model, which had ‘good’ performance.

The standalone LWL and SMO models gave the lowest performance with higher *RMSE* and *MAE*, and low *NSE* values. Among the standalone models, ICO produced the most accurate predictions. All the metrics reveal the hybridized version of the machine learning algorithm outperformed the standalone counterpart. For example, the standalone LWL model had a *RMSE* value of 0.033 m, while the DA-LWL and RS-LWL models produced a *RMSE* value of 0.014 m. This hybridization represented a 58% improvement in standalone LWL performance. In terms of *RMSE* values, hybridization with DA and RS algorithms enhanced the performance of SMOR models by 23%, ICO by 41% (DA) and 21% (RS), LMS models by 10% (DA) and 19% (RS), and ISOR by 9% (RA). The *RMSE* values were the same for ISOR and DA-ISOR models.



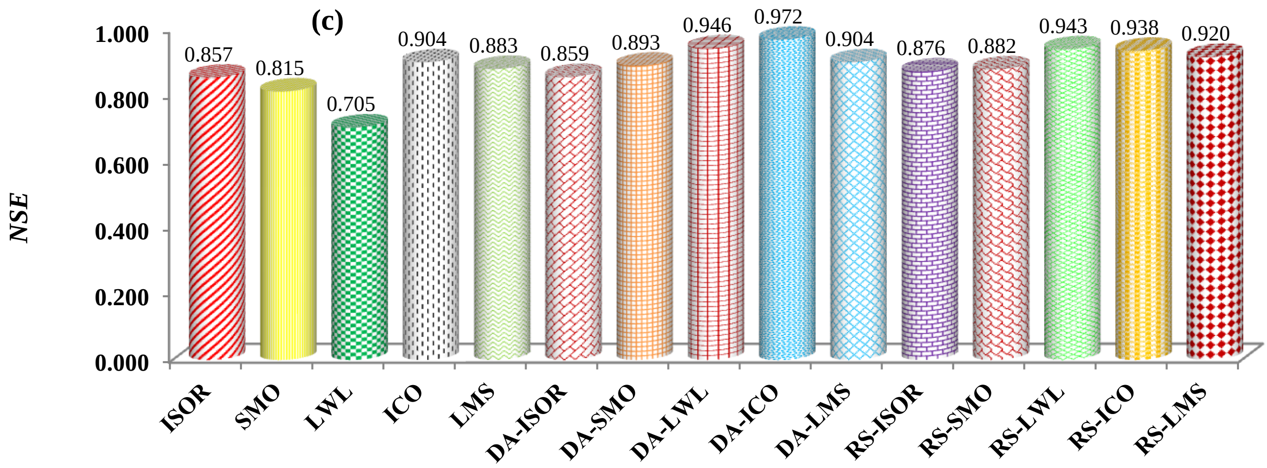


Fig. 5. Model evaluation using quantitative criteria

A Reliability analysis showed that the DA-ICO hybrid model had the highest level of reliability (83.3%), while the standalone model of LWL had the lowest degree of reliability (45.2%). In all but one case (DA-ISOR) hybridization enhanced reliability (Table 5).

Table 5. Model reliability.

Model	Reliability (%)
ICO	71.4
ISOR	73.8
LMS	64.3
LWL	45.2
SMOR	59.5
DA-ICO	83.3
DA-ISOR	69.0
DA-LMS	69.0
DA-LWL	78.6
DA-SMOR	64.3
RS-ICO	81.0
RS-ISOR	76.2
RS-LMS	71.4
RS-LWL	81.0
RS-SMOR	69.0

4. Discussion

Scouring takes place in clear water flow due to a contraction resulting from a change in channel cross-section shape. For a given discharge, reducing the channel width causes an increase in flow velocity, bed shear stress and thus scour. In order to compute the scour depth in such a hydraulic condition, several empirical equations have been proposed in the literature, developed using conventional regression analysis. These equations are based on the most important variables considered to affect scour depth, such as channel width, flow depth and velocity in the un-contracted and contracted sections of the channel. These empirical equations do not perform as accurately as machine learning algorithms (Najafzadeh et al. 2018). However no study has examined the performance of hybrid machine learning algorithms in predicting scour depth. To this end, this study sought to provide the first comparison of the accuracy of these novel algorithms with standalone machine learning algorithms and empirical equations.

The determination of the best input variable combination is one of the most critical steps in producing an accurate machine learning model. Some researchers have determined the best input combination according to the highest correlation coefficient in a multiple regression (Barzegar et al. 2016). However the current paper shows this approach for scour prediction is not the best to take because the best input combination was not the same for all algorithms and, due to nonlinearity between variables, the variables with low correlation coefficients with scour depth enhanced the prediction power of the some of the models. Thus a range of different input variable combinations must be considered in the optimization of machine learning models.

In order to find the best input combination, a sensitivity analysis was carried out to investigate the importance of each dimensionless parameter in scour depth computation. The results revealed that upstream, unconstructed densimetric particle Froude number had the greatest impact on scour depth prediction, in accordance to previous laboratory experiments (Li et al., 2002). This result is also consistent with Dey and Raikar (2005) who illustrated that the upstream densimetric particle Froude number had the greatest effect on scour depth, in which an increase in this number decreased the scour depth. This parameter represents the impact of the mobility of submerged sediment particles on scour depth (Dey and Raikar, 2005) and has been proposed as suitable for defining the initiation of sediment motion (Aguirre-Pe et al., 2003). Thus it is reasonable to expect the particle Froude number to have this important impact on scour depth. In our study we estimated the densimetric particle Froude number based on the ratio between the approach flow velocity and submerged weight of the sediment. The particle Froude number can also be obtained, when data is available, through analyzing the hydrodynamic forces - lift, drag resistance and submerged weight - acting on the sediment particle in equilibrium conditions (Safari et al., 2017). The drag and lift forces have a positive effect on sediment motion, and the buoyed weight of the sediment and resistance force have the opposite effect. Therefore, to initiate scour, the flow must have adequate force to overcome this buoyed weight and resistance force. An alternative method to estimate the particle Froude number is to combine the Shields sediment threshold equation (1936) with the Manning (1891) flow resistance formula. In this formulation the particle Froude number is an alternative type of Shields (1936) threshold parameter, expressed in terms of velocity rather than shear stress (Safari et al., 2015).

The DA-ICO was found to provide the most accurate predictions of scour depth using four input parameters: Fr_0, b_2/b_1, h_1/b_1 and σ_g. For most of the other models, six input parameters were required to provide the optimum prediction performance. Thus a further advantage of the DA-ICO is in its relative simplicity, using parameters that are more readily and easily measured, removing the requirement to measure the approach and critical flow velocity. The

ISOR, DA-ISOR and RS-ISOR algorithms gave reasonable prediction accuracy with only Fr_0 as an input, demonstrating their efficiency.

The hybrid models had a higher prediction power than standalone models because hybrid models are more flexible than standalone models and have a nonlinear structure (De'ath and Fabricius, 2000). These two model properties are particularly important in the prediction of scour depth because of the nonlinearity between variables. The LWL algorithm was the worst performing because in the LWL algorithm, fitting the noise data causes a higher prediction error if noise in the dataset is not filtered out well, and the algorithm does not have the ability to interpolate smoothly between datasets (Schneiderm and Moore, 1997).

The difference in performance between algorithms is attributable to their different computational structures. The DA-ICO model provided the most accurate predictions for two reasons. First, the ICO algorithm uses a cross-validation or percentage split approach to optimize the number of iterations. Secondly, the Dagging algorithm benefits from ensemble learning in its structure (multiple weak learners) which outperforms a single strong learner (Dietterichm 1997). This learning helps to reduce variance and avoid the over-fitting problem caused by the use of a bootstrap procedure.

Najafzadeh et al. (2016, 2018), using the same datasets as the current paper, compared the performance of traditional AI-based algorithms (SVM, ANFIS, GEP, EPR and MT) with the empirical equations of Laursen (1963), Komura (1966), Gill (1981) and Lim (1993). Table 6 shows this comparison, along with the current paper's *RMSE* value for the best performing hybrid machine learning algorithm. The table shows that the GEP model (*RMSE* = 0.027 m) outperformed all empirical equations, and the newly developed DA-ICO model (*RMSE* = 0.010 m) performed significantly better than the GEP model (a 62% improvement in *RMSE*). When compared to the empirical equations in terms of *RMSE* values, the DA-ICO model had

82% (Laursen, 1963), 93% (Lim, 1993), 95% (Komura, 1966) and 99% (Gill, 1981) better accuracy.

Table 6. A comparison of model performance between traditional AI-based algorithms, GEP, MT and EPR Najafizadeh et al. (2016, 2017), empirical equations (Laursem, 1963; Komura, 1966; Gill, 1981; Lim, 1993) and DA-ICO, the best performing hybrid machine learning algorithm in the current study.

Model	*RMSE* (m)	Model	*RMSE* (m)	Model	*RMSE* (m)
GEP	0.0260	SVM	0.028	Komura (1966)	0.0833
MT	0.0296	ANFIS	0.0281	Gill (1981)	0.200
EPR	0.0263	Laursen (1963)	0.0543	Lim (1993)	0.134

Overall, the results show that DA-ICO models have great potential to produce robust predictions of scour depth in long contractions in clear-water conditions. As well as offering far superior prediction accuracy than existing empirical and traditional AI-models, a distinct strength of this model is the need for just four readily measured dimensionless variables: densimetric particle Froude number, width of the un-contracted section, approach flow depth and sediment geometric standard deviation. Thus this type of data-driven model could be of real practical benefit to engineers required to estimate maximum scour depth when designing bridges, weirs, spur dike and cofferdams. Future studies should consider the performance of these algorithms in the prediction of scour depth in more complex conditions, such as within natural rivers, than those featured in the studied datasets, such as, live-bed scour, non-rectangular channels, unsteady flows, non-equilibrium transport conditions, and water-worked beds that mimic better the surface topographies of natural coarse-grained rivers (Cooper and Tait, 2009).

5. Conclusion

The accurate prediction of scour depth is vital for preventing the collapse of in-channel structures. Due to the non-linear behavior of sediment transport in a river, hybrid machine learning algorithms have great potential to produce accurate predictions of scour depth in long contractions. Using previously collected scour depth data from laboratory flume experiments, this study tested this potential for the first time by comparing the prediction power of five standalone algorithms, Isotonic Regression (ISOR), Sequential Minimal Optimization (SMO), Iterative Classifier Optimizer (ICO), Locally Weighted learning (LWL) and Least Median of Squares Regression (LMS) and their hybrid versions with Dagging (DA) and Random Subspace (RS) algorithms (i.e., DA-ISOR, DA-SMOR, DA-LWL, DA-ICO, DA-LMS, RS-ISOR, RS-SMOR, RS-LWL, RS-ICO, RS-LMS). The main findings were as follows:

(1) A test of model performance showed that the DA-ICO model had the highest prediction power followed by DA-LWL, RS-LWL, RS-ICO, RS-LMS, DA-LMS, ICO, DA-SMOR=RS-SMOR, RS-ISOR=LMS, DA-ISOR=ISOR, SMOR and LWL. All models displayed ‘very good’ performance except the LWL model, which had ‘good’ performance.

(2) The hybrid models had a higher prediction power than standalone models because the hybrid models are more flexible and have a nonlinear structure that better represents the nonlinear behavior of sediment transport.

(3) All algorithms underestimated the maximum scour depth, except DA-ICO which predicted scour depth almost perfectly.

(4) A sensitivity analysis revealed that scour depth was most sensitive to the densimetric particle Froude number followed by the non-dimensionalized contraction width, flow

depth within the contraction, sediment geometric standard deviation, approach flow velocity and median grain size.

(5) Most of the algorithms performed best when all the input parameters were involved in the building of the model. An important exception was the best performing model, which required only four input parameters: densimetric particle Froude number and non-dimensionalized contraction width, flow depth within the contraction and sediment geometric standard deviation.

(6) Variables with low correlation coefficients with scour depth enhanced the prediction power of the some of the models. Thus a range of different input variable combinations must be considered in the optimization of machine learning models.

Overall the results revealed that hybrid machine learning algorithms provide more accurate predictions of scour depth than empirical equations and traditional AI-algorithms. The DA-ICO model not only created the most accurate predictions but also used the fewest easily and readily measured input parameters. Thus this type of model could be of real benefit to practicing engineers required to estimate maximum scour depth when designing in-channel structures. In this case, understanding more about the potential for hybrid machine learning algorithms to provide relatively cheap and fast predictions of scour depth in more complex hydro-sedimentary conditions represents a vital research avenue.

References

Aguirre-Pe J, Olivero ML, Moncada AT (2003) Particle densimetric Froude number for estimating sediment transport. J Hydraul Eng 129(6):428–437

Ashida, K. 1963. Study on the stable channel through constrictions. Annual Report, Disaster Prevention Research Institute, Kyoto University, Kyoto, Japan, 1-15.

Atkeson, C., Moore, A., Schaal, S. 1997. Locally Weighted Learning, Artificial Intelligence Review, 11:11-73.

Ayoubloo, M. K., A. Etemad-Shahidi, and J. Mahjoobi. 2010. Evaluation of regular wave scour around a circular pile using data mining approaches. Applied Ocean Research 32 (1), 34–39. doi:10.1016/j.apor.2010.05.003

Azamathulla, H. M., A. Ab Ghani, N. A. Zakaria, and A. Guven. 2010. Genetic programming to predict bridge pier scour. Journal of Hydraulic Engineering, ASCE 136 (3), 165–69.

Azmathullah, H.M., Ghani, A.A.B., Zakaria, N.A., 2009. ANFIS based approach forpredicting maximum scour location of spillway. Water Manag. ICE Lond. 162(6), 399–407.

Barlow, R. E.; Bartholomew, D. J.; Bremner, J. M.; Brunk, H. D. (1972). Statistical inference under order restrictions; the theory and application of isotonic regression. New York: Wiley. ISBN 978-0-471-04970-8.

Bui, A Shirzadi, K Chapi, H Shahabi, B Pradhan, BT Pham, 2019b A hybrid computational intelligence approach to groundwater spring potential mapping. Water 11 (10), 2013

Bui, D., Khosravi, K., Karimi, M., Khozani, Z., Nguyen, H. et al. 2020b.Enhancing nitrate and strontium concentration prediction in groundwater by using new data mining algorithm. Science of The Total Environment, 136836

Bui, D., Khosravi, K., Tiefenbacher, J., Nguyen, H., Kazakis, N.2020. Improving prediction of water quality indices using novel hybrid machine-learning algorithms. Science of The Total Environment, 13761.

Bui, D., Pradhan, B., Nampak, H., Bui, Q.-T., Tran, Q.-A., Nguyen, Q.-P., 2016. Hybrid artificial intelligence approach based on neural fuzzy inference model and metaheuristic optimization for flood susceptibility modeling in a high-frequency tropical cyclone area. Journal of hydrology.

Bui,D., Shahabi, H., Omidvar,E., Shirzadi, A., Geertsema, M., Clague, J., Khosravi,K., Pradhan, B., Pham, B., Chapi, K., Barati, Z., Bin Ahmad, Rahmani, H., Lee, S. 2019a. Shallow landslide prediction using a novel hybrid functional machine learning algorithm. Remote Sensing, 11 (8): 931

Chan-Yun Yang, Kuo-Ho Su and Gene Eu Jan, "An elaboration of sequential minimal optimization for support vector regression," *2014 IEEE International Conference on System Science and Engineering (ICSSE)*, Shanghai, 2014, pp. 88-93

Chen, W., Panahi, M., Khosravi, K., Poughasemi, H.R., Rezaei, F. 2019. Spatial prediction of groundwater potentiality using ANFIS ensembled with teaching-learning-based and biogeography-based optimization, Journal of Hydrology 572, 435-448

Cheng, L., Qu, X. 2013. Advanced Technologies on Measure and Diagnosis, Manufacturing Systems and Environment Engineering, Applied Mechanics and Materials, 329: https://doi.org/10.4028/www.scientific.net/AMM.329.472

Choubin, B., Darabi, H., Rahmati, O., Sajedi-Hosseini, F., Kløve, B., 2018. River suspended sediment modelling using the CART model: A comparative study of machine learning techniques. Sci. Total Environ. 615, 272–281.https://doi.org/10.1016/j.scitotenv.2017.09.293

Cooper, J.R., Tait S.J. 2009. Water-worked gravel beds in laboratory flumes: a natural analogue? Earth Surface Processes and Landforms. 34, 384–397, doi: 10.1002/esp.1743D Tien Bui, H Shahabi, E Omidvar, A Shirzadi, M Geertsema, JJ Clague, ...2019a Shallow landslide prediction using a novel hybrid functional machine learning algorithm. Remote Sensing 11 (8), 931

De Leeuw J, Hornik K, Mair P (2009). "Isotone Optimization in R: Pool-Adjacent-Violators Algorithm (PAVA) and Active Set Methods." Journal of Statistical Software, 32(5), 1–24. URL http://www.jstatsoft.org/v32/i05/.

Dey, S. 1997. "Local scour at piers, part 1: A review of development of research." Int. J. Sediment Res., 122, 23–44.

Dey, S., and R. V. Raikar. 2005. Scour in long contractions. Journal of Hydraulic Engineering, ASCE 131 (12), 1036–49. doi:10.1061/(asce)0733-9429(2005)131:12(1036)

Dietterich, T.G. 1997. Machine learning research: Four current directions. AI Mag. 18(4), 97–136.

Etemad-Shahidi, A., and N. Ghaemi. 2011. Model tree approach for prediction of pile groups scour due to waves. Ocean Engineering. 38, 1522–27. doi:10.1016/j.oceaneng.2011.07.012

Firat, M. 2009. Scour depth prediction at bridge piers by ANFIS approach. Proceedings of the Institution of Civil Engineers - Water Management, Volume 162 Issue 4, August 2009, pp. 279-288

Ghaleh Nou, M. Azhdary Moghaddam, M. Shafai Bajestan, H. Md. Azamathulla; Estimation of scour depth around submerged weirs using self-adaptive extreme learning machine. *Journal of Hydroinformatics* 1 November 2019; 21 (6): 1082–1101. doi: https://doi.org/10.2166/hydro.2019.070

Gi;oni, A., Padberg, M. 20002.Least trimmed squares regression, least median squares regression, and mathematical programming. Mathematical and Computer Modelling, 35:1043-1060

Gill, M. A. 1981. Bed erosion in rectangular long contraction. Journal of the Hydraulics Division ASCE 107 (3), 273–84.

Guven, A., and M. Gunal. 2008. Genetic programming approach for prediction of local scour downstream hydraulic structures. Journal of Irrigation and Drainage Engineering, 134 (2), 241–49. doi:10.1061/(asce)0733-9437(2008)134:2(241)

Ho, T.K., 1998. The random subspace method for constructing decision forests. IEEE Trans. Pattern Anal. Mach. Intell. 20, 832–844. https://doi.org/10.1109/34.709601

K Chapi, VP Singh, A Shirzadi, H Shahabi, DT Bui, BT Pham, K Khosravi. A novel hybrid artificial intelligence approach for flood susceptibility assessment. Environmental modelling & software 95, 229-245

Khosravi K, L Mao, O Kisi, ZM Yaseen, S Shahid. 2018a. Quantifying hourly suspended sediment load using data mining models: case study of a glacierized Andean catchment in Chile. Journal of Hydrology 567, 165-179

Khosravi, K., Pham, B.T., Chapi, K., Shirzadi, A., Shahabi, H., Revhaug, I., Prakash, I., Tien Bui, D., 2018a. A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran. Sci. Total Environ. https://doi.org/10.1016/j.scitotenv.2018.01.266

Khosravi, K., Barzegar R, Miraki S, Adamowski J, Daggupati P, Pham, B. 2020b. Stochastic Modeling of Groundwater Fluoride Contamination: Introducing Lazy Learners, Groundwater. In press.

Khosravi, K., Cooper, J.R., Daggupati, P., Pham, B., Bui, D.T. 2020a. Bedload transport rate prediction: Application of novel hybrid data mining techniques. Journal of Hydrology, 124774. In press. https://doi.org/10.1016/j.jhydrol.2020.124774

Khosravi, K., Daggupati P, Alami MT, Awadh SM, Ghareb MI, Panahi M, et al. 2019. Meteorological data mining and hybrid data-intelligence models for reference evaporation simulation: A case study in Iraq. Computers and Electronics in Agriculture 167, 105041

Khozani, Z., Khosravi, K., Pham, B. T., Kløve, B., Wan Mohtar, W. H. M. & Yaseen, Z. M. 2019. Determination of compound channel apparent shear stress: application of novel data mining models. Journal of Hydroinformatics, doi:10.2166/hydro.2019.037

Kisi, O., Ozkan, C., Akay, B., 2012b. Modeling discharge-sediment relationship using neural networks with artificial bee colony algorithm. J. Hydrol. 428–429, 94–103. https://doi.org/10.1016/j.jhydrol.2012.01.026

Komura, S. 1966. Equilibrium depth of scour in long constrictions. Journal of the Hydraulics Division ASCE 92 (5), 17–38.

Komura, S. (1966). Equilibrium depth of scour in long constrictions."J. Hydr. Div., ASCE, 92(5), 17-37.

Kruskal, J. B. 1964. Nonmetric Multidimensional Scaling: A numerical method. Psychometrika. 29 (2): 115–129. doi:10.1007/BF02289694.

Laursen, E. M. 1963. An analysis of relief bridge scour. Journal of the Hydraulics Division ASCE 89 (3), 93–118.

Laursen, E. M. (1960). Scour at bridge crossing. J. Hydr. Div., ASCE, 86(2), 39-54.

Legates, D.R., Mccabe, G.J., 1999. Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. water Resour. Res. 35, 233–241.

Lim, S. Y. 1993. Clear water scour in long contractions. Proceeding of the Institution Civil Engineerings-Waters, Maritime and Energy, London, UK 101 (6), 93–98.

Lim, S. Y., and Cheng, N. S. 1998. Scouring in long contractions. Journal of Irrigation and Drainage Engineering ASCE 124 (5), 258–61. doi:10.1061/(asce)07339437(1998)124:5(258)

Maier, H.R., Kapelan, Z., Kasprzyk, J., Kollat, J., Matott, L.S., Cunha, M.C., Dandy, G.C., Gibbs, M.S., Keedwell, E., Marchi, A., Ostfeld, A., Savic, D., Solomatine, D.P., Vrugt, J.A., Zecchin, A.C., Minsker, B.S., Barbour, E.J., Kuczera, G., Pasha, F., Castelletti, A., Giuliani, M., Reed, P.M., 2014. Evolutionary algorithms and other metaheuristics in water resources: current status, research challenges and future directions. Environ. Modell. Software 62, 271–299. https://doi.org/10.1016/j.envsoft.2014.09.013

Manning, R. (1891). "On the flow of water in open channels and pipes". Transactions of the Institution of Civil Engineers of Ireland. 20: 161–207.

Melesse, A.M., Ahmad S, McClain ME, Wang X, Lim YH. 2011. Suspended sediment load prediction of river systems: An artificial neural network approach. Agricultural Water Management 98 (5), 855-866

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Binger, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50, 885–900. https://doi.org/10.13031/2013.23153

Moussa, Y.A. 2013. Modeling of local scour depth downstream hydraulic structures in trapezoidal channel using GEP and ANNs. Ain Shams Engineering Journal, 4:717-722.

Muzzammil, M. 2008. Application of Neural Networks To Scour Depth Prediction at The Bridge Abutments, Engineering Applications of Computational Fluid Mechanics, 2:1, 30-40, DOI:10.1080/19942060.2008.11015209

Najafzadeh, M., A. Etemad-Shahidi, and S.-Y. Lim. 2016. Prediction of scour depth in long contractions using ANFIS and SVM. Ocean Engineering, Elsevier 111, 128–35.

Najafzadeh, m., Shiri, J., Rezaei Balf, M. 2018. New Expressions-Based Models to Estimate Scour Depth at Clear Water Conditions in Rectangular Channels. Marine Georesources & Geotechnology, http://dx.doi.org/10.1080/1064119X.2017.1303009

Najafzadeh, M., G. A. Barani, and H. M. Azamathulla. 2013. GMDH to prediction of scour depth around vertical piers in cohesive soils. Applied Ocean Research 40, 35–41. doi:10.1016/j.apor.2012.12.004

Najafzadeh, M., G. A. Barani, and M. R. Hessami-Kermani. 2014. Estimation of pipeline scour due to waves by the group method of data handling. Journal of Pipeline Systems Engineering and Practice, ASCE 5 (3), 06014002. doi:10.1061/(asce)ps.1949-1204.0000171

Najafzadeh, M., Saberi-Movahed, F. 2019. GMDH-GEP to predict free span expansion rates below pipelines under waves. Marine Georesources & Geotechnology, 37(3), 375-392.

Onen, F. 2014. Prediction of Scour at a Side-Weir with GEP, ANN and Regression Models. Arabian Journal for Science and Engineering, 39:6031-6041.

Osuna, E.and Freund, R. "An Improved Training Algorithm for Support Vector Machines", Proceeding of the 1997IEEE Workshop on Neural Network for Signal Processing, New York: IEEE, 1997, pp. 276-285.

Pal, M., Singh, N.K., Tiwari, N.K., 2014. Kernel methods for pier scour modeling using fi eld data. J. Hydroinf. 16 (4), 784–796.

Parsaie, A., Haghiabi, A.H. & Moradinejad, A. 2019. Prediction of Scour Depth below River Pipeline using Support Vector Machine. KSCE J Civ Eng 23, 2503–2513. https://doi.org/10.1007/s12205-019-1327-0

Pham, B., Prakash, I., Khosravi, K., Chapi, K., Trinh, T., Ngo, T., Hosseini, V., Bui, D. 2019. A comparison of Support Vector Machines and Bayesian algorithms for landslide susceptibility modelling. Geocarto international, 34 (13): 1385-1407

Platt JC. 1999. Fast training of support vector machines using sequential minimal optimization. In Advances in Kernel Methods: Support Vector Machines, Schölkopf B, Burges C, Smola AJ (eds). MIT press: Cambridge, MA; 185–208.

Rady, R. 2020. Prediction of local scour around bridge piers: artificial-intelligence-based modeling versus conventional regression methods. Appl Water Sci 10, 57. https://doi.org/10.1007/s13201-020-1140-4

Raikar, R. V. 2004. Local and general scour of gravel beds. PhD thesis, Dept. of Civil Engineering, Indian Institute of Technology, Kharagpur, India.

Rousseeuw, P., Leory, A. 1987. Robust regression and outlier detection. Wiley Series in Probability and Statistics, |DOI:10.1002/0471725382

Saad, A. 2018. An Efficient Classification Algorithms for Image Retrieval Based Color and Texture Features Iman. Journal of AL-Qadisiyah for computer science and mathematics, 10:42 – 53

Saberi-Movahed, F., Najafzadeh, M., Mehrpooya, A. 2020. Receiving More Accurate Predictions for Longitudinal Dispersion Coefficients in Water Pipelines: Training Group Method of Data Handling Using Extreme Learning Machine Conceptions. Water resources management. 34:529–561. https://doi.org/10.1007/s11269-019-02463-w

Safari, M. J. S., Aksoy, H., Mohammadi, M. 2015. Incipient deposition of sediment in rigid boundary open channels. Environmental Fluid Mechanics, 15(5), 1053-1068.

Safari, M. J. S., Aksoy, H., Unal, N. E., Mohammadi, M. 2017. Experimental analysis of sediment incipient motion in rigid boundary open channels. Environmental Fluid Mechanics, 17(6), 1281-1298.

Shields, A. 1936. Application of similarity principles and turbulence research to bed-load movement. Preussiischen Research Institute of Hydraulic Engineering, Berlin, Germany, Issue 26.

Salih, SQ, Sharafati, A, Khosravi, K, Faris, H, Kisi, O, Tao H, M Ali, Yaseen ZM. 2019. River suspended sediment load prediction based on river discharge information: application of newly developed data mining models. Hydrological Sciences Journal. Under press.

Schneiderm J., Moore, A. 1997. A Locally Weighted Learning Tutorial using Vizier 1.0

Sharafati, A., Khosravi, K., Khosravinia, P., Ahmed, K., Salman, S.A., Mundher, Z.,Shamsuddin, Y., 2019. The potential of novel data mining models for global solar radiation prediction. Int. J. Environ. Sci. Technol. https://doi.org/10.1007/s13762-019-023440

Shevade, S. K., Keerthi, S. S., Bhattacharyya, C., & Murthy, K. R. K. (2000). Improvements to the SMO algorithms for SVM regression. IEEE Transactions on Neural Networks, 11:5, 1188–1193.

Smola A.J. and Sch¨olkopf B. 1998a. On a kernel-based method for pattern recognition, regression, approximation and operator inversion. Algorithmica, 22: 211–231.

Straub, L. G. 1934. Effect of channel contraction works upon regimen of movable bed streams. Transactions, American Geophysical Union 2, 454–463. doi:10.1029/tr015i002p00454

Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res. Atmos. 106, 7183–7192. https://doi.org/10.1029/2000JD900719

Tin Kam Ho (1998). The Random Subspace Method for Constructing Decision Forests. IEEE Transactions on Pattern Analysis and Machine Intelligence. 20(8):832-844. URL http://citeseer.ist.psu.edu/ho98random.html.

Ting, K. M., Witten, I. H.: Stacking Bagged and Dagged Models. In: Fourteenth international Conference on Machine Learning, San Francisco, CA, 367-375, 1997.

Ting, K.M., Witten, I.H., 1997. Stacking bagged and dagged models. (Working paper 97/09). University of Waikato, Department of Computer Science, Hamilton, New Zealand.

Webby, M. G. 1984. General scour at contraction. RRU Bulletin 73, National Roads Board, Bridge Design and Research Seminar, New Zealand, 109-118.

Yang, J. F., Zhai, Y. J., Xu, D. P., & Han, P. (2007). SMO algorithm applied in time series model building and forecast. In 2007 International Conference on Machine Learning and Cybernetics (Vol. 4, pp. 2395-2400). IEEE.

Yaseen ZM, Jaafar O, Deo RC, Kisi O, Adamowski J, Quilty J, El-Shafie A. 2016.Stream-flow forecasting using extreme learning machines: a case study in a semi-arid region in Iraq. Journal of Hydrology 542, 603-614

Supplementary information material

Table A. Optimum values of each model's parameter

Models parameters	Developed algorithms						
	ISOR	SMOR	LWL	ICO	LMS	Dagging	RS
Debug	No	-	NO	NO	No	No	NO
C	-	1	-	-	-	-	-
Filter type	-	Normalize training data	-	-	-	-	-
Kernel	-	Poly-kernel	-	-	-	-	-
Reg-optimized	-	Reg-SMO Improved	-		-	-	-
KNN	-	-	-1		-	-	-
Nearest Neighbor search algorithm	-	-	Linear NN Search		-	-	-
Weighting Kernel	-	-	0			-	-
Random seed	-	-	-		0	-	-
Sample size	-	-	-		4	-	-
Classifier	-	-	-	-	-	ISOR, SMOR, LWL, ICO, LMS	ISOR, SMOR, LWL, ICO, LMS
Number of folds	-	-	-	10	-	10	-
Verbose	-	-	-		-	No	-
SubSpace size	-	-	-	-	-	-	0.5
Batch size	-	-	-	100	-	-	-
Class value index	-	-	-	-1	-	-	-
Do not check capability	-	-	-	No	-	-	-
Evaluation metric	-	-	-	*RMSE*	-	-	-
Iterative classifier	-	-	-	Additive regression	-	-	-
Look ahead iterations	-	-	-	50	-	-	-

Table B. definition of each models parameter

Parameters	definition
Debug	Is set to yes, classifier may output addition information to the console
C	The complexity parameter C
Filter type	Determines how/if the data will be transformed
KNN	How many neighbors are used to determine the width of the weighting function
Weighting Kernel	Determines weighting function
Random seed	Set the seed for selecting random subsamples of the training data
Sample size	Set the size of random samples used to generate the least squared regression functions
Number of folds	Number of fold for cross-validation
Verbose	Whether to output some additional information during building
SubSpace size	Set of each subspace: if less than 1 as a percentage of the number of attributes, otherwise the absolute number of attributes
Batch size	The preferred number of instances to process if batch prediction is being performed
Class value index	The class value index to use with information retrieval type metrics
Do not check capability	If set to yes, classifier capabilities are not checked before classifier is built
Evaluation metric	Evaluation metric to use
Iterative classifier	The Iterative classifier to be optimized
Look ahead iterations	The number of iteration to look ahead for to find a better optimum