Insights into the prediction capability of roughness coefficient in current ripple bedforms under varied hydraulic conditions

Kiyoumars Roushangar [ID][a,b,*] and Saman Shahnazi[a]

[a] Department of Water Resources Engineering, Faculty of Civil Engineering, University of Tabriz, 29 Bahman Ave, Tabriz, Iran
[b] Center of Excellence in Hydroinformatics, University of Tabriz, Tabriz, Iran
*Corresponding author. E-mail: kroshangar@yahoo.com

[ID] KR, 0000-0002-2620-5106

ABSTRACT

Ubiquitous flow bedforms such as ripples in rivers and coastal environments can affect transport conditions as they constitute the bed roughness elements. The roughness coefficient needs to be adequately quantified owing to its significant influence on the performance of hydraulic structures and river management. This work intended to evaluate the sensitivity and robustness of three machine learning (ML) methods, namely, Gaussian process regression (GPR), artificial neural network (ANN), and support vector machine (SVM) for the prediction of the Manning's roughness coefficient of channels with ripple bedforms. To this end, 840 experimental data points considering various hydraulic conditions were prepared. According to the obtained results, GPR was found to accurately predict the Manning's coefficient with input parameters of Reynolds number (Re), depth to width ratio (y/b), the ratio of the hydraulic radius to the median grain diameter (R/D_{50}), and grain Froude number ($V/\sqrt{g(s-1)D_{50}}$). Moreover, sensitivity analysis was implemented with proposed ML approaches which indicated that the ratio of the hydraulic radius to the median grain diameter has a considerable role in modeling the Manning's coefficient in channels with ripple bedforms.

Key words: artificial neural network (ANN), Gaussian process regression (GPR), Manning's coefficient, ripple bedforms, support vector machine (SVM)

HIGHLIGHTS

- GPR, SVM and ANN were selected to identify influential parameters for prediction of roughness coefficient of ripple bedforms.
- 840 experimental data points from different sources were used to feed the utilized models.
- Prediction capability of roughness coefficient was investigated under varied hydraulic conditions.

INTRODUCTION

Accurate prediction of the flow resistance (i.e., roughness coefficient) in open channel hydraulics has a significant effect on flow conditions and can be identified as a crucial part of designing and operating hydraulic structures. It is way more complicated to determine roughness since it is affected by several factors including irregular channel bed properties, bed material, vegetation, cross-sectional, plan form variability, etc.). The problem of predicting flow resistance and roughness coefficient depends, to a large extent, on the bedform. Assessments of the bedform such as dunes and ripples in rivers and marine environments require information on the instability mechanism, as the development of the bedform is inversely proportional to the lag between bed shear stress, sediment transport, and bed elevation. When the tractive force is sufficient to begin the sediment transport, an initially flat bed will be unstable and deformed into irregular features (Kennedy 1969). In the case of fine sediment, ripples are formed, while coarser sediments and higher subcritical velocity (Froude number <1) will usually form dunes (Figure 1). Ripples refer to triangular sand waves with small dimensions, typically shorter than about 0.6 meters and higher than about 60 mm, whereas dunes are associated with larger dimensions formed in natural streams (Engelund & Fredsoe 1982).

In the past half-century, many analytical and semi-empirical approaches have been presented in order to predict the total roughness coefficient owing to bedform roughness (Meyer-Peter & Müller 1948; Einstein & Barbarossa 1952; Tylor & Brooks
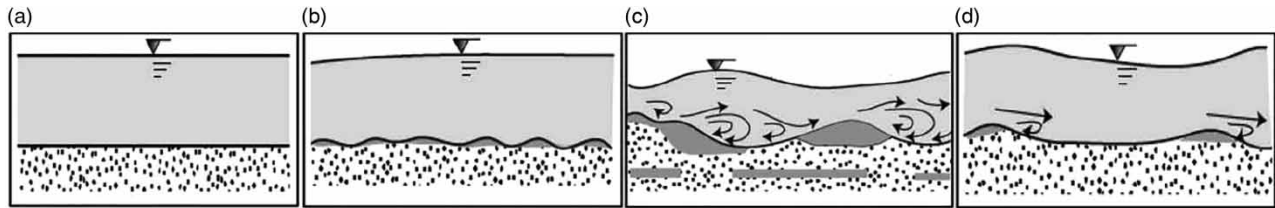
Figure 1 | Different types of bedforms in alluvial channels (lower regime): (a) plane bed, (b) ripple, (c) dune, (d) washed-out dune (Julien 2010).

1962; Raudkivi 1967; Richardson & Simons 1967; Smith 1968; Van Rijn 1984; Karim 1995; Yang *et al.* 2005; Van der Mark *et al.* 2008). There are also a large number of studies characterizing the effect of the Reynolds and Froude numbers on the roughness coefficient (Rouse *et al.* 1963; Brownlie 1983; Colosimo *et al.* 1988). Utilizing the Froude number as an independent parameter, Ugarte & Madrid (1994) prepared an expression for Manning's *n*. Afzalimehr & Anctil (1998) conducted dimensional analysis on different dimensionless parameters with the Froude number as one of them, which had a significant effect on friction factor. García Díaz (2005) suggested that the Froude number is inversely proportional to Manning's *n*. Zhang *et al.* (2010) studied a curve-fitting relationship between Manning's *n* with the Reynolds and Froude numbers. However, the existing relationships for the prediction of roughness coefficient related to bedforms differ from each other and no universal equation for roughness coefficient was established. This can be due to the complicated process of interaction between a large number of variables, 3D nature of bedform development, and also the lag in the adjustment of bedform in reaction to changing flow conditions (Karim 1999). Concerning the point emphasized in the foregoing discussion, development of a flexible and robust methodology that is capable of predicting roughness coefficient for the channel with different types of bedform is deemed a crucial problem. In the recent decade, AI (artificial intelligence) methods have been introduced as reliable tools in providing persistent success in various fields of hydraulic engineering. In AI models we are looking for a learning machine capable of finding an accurate approximation of a natural phenomenon, as well as expressing it in the form of an interpretable equation. However, this bias towards interpretability creates several new issues. The computer-generated hypotheses should take advantage of the already existing body of knowledge about the domain in question. However, the method by which we express our knowledge and make it available to a learning machine remains rather unclear (Babovic 2009). More recently, AI methods including artificial neural network (ANN), support vector machine (SVM), gene programming (GP), and group method of data handling (GMDH) have been applied for modeling flow resistance and bedform dimensions of alluvial channels. The effectiveness of GP-based approaches was revealed in the experiments of developing a formula for the description of vegetation-induced roughness (Babovic & Keijzer 2000; Keijzer & Babovic 2002; Giustolisi 2004; Baptist *et al.* 2007). Azamathulla *et al.* (2013) accurately solved a high nonlinear relationship between Manning's *n* and input parameters of the Froude and Reynolds numbers, width to depth ratio, bed slope of channel, and relative roughness through the gene expression programming (GEP) approach. Roushangar *et al.* (2017) offered a useful prediction method based on the least squares support vector machine (LSSVM) coupled with particle swarm optimization (PSO). Compared with semi-empirical equations, their hybrid model enjoyed higher performance when it came to predicting the Manning and Darcy–Weisbach roughness coefficients in open channels with dune bedforms. In another investigation into the application of AI methods on the modeling characteristics of dune bedforms, Roushangar *et al.* (2018a, 2018b) developed GEP-based equations for the prediction of the Manning's roughness coefficient and relative dune height. Javadi *et al.* (2015) found that SVM surpasses ANN in terms of predicting dune bedform dimension. Qaderi *et al.* (2017) used a combination of GMDH with shuffled complex evolution (SCE) and harmony search (HS) in simulating bedform dimensions, and concluded that the developed hybrid models outperform all other empirical approaches for predicting bedform dimensions. Roushangar *et al.* (2018a, 2018b) applied extreme learning machine (ELM) in order to find the nonlinear interaction among different input variables for the prediction of coefficient of friction of overland flows. More recently, Saghebian *et al.* (2020) presented the applicability of Gaussian process regression (GPR) for the prediction of total and bedform resistance of dune bed channels.

A great deal of previous research into the application of AI methods for predicting flow resistance of alluvial channels shows that much attention has been paid to dunes' bedform and, to the best of our knowledge, there is a lack of research on the comprehensive study of predicting the roughness coefficient in channels with ripple bedform. Therefore, the present study aims to investigating the generalization capability of SVM and GPR as effective kernel-based techniques for modeling

the total Manning's coefficient (which includes both grain friction and form resistance) in channels with ripple bedform. The proposed techniques were developed using experimental datasets. Performance of the different input combinations was evaluated in both training and testing phases and under two scenarios through several statistical measures. Moreover, in order to undertake the comprehensive assessment of the bedform's role in modeling the Manning's coefficient, prediction results of the employed methods for channels with dune bedform were discussed using four experimental datasets (one original experiment was performed by the author and three other experiments by other researchers). In the last stage, the most influential parameters in predicting the total Manning's coefficient were determined using sensitivity analysis.

MATERIALS AND METHODS

Experimental data used in the study

Due to the fact that employing more datasets from varied hydraulic conditions can challenge the ML methods and enjoy more reliable evaluation, a total of 26 related data sources with the appropriate experimental data were explored and the relevant data were considered for the modeling process. As a result, there are 840 records for open channels with ripple bedforms. The sources of data as well as the ranges of measured and calculated parameters are presented in Table 1. It should be pointed out that the measured values of the Manning's coefficient are calculated from Manning's formula which is simple in form and is well confirmed by much practical experience.

Table 1 | Hydraulic characteristics of the utilized experimental data for ripple bedforms

Source	D_{50} (mm)	y (m)	B (m)	Fr	Re	n	Number of data
Athaullah (1968)	0.018–0.047	0.08–0.32	2.43	0.03–0.36	2,012–113,789	0.015–0.026	38
Jopling & Forbes (1979)	0.045	0.02–0.1	0.2	0.22–0.67	3,793–29,494	0.014–0.027	11
Guy *et al.* (1966)	0.18–0.5	0.08–0.31	0.6–2.43	0.14–0.36	14,505–47,429	0.007–0.019	44
Mantz (1983)	0.017–0.35	0.02–0.12	0.3	0.19–0.61	5,824–39,494	0.005–0.028	26
Lau (1988)	0.08–0.4	0.05–0.14	0.75	0.17–0.45	8,956–27,639	0.013–0.058	35
Banks & Collinson (1975)	0.29	0.07–0.22	0.81	0.19–0.54	17,215–77,875	0.017–0.025	23
Costello & Southard (1981)	0.51–0.66	0.14–0.16	0.92	0.17–0.26	35,149–46,557	0.007–0.017	8
Ueno (1981)	0.23–0.53	0.02–0.11	0.4–1	0.06–0.38	2,142–16,482	0.02–0.038	14
Taylor (1972)	0.22	0.11–0.18	0.85	0.23–0.45	41,324–76,954	0.017–0.020	9
Barton & Lin (1955)	0.18	0.09–0.42	1.2	0.16–0.39	19,432–137,629	0.016–0.032	17
Brooks[a] (1957)	0.08–0.14	0.05–0.09	0.26	0.27–0.49	14,532–26,708	0.015–0.021	10
Chyn (1935)	0.59–0.84	0.04–0.07	0.61	0.49–0.76	19,433–32,686	0.011–0.013	25
Davies (1971)	0.15	0.07–0.3	1.3	0.17–0.49	16,749–110,839	0.013–0.023	27
Franco (1968)	0.23	0.12–0.16	0.91	0.28–0.40	27,676–52,607	0.017–0.022	11
Jorissen (1938)	0.6–0.91	0.02–0.10	0.61	0.48–0.67	7,262–46,380	0.010–0.017	15
Laursen (1958)	0.04–0.11	0.07–0.30	0.91	0.24–0.47	22,751–134,842	0.013–0.022	20
Mutter (1971)	0.26	0.01–0.1	1.21	0.16–1.57	8,236–22,856	0.008–0.041	23
Nomicos (1956)	0.07–0.08	0.09–0.15	0.26	0.28–0.72	11,578–33,069	0.009–0.023	15
Nordin (1976)	0.12–0.24	0.32–0.85		0.17–0.30	96,870–315,067	0.014–0.024	5
Pratt (1970)	0.47	0.07–0.45	1.37	0.11–0.30	14,502–121,741	0.015–0.028	25
Straub (1954), Straub *et al.* (1958)	0.16–0.19	0.04–0.07	0.3	0.39–0.83	17,478–38,922	0.010–0.020	10
Vanoni & Brooks (1957)	0.13	0.07–0.16	0.85	0.19–0.5	15,695–64,141	0.014–0.025	12
Vanoni & Hwang (1967)	0.20–0.23	0.07–0.37	0.26–1.1	0.22–0.50	9,306–100,502	0.015–0.024	16
Shinohara (1959)	0.21	0.01–0.04	0.34	0.33–0.9	2,446–17,750	0.014–0.042	15
U.S. Corps of Engineers (1935)	0.18–0.47	0.01–0.26	0.73	0.11–0.73	5,218–61,013	0.003–0.051	215
Singh (1960)	0.62	0.01–0.2	0.25–0.75	0.27–0.85	3,100–36,129	0.008–0.024	171

[a] Data source: Vanoni & Brooks (1957).

Feed-forward neural network (FFNN)

Artificial neural networks are a family of machine learning algorithms originally inspired by biological neural networks that can be employed to approximate any measurable function with an arbitrary number of inputs (Tayfur 2014). The feed-forward neural network (FNN) with back propagation (BP) is a widely known utilized strategy in water resources engineering issues (Karami *et al.* 2012; Li *et al.* 2018). The employed ANN is the common FFNN algorithm with three layers of input, hidden, and target (Figure 2). The Levenberg–Marquardt preparing calculation (Hagan & Menhaj 1994) was utilized, and the mean square error (MSE) between the calculated and observed values served as the cost function. Different numbers of parameters were used as input and the optimum number of neurons was obtained through trial and error process. Furthermore, in the proposed work, the tan-sigmoid was used as an activation function in the hidden and output layers.

Support vector machine (SVM)

Here, fundamental aspects of SVM are presented since this kernel-based method has been conducted in various fields of engineering. It may be presumed that for dataset$\{x_i, y_i\}$, SVM formulations function as (Vapnik 1998):

$$f(x) = W\varphi(x) + b \tag{1}$$

where W (weight factor) and b (bias) are known as the parameters of the regression function and $\varphi(x)$ stands as the transfer function.

$$R_{SVM_s}(C) = \frac{1}{2}\|w\|^2 + C\frac{1}{2}\sum_{i=1}^{n} L(x_i, y_i) \tag{2}$$

where $C\frac{1}{2}\sum_{i=1}^{n} L(x_i, y_i)$ denotes the empirical risk. The minimization process of regularized risk function is used in order to calculate the parameters W and b. This process is implemented after introducing positive slack variables ξ_i and ξ_i^* as representative of upper and lower excess deviation.

$$\text{Minimize } R_{SVM_s}(w, \xi^*, \xi) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n} (\xi_i, \xi_i^*) \tag{3}$$

$$\text{Subject to} \begin{cases} y_i - w\varphi(x_i) + b_i \leq \xi + \xi_i \\ w\varphi(x_i) + b_i - y_i \leq \xi + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, \ i = 1, \ldots, n \end{cases}$$
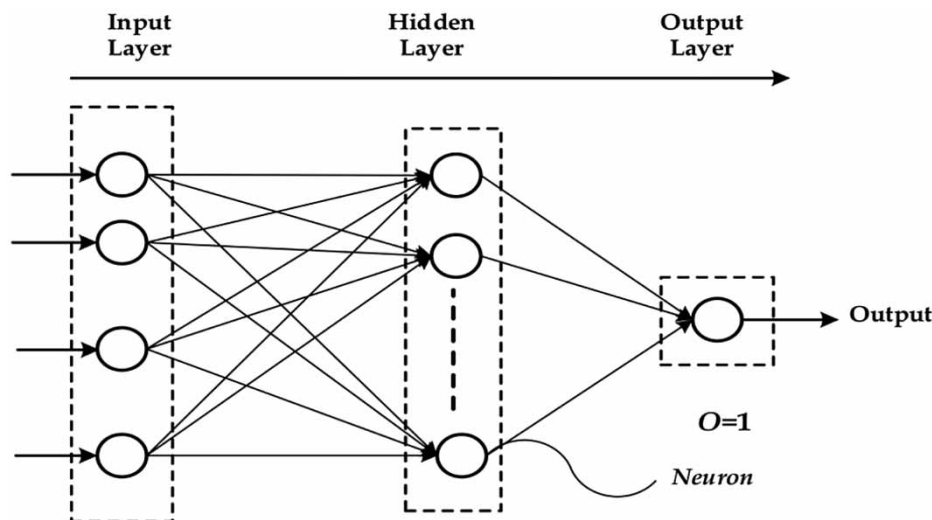

Figure 2 | Structure of employed ANN model.

where $\frac{1}{2}\|w\|^2$ stands for the regularization term, C is the cost factor, ε is known as the loss function, and n represents the sample size.

Equation (1) can be resolved through applying Lagrange multiplier and optimality constraints. Consequently, a general form of function can be obtained by:

$$f(x) = \sum_{i=1}^{n} (\beta_i - \beta_i^*) K(x_i, x_j) + b \tag{4}$$

where $K(x_i, x_j) = \varphi(x_i)\varphi(x_j)$ and the term $K(x_i, x_j)$ stand as the kernel function, which is an inner product of two vectors x_i and x_j in the feature space $\varphi(x_i)$ and $\varphi(x_j)$, respectively. The kernel function plays a significant role in the performance of SVM. Some of the kernel functions and their parameters are listed in Table 2. In this study, radial basis function (RBF) kernel function was used for implementation of the SVM algorithm as suggested by researchers (Azamathulla *et al.* 2016; Roushangar & Shahnazi 2020). The optimum values of RBF kernel parameter (γ) were obtained after a trial and error process. Furthermore, optimization of related hyper parameters (C and ε) has been carried out by a systematic grid search of the parameters using cross-validation on the training dimensionless measures.

Gaussian process regression (GPR)

Deriving from the Bayesian framework, GPR can be considered as a random process to carry out the nonparametric regression with the Gaussian process. For regression, consider that $y \subseteq Re$; then, a Gaussian process on domain of inputs is determined by a mean function $\mu : \chi \to Re$. In GPR, the response variable (y) is calculated as $y = f(x) + \xi$, where $f(x)$ is latent function and $\xi \sim N(0, \sigma^2)$ is the additive noise and is considered as normal independent and identically distributed noise contribution with a mean value of zero. The standard deviation of the noise (σ^2) and $f(x)$ drawn from the Gaussian process on χ is determined by K, as:

$$Y = (y_1, \ldots, y_n) \sim N(0, K_{ij} + \sigma^2 I)$$

where $K_{ij} = K(x_i, x_j)$, and I refers to the identity matrix. Since $Y/X \sim N(0, K + \sigma^2 I)$ is normal, it can be considered as the contingent distribution of test labels provided training and test data of $p(Y^*/Y, X, X^*)$. Therefore, one has $Y^*/Y, X, X^* \sim N(\mu, \sum)$, where:

$$\mu = K(X_*, X)(K, (X, X) + \sigma^2 I)^{-1} Y \tag{5}$$

$$\sum = K(X_*, X_*) - \sigma^2 I - K(X_*, X)(K(X, X) + \sigma^2 I)^{-1} K(X, X_*) \tag{6}$$

where $K(X, X_*)$ represents the matrix of covariance between the training set X and the test set X_*. Moreover, $K(X_*, X_*)$ represents the covariance matrix of the test set itself. Here, X and Y are the vector of the training data and training data labels y_i, whereas X_* is the vector of the test data. A particular covariance function is needed for producing a positive semi-definite covariance matrix K, where $K_{ij} = K(x_i, x_j)$. The covariance function and associated parameters with the degree of noise should be optimally determined through the training process of the GPR model. The Gaussian process allows the utilization of Bayesian inference over the noise variance σ_2 and the kernel parameters to be applied. The process begins with the calculation of the log-likelihood of the regressors y. Then, maximization of this marginal likelihood can be obtained by taking derivatives over the parameters and using gradient descent (Kuss 2006). The covariance function $k(x_i, x_j)$ can be defined by various kernel functions. It can be parameterized in terms of the kernel parameters in vector θ. Hence, it is possible to express the covariance function as $k(x_i, x_j|\theta)$.

Table 2 | Different kernel functions of SVM

Kernel type	Function	Kernel parameter
Linear	$K(x_i, x_j) = (x_i, x_j)$	–
Polynomial	$K(x_i, x_j) = ((x_i, x_j) + 1)^d$	d
RBF	$K(x_i, x_j) = \exp(- \| x_i, x_j \|^2 / 2\gamma^2)$	γ
Sigmoid	$K(x_i, x_j) = \tan h(-\alpha(x_i, x_j) + c)$	α, c

Considering $r = \sqrt{(x_i - x_j)^T(x_i - x_j)}$ and $R = \sqrt{\sum_{m=1}^{d}(x_{im} - x_{jm})^2/\sigma_m^2}$, different kernel functions can be defined as shown in Table 3. In this table, σ_l represents the length scale parameter, σ_f represents the signal standard deviation, and σ_m represents the separate length scale. In the present study, in order to tune the related hyper parameters, a standard gradient descent optimizer was utilized through maximizing the log marginal likelihood.

Performance metrics

Assessment and comparison of the performance of each proposed model developed in the present study is evaluated using statistical parameters of the Nash–Sutcliffe efficiency (NSE), correlation coefficient (R), root mean squared error (RMSE), and the logarithmic transformation variable (e), which are formulated as follows:

$$R = \frac{\sum_{i=1}^{N}(X_i - \bar{X}) \times (Y_i - \bar{Y})}{\sum_{i=1}^{N}(X_i - \bar{X})^2 \times (Y_i - \bar{Y})^2} \tag{7}$$

$$NSE = 1 - \frac{\sum_{i=1}^{N}(X_i - Y_i)^2}{\sum_{i=1}^{N}(X_i - \bar{X})^2} \tag{8}$$

$$RMSE = \sqrt{\sum_{i=1}^{N}\frac{(X_i - Y_i)^2}{N}} \tag{9}$$

$$e = \sum_{i=1}^{N}\frac{\log Y_i}{\log X_i} \tag{10}$$

Table 3 | Different kernel functions of GPR

Kernel type	Function
Squared exponential	$k(x_i,\ x_j\|\theta) = \sigma_f^2 \exp\left[-\frac{1}{2}\frac{(x_i - x_j)^T(x_i - x_j)}{\sigma_l^2}\right]$
Exponential	$k(x_i,\ x_j\|\theta) = \sigma_f^2 \exp\left(-\frac{r}{\sigma_l}\right)$
Matern 3/2	$k(x_i,\ x_j\|\theta) = \sigma_f^2\left(1 + \frac{\sqrt{3}r}{\sigma_l}\right)\exp\left(-\frac{\sqrt{3}r}{\sigma_l}\right)$
Matern 5/2	$k(x_i,\ x_j\|\theta) = \sigma_f^2\left(1 + \frac{\sqrt{5}r}{\sigma_l} + \frac{5r^2}{3\sigma_l^2}\right)\exp\left(-\frac{\sqrt{5}r}{\sigma_l}\right)$
Rational quadratic	$k(x_i,\ x_j\|\theta) = \sigma_f^2\left(1 + \frac{r^2}{2\alpha\sigma_l^2}\right)^{-\alpha}$
ARD squared exponential	$k(x_i,\ x_j\|\theta) = \sigma_f^2 \exp\left[-\frac{1}{2}\sum_{m=1}^{d}\frac{(x_{im} - x_{jm})^2}{\sigma_m^2}\right]$
ARD exponential	$k(x_i,\ x_j\|\theta) = \sigma_f^2 \exp(-R)$
ARD Matern 3/2	$k(x_i,\ x_j\|\theta) = \sigma_f^2(1 + \sqrt{3}R)\exp(-\sqrt{3}R)$
ARD Matern 5/2	$k(x_i,\ x_j\|\theta) = \sigma_f^2\left(1 + \sqrt{5}R + \frac{5}{3}R^2\right)\exp(-\sqrt{5}R)$
ARD rational quadratic	$k(x_i,\ x_j\|\theta) = \sigma_f^2\left(1 + \frac{1}{2\alpha}\sum_{m=1}^{d}\frac{(x_{im} - x_{jm})^2}{\sigma_m^2}\right)^{-\alpha}$

where N stands for the number of data, X_i is the observed value, Y_i is the predicted value, $\bar{X}$ and $\bar{Y}$ represent the mean values of the observed and predicted values. Since using non-normalized data would decrease the speed and accuracy of AI approaches and may lead to zero and minus predictions, the following equation was used to normalize input and output variables by scaling between 0.1 and 1.

$$x_{norm} = 0.1 + 0.9 \times \left(\frac{x_i - x_{imin}}{x_{imax} - x_{imin}} \right) \tag{11}$$

For the purpose of predicting Manning's coefficient, due to training and testing goals, data were divided into training set (75% of total data) and testing set (remaining 25% of data). As a result, there are 630 measurements for training and 210 measurements for testing.

RESULTS AND DISCUSSION

In order to implement the employed ANN, SVM, and GPR for prediction of Manning's roughness coefficient, the most significant step is to find the best architecture of the model, which can properly simulate the relationship between input and output variables. It is assumed that the roughness coefficient (n) as a dependent variable can be described through a function of dimensionless variables as follows (Roushangar *et al.* 2017, 2018a):

$$n = f\left(Fr, \ Re, \ \frac{y}{b}, \ \frac{R}{D_{50}}, \ \frac{V}{\sqrt{g(s-1)D_{50}}}, \ \frac{Vy}{\sqrt{g(s-1)D_{50}^3}} \right) \tag{12}$$

where Fr is the Froude number, Re is Reynold's number of water flow, y is flow depth, b is width of channel, R is hydraulic radius, D_{50} is mean grain diameter, and the last two parameters of Equation (12) are included as independent variables since they include most of the dimensional sediment and flow variables (except viscosity). To quantitatively assess the influence of each parameter, different combinations of the aforementioned parameters were considered and based on trial and error process, the models of Table 4 were suggested for modeling the Manning's roughness coefficient of alluvial channels with ripple bedforms.

The performances of the employed ANN, SVM, and GPR methods were compared with each other and the associated results appear in Table 5 representing the evaluation indices of the applied models. In the first step, it was attempted to express the Manning's coefficient modeling process through variables based on hydraulic characteristics. To achieve this goal, the employed ML approaches were fed with double input variables (models I and II). According to the results of Table 5, it can be seen that introducing the Froude number with depth to width ratio yielded better prediction accuracy. Considering the results

Table 4 | Applied input combination for prediction of Manning's n

Models	Input parameters
(I)	$Fr, \ \dfrac{y}{B}$
(II)	$Re, \ \dfrac{y}{B}$
(III)	$Fr, \ \dfrac{y}{B}, \ \dfrac{R}{D_{50}}$
(IV)	$Re, \ \dfrac{y}{B}, \ \dfrac{R}{D_{50}}$
(V)	$Re, \ \dfrac{y}{B}, \ \dfrac{R}{D_{50}}, \ \dfrac{V}{\sqrt{g(s-1)D_{50}}}$
(VI)	$Re, \ \dfrac{y}{B}, \ \dfrac{R}{D_{50}}, \ \dfrac{Vy}{\sqrt{g(s-1)D_{50}^3}}$

Table 5 | Statistical comparison of the obtained results of employed methods (the best results are italicized and in bold)

	Performance criteria for test series							
	Train				Test			
Models	R	NSE	RMSE	e	R	NSE	RMSE	e
ANN (I)	0.742	0.551	0.067	3.63	0.747	0.548	0.073	−0.541
ANN (II)	0.525	0.275	0.085	6.87	0.521	0.259	0.093	2.98
ANN (III)	0.778	0.605	0.063	3.90	0.730	0.481	0.078	−0.113
ANN (IV)	0.789	0.623	0.061	2.84	0.793	0.617	0.067	0.249
ANN (V)	*0.852*	*0.726*	*0.052*	*2.43*	*0.854*	*0.719*	*0.057*	*0.586*
ANN (VI)	0.817	0.667	0.057	2.19	0.819	0.650	0.064	1.51
SVM (I)	0.661	0.432	0.075	−0.276	0.646	0.360	0.087	−2.52
SVM (II)	0.637	0.404	0.077	3.81	0.614	0.375	0.086	0.935
SVM (III)	0.822	0.675	0.057	1.36	0.801	0.636	0.065	−0.868
SVM (IV)	0.882	0.766	0.048	5.73	0.838	0.687	0.060	0.764
SVM (V)	*0.888*	*0.788*	*0.046*	*0.367*	*0.880*	*0.775*	*0.051*	*0.687*
SVM (VI)	0.922	0.850	0.038	1.49	0.877	0.767	0.052	−0.097
GPR (I)	0.661	0.432	0.075	−0.276	0.646	0.360	0.087	−2.52
GPR (II)	0.637	0.404	0.077	3.81	0.614	0.375	0.086	0.935
GPR (III)	0.915	0.807	0.044	4.12	0.826	0.669	0.062	0.262
GPR (IV)	0.988	0.969	0.017	1.84	0.853	0.714	0.058	−0.567
GPR (V)	*0.982*	*0.956*	*0.021*	*2.05*	*0.880*	*0.769*	*0.052*	*0.286*
GPR (VI)	0.993	0.983	0.013	1.39	0.863	0.733	0.056	−0.321

of the modeling process with hydraulic characteristics, ANN confirms its superiority over kernel-based approaches with respect to statistical indices (R = 0.747, NSE = 0.548, RMSE = 0.073, and e = −0.541) for the testing part.

It is apparent from the results that the model (V) with four input parameters including Reynolds number (Re), the ratio of depth to width of channel (y/B), the ratio of the hydraulic radius to median grain diameter (R/D_{50}) and grain Froude number ($V/\sqrt{g(s-1)D_{50}}$) has the best performance for prediction of roughness coefficient in alluvial channels with ripple bedforms. The obtained results indicate that the implementation of model (V) as input combination of the SVM method provided very good outcomes (R = 0.880, NSE = 0.775, RMSE = 0.051, and e = 0.687), superior to the other machine learning methods employed, while ANN generated poor results. According to NSE values, when comparing model (I) and model (II), considering Re (in model II) increases the model accuracy by approximately 28% (ANN), 8% (SVM), and 6.7% (GPR). It can indicate the merits of each input combination and sensitivity of employed kernel-based approaches to input parameters. It can be seen that omitting $V/\sqrt{g(s-1)D_{50}}$ and introducing relative discharge ($Vy/\sqrt{g(s-1)D_{50}^3}$) in model (VI) reduces the performance of ANN by approximately 9%. On the other hand, considering the obtained results of GPR and SVM approaches, models (V) and (VI) show similar potential, indicating more flexibility and generalization capability of kernel-based approaches in quantification of Manning's roughness coefficient. Figure 3 shows the variation of statistical parameters of NSE via the numbers of neurons in the hidden layer (fed with model (V)). Taking into account the NSE values, significant variability can be seen throughout the performance of the employed ANN approach. ANN performance ranges between NSE = 0.330 (for 1 neuron) and NSE = 0.719 (for 21 neurons). It is observed that the best network structure is 4-21-1. Furthermore, it can be seen that for solving the objective problem with employed datasets, the ANN model with numbers of neurons in the hidden layer less than 21 leads to over-fitting.

Based on the value of the logarithmic transformation variable, GPR presents relatively better prediction accuracy (e = 286) in comparison to SVM (e = 0.687). With the aid of standard gradient descent optimizer, the best values of related parameters for the different kernels were achieved as the length scale parameter (σ_l) ranging from 0.0240 to 0.1320 and the signal standard deviation (σ_f) ranging from 0.0664 to 0.1149.
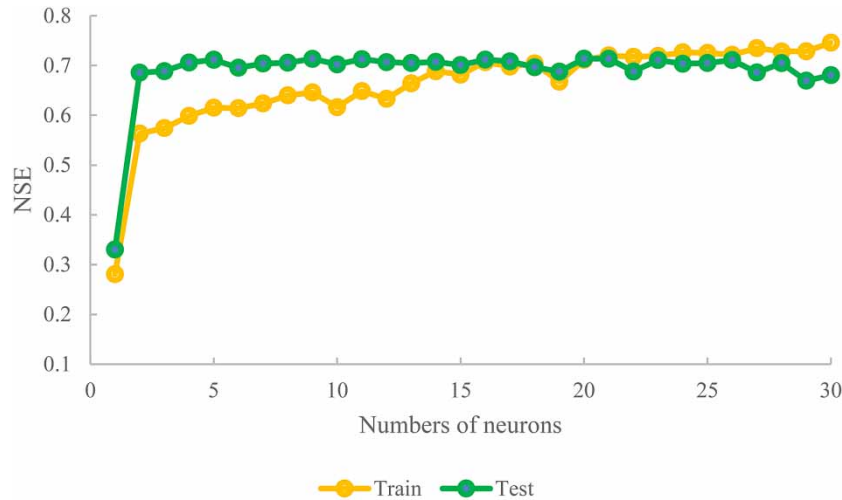
Figure 3 | Variation of NSE vs. numbers of neurons for model (V).

Figure 4 depicts the scatter plots between the observed data and model results. The results of previous works (Roushangar *et al.* 2017; 2018a; Roushangar and Shahnazi 2020 and Saghebian *et al.* 2020) show that the best results for dune-bed channel studies for prediction of Manning's roughness coefficient were obtained from GEP (R = 0.866, NSE = 0.742, and RMSE = 0.0035), least squares SVM (R = 0.839, NSE = 0.705, and RMSE = 0.0036), and GPR (R = 0.784 and NSE = 0.715), respectively. Hence, it seems that hydraulic conditions governing ripple bedforms provide better predictive ability for machine learning approaches in comparison to channels with dune bedforms.

In order to assess the prediction capability of the Manning's coefficient under varied hydraulic conditions, different intervals of the Reynolds number were considered based on trial and error. Then, the best input combination (Re, y/b, R/D_{50}, and $V/\sqrt{g(s-1)D_{50}}$) was rerun for selected data categories. Results of the testing parts are plotted in Figure 5, which shows a clear ascending trend in performance of kernel-based approaches. These findings indicate that the performance of SVM and GPR approaches with the selected best input combination tends to be more robust with increasing Reynolds number. Prediction process for Reynolds number greater than 26,000 (which includes 339 data points) gave the most accurate results with SVM (R = 0.941, NSE = 0.884, RMSE = 0.036, and e = 0.326) and GPR (R = 0.945, NSE = 0.890, RMSE = 0.035, and e = 0.928). According to the obtained results, the ANN method showed less stability and presented a poor performance in predicting the Manning's coefficient in different intervals of the Reynolds number. In addition, for a Reynolds number greater than 11,000 (which includes 88% of employed data), SVM demonstrated satisfactory outcomes with R = 0.873, NSE = 0.761, RMSE = 0.048, and e = 0.618.

Since the Froude number is an effective parameter for illustrating the hydraulic properties of rivers, it may be beneficial to check out the effectiveness of the best input combination for the prediction of Manning's coefficient in different intervals of this parameter. In accordance with the obtained results, prediction of Manning's coefficient with lower Froude number ($Fr <$ 0.25) increases the modeling performance of GPR and ANN by 4.2% and 5.5%, respectively. On the contrary, hydraulic conditions governing the flow with the higher Froude number ($Fr > 0.55$) decreased the modeling accuracy with R = 0.539, NSE = 0.267, and RMSE = 0.033 for SVM and R = 0.532, NSE = 0.279, and RMSE = 0.032 for GPR. Generally, it can be inferred that Manning's coefficient in channels with ripple bedforms has better prediction capability in flows with lower Froude number. Results of predicted Manning's coefficient in different intervals of the Froude number are presented in Figure 6.

The functionality of the employed ML approaches was investigated in different ranges of the R/D_{50} as the most influential parameter in determination of Manning's coefficient of ripple bedform channels. As shown in Figure 7, variation of the R/D_{50} from 50 to 400 caused fluctuation in performance of ANN from NSE = 0.444 to NSE = 0.756, SVM from NSE = 0.629 to NSE = 0.806, and GPR from NSE = 0.692 to NSE = 0.807. Results revealed that the GPR model offered a more consistent performance with the variation of R/D_{50} values. Moreover, as can be seen in Figure 7, when the depth is larger compared with bed material size so that $R/D_{50} > 450$, the performance of the employed ML approaches decreases significantly.
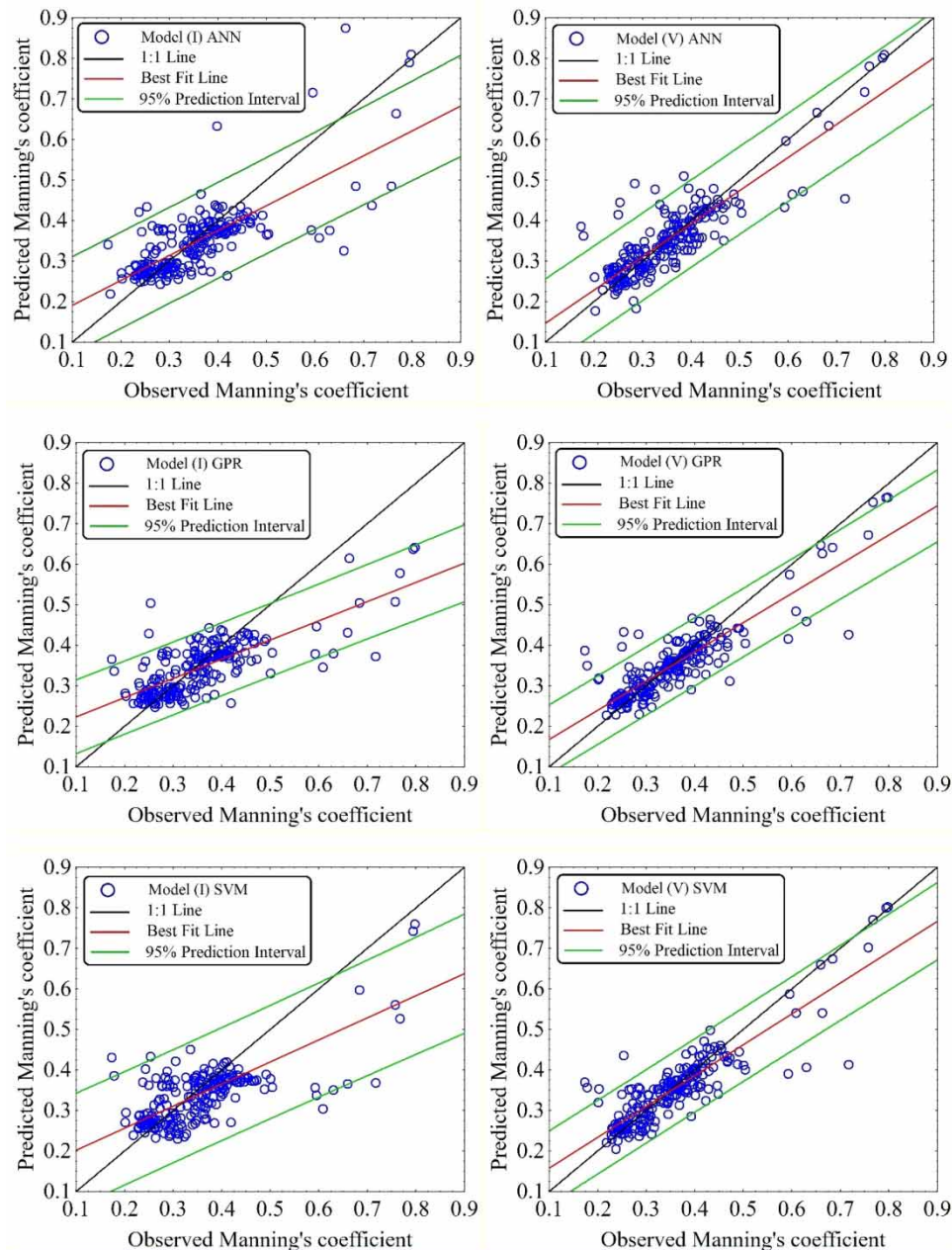
Figure 4 | Scatter plots of observed and predicted values of *n* for the test sets.

In SVM, model behavior is largely dependent on the RBF kernel parameter (γ), which can lead to under-fitting and over-fitting in the prediction process (Roushangar & Shahnazi 2019). Figure 8 illustrates the statistical indices via gamma values of the SVM model (fed with model (V)). From the figure, it can be seen that the best fitting gamma values are obtained when $\gamma \geq 50$. When gamma is small (in our example $\gamma \leq 50$), the SVM model tends to memorize all the training data but is capable of generalizing unseen data: hence, only trained data points can be predicted. Thus, for solving the objective problem with the employed datasets, SVM model with gamma values less than 50 leads to over-fitting. According to the output, the optimum value of RBF kernel function was assessed as 300 for model (V). It is worth noting that statistical indices can show different behavior with variations of gamma value for different input combinations.

Since a well-advised application of kernel-based modeling methods is to find an appropriate kernel function and tune associated hyper parameters, various kernel functions were used as a core tool of the employed GPR methods. Table 6
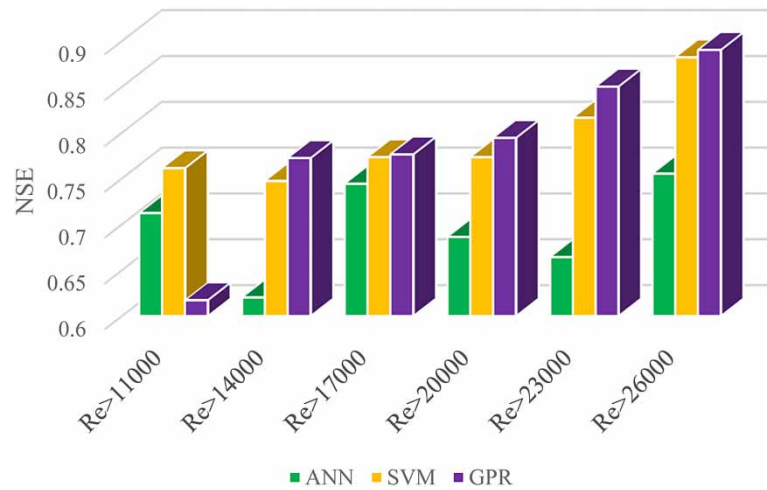
Figure 5 | Results of predicted Manning's coefficient in different intervals of Reynolds number.
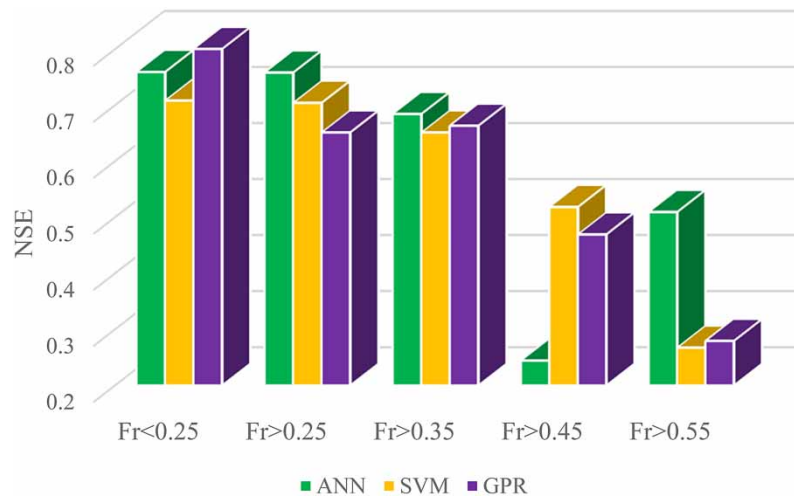

Figure 6 | Results of predicted Manning's coefficient in different intervals of the Froude number.
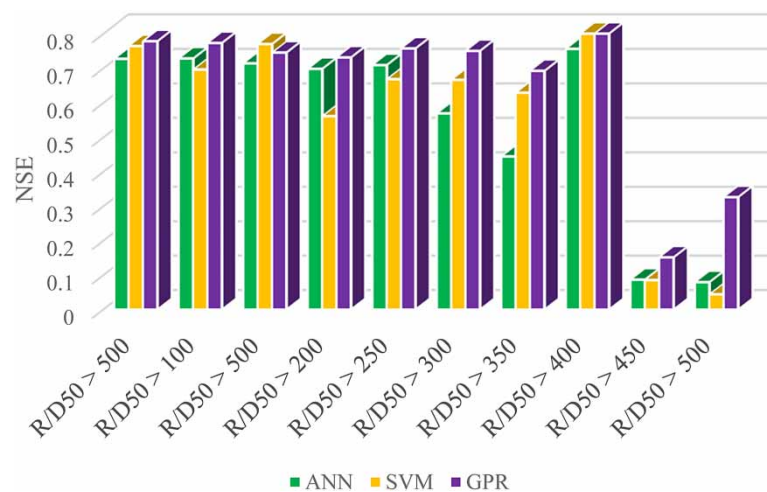

Figure 7 | Results of predicted Manning's coefficient in different intervals of the relative roughness.
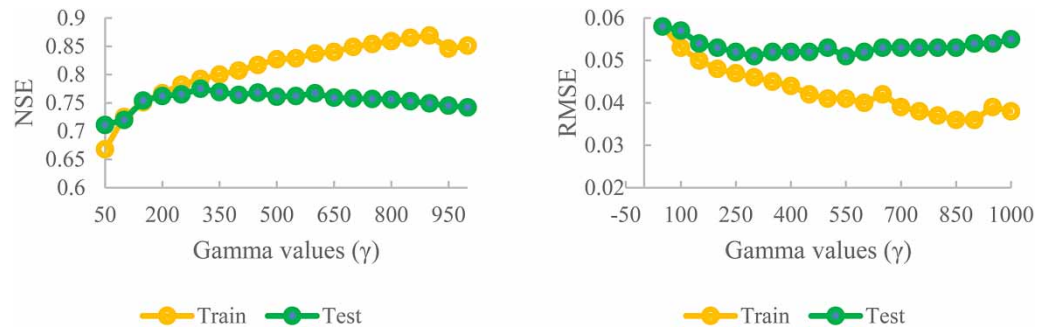
Figure 8 | Variation of NSE and RMSE vs. gamma values for model (V).

lists the results of statistical indices of different kernels for model (III) as the best input combination for prediction of the Manning's roughness coefficient.

Attending to the NSE values, the trivial variability can be seen throughout the considered kernel functions. GPR performance ranges between 0.722 (for rational quadratic kernel) and 0.798 (for ARD Matern 3/2 kernel). In addition to exponential kernel, the constructed GPR structure with squared exponential kernel seems to be robust since their performance in the test set is encouraging. The comparison demonstrated that the Matern 3/2 Matern 5/2 kernel functions performed equally well. Moreover, based on the results of Table 6, utilization of automatic relevance determination (ARD) kernels improved the effect of a good nonlinear interpolation of the employed GPR method and increased the global average accuracy by approximately 3% in terms of NSE values. As mentioned before, the result proved that SVM–RBF had the most success rate values (R = 0.880, NSE = 0.775, and RMSE = 0.051). It was followed by polynomial kernel function (R = 0.513, NSE = 0.108, and RMSE = 0.102) and linear kernel (R = 0.271, NSE = 0.038, and RMSE = 0.106). The results showed that the SVM with sigmoid kernel function achieved the worst performance for both training and testing phases (R = 0.116, NSE = −0.00056, and RMSE = 0.149).

In the last step, a simple sensitivity analysis of the input parameters on the prediction accuracy of employed ML approaches for the Manning's roughness coefficient of channel with ripple bedforms is presented. The selected input combination for the analysis is model (III), as its prediction accuracy was proved to be the best in the previous sections. In addition, the ANN model with Levenberg–Marquardt learning algorithm, SVM with RBF kernel, and GPR with exponential kernel have also been selected for the sensitivity analysis of inputs. The sensitivity analysis was implemented by successively omitting each input from the model (V). Consequently, the statistical behavior of the eliminated input is reduced in terms of employed criteria such as R and NSE, allowing the prediction models to quantify the effect of excluded input on the prediction targets. Figure 9 shows the results of the sensitivity analysis. In this figure ΔNSE stands for the values of percent reduction of NSE pertaining to each excluded parameter.

Table 6 | The statistical indices of GPR method with different kernel functions (model (V))

| Kernel types	Performance criteria for test series			
	R	NSE	RMSE	e
Exponential	0.880	0.769	0.052	0.286
Squared exponential	0.887	0.776	0.051	0.896
Rational quadratic	0.851	0.722	0.057	−0.123
Matern 3/2	0.888	0.782	0.050	0.704
Matern 5/2	0.888	0.781	0.050	0.796
ARD exponential	0.893	0.795	0.049	0.189
ARD squared exponential	0.872	0.759	0.053	0.774
ARD rational quadratic	0.870	0.754	0.054	−0.397
ARD Matern 3/2	*0.895*	*0.798*	*0.048*	*0.695*
ARD Matern 5/2	0.889	0.788	0.050	0.834
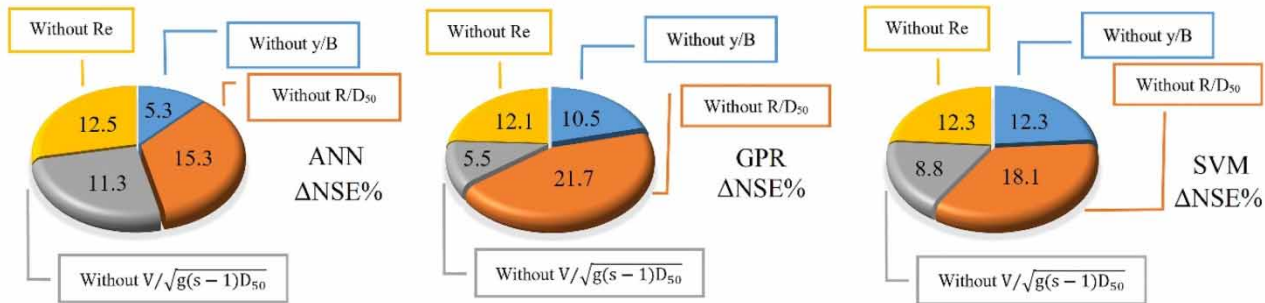
Figure 9 | The results of sensitivity analysis.

In view of the analysis results, depicted in Figure 9, it can be deduced that R/D_{50} is the most sensitive input parameter and plays a significant role in the prediction process. This conclusion is reached using ANN, GPR, and SVM methods with the highest ΔNSE values (15.3%, 21.7%, and 18.1%, respectively). The Reynolds number is the second most effective parameter on the Manning's coefficient of ripple bedforms. With its elimination ΔNSE increased to 12.5%. In the case of SVM, given the ΔNSE measure, the two parameters of y/B and Re had a similar effect on prediction accuracy, while the parameter y/B has the least effect on prediction accuracy of ANN. It can be observed that different ML approaches use different degrees of input parameter features for appropriate modeling of the relationship between input and output parameters. It can be seen that, in the case of the ANN method, the parameter y/b had the least impact on the modeling process, but, on the other hand, this parameter was the second most effective parameter on the modeling accuracy of the employed SVM method.

CONCLUSIONS

Although there have been some useful steps towards the application of ML approaches in determining the characteristics of dune bedforms, the effectiveness of the aforementioned approaches for modeling flow resistance in alluvial channels with ripple bedforms remains elusive. In this regard, three ML approaches, namely, GPR, ANN, and SVM, were employed in this study for the Manning's coefficient prediction. An extensive dataset consisting of 840 experimental samples from 24 sources was used in order to interpret their embodied knowledge by employed techniques. Six input combinations were tested, based on flow and sediment characteristics. All the variables were used in non-dimensional form in order to ensure dimensional consistency between inputs and outputs. Regarding the results, the model (V) with parameters Re, y/b, R/D_{50}, and $V/\sqrt{g(s-1)D_{50}}$ was the most accurate model. SVM yielded the best results in all the employed statistical indices (R = 0.880, NSE = 0.775, RMSE = 0.051, and e = 0.687). Moreover, the proper utilization of ML techniques in different intervals of the Reynolds number and relative roughness dictates more stability and generalization capability of GPR in predicting the Manning's coefficient under varied hydraulic conditions. It was deduced that the prediction of the Manning's coefficient for Reynolds values of greater than 26,000 was more precise than those with lower Reynolds values. In addition, the obtained results demonstrated that complicated hydraulic conditions governing channel systems caused a considerable decrease in the performance of the employed ML approaches. However, it should be noted that the employed GPR, SVM, and ANN are data-driven models and the ML-based models are data sensitive, so further studies using data ranges beyond this study and field data should be conducted in order to prove the merits of the proposed models to estimate roughness coefficient of ripple bedforms in real flow conditions.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

REFERENCES

Afzalimehr, H. & Anctil, F. 1998 Estimation of gravel-bed river flow resistance. *Journal of Hydraulic Engineering* **124** (10), 1054–1058.
Athaullah, M. 1968 *Prediction of Bed Forms in Erodible Channels*. PhD thesis, Colorado State University, Fort Collins, CO, USA.
Azamathulla, H. M., Ahmad, Z. & Ghani, A. A. 2013 An expert system for predicting Manning's roughness coefficient in open channels by using gene expression programming. *Neural Computing and Applications* **23** (5), 1343–1349.

Azamathulla, H. M., Haghiabi, A. H. & Parsaie, A. 2016 Prediction of side weir discharge coefficient by support vector machine technique. *Water Science and Technology: Water Supply* **16** (4), 1002–1016.

Babovic, V. 2009 Introducing knowledge into learning based on genetic programming. *Journal of Hydroinformatics* **11** (3–4), 181–193.

Babovic, V. & Keijzer, M. 2000 Genetic programming as a model induction engine. *Journal of Hydroinformatics* **2** (1), 35–60.

Banks, N. L. & Collinson, J. D. 1975 The size and shape of small-scale current ripples: an experimental study using medium sand. *Sedimentology* **22** (4), 583–599.

Baptist, M. J., Babovic, V., Rodríguez Uthurburu, J., Keijzer, M., Uittenbogaard, R. E., Mynett, A. & Verwey, A. 2007 On inducing equations for vegetation resistance. *Journal of Hydraulic Research* **45** (4), 435–450.

Barton, J. R. & Lin, P. N. 1955 *Study of the Sediment Transport in Alluvial Channels*. Colorado A and M College, Fort Collins, CO, USA.

Brownlie, W. R. 1983 Flow depth in sand-bed channels. *Journal of Hydraulic Engineering* **109** (7), 959–990.

Chyn, S. D. 1935 *An Experimental Study of the Sand Transporting Capacity of Flowing Water on Sandy Bed and the Effect of the Composition of the Sand*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, USA.

Colosimo, C., Copertino, V. A. & Veltri, M. 1988 Friction factor evaluation in gravel-bed rivers. *Journal of Hydraulic Engineering* **114** (8), 861–876.

Costello, W. R. & Southard, J. B. 1981 Flume experiments on lower-flow-regime bed forms in coarse sand. *Journal of Sedimentary Research* **51** (3), 849–864.

Davies, T. R. 1971 *Summary of Experimental Data for Flume Tests Over Fine Sand*. Report CE/3/71, p. 2. Department of Civil Engineering, University of Southampton, Southampton, UK.

Einstein, H. A. & Barbarossa, N. L. 1952 River channel roughness. *Transactions of the ASCE* **117**, 1121–1132.

Engelund, F. & Fredsoe, J. 1982 Sediment ripples and dunes. *Annual Review of Fluid Mechanics* **14** (1), 13–37.

Franco, J. J. 1968 Effects of water temperature on bed-load movement. *Journal of Waterways and Harbors Division* **94** (3), 343–352.

García Díaz, R. 2005 Analysis of Manning coefficient for small-depth flows on vegetated beds. *Hydrological Processes* **19** (16), 3221–3233.

Giustolisi, O. 2004 Using genetic programming to determine Chezy resistance coefficient in corrugated channels. *Journal of Hydroinformatics* **6** (3), 157–173.

Guy, H. P., Simons, D. B. & Richardson, E. V. 1966 *Summary of Alluvial Channel Data From Flume Experiments*. US Geological Survey Professional Paper 462-I, 11–96. US Government Printing Office, Washington DC, USA.

Hagan, M. T. & Menhaj, M. B. 1994 Training feedforward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks* **5** (6), 989–993.

Javadi, F., Ahmadi, M. M. & Qaderi, K. 2015 Estimation of river bedform dimension using artificial neural network (ANN) and support vector machine (SVM). *Journal of Agricultural Science and Technology* **17** (4), 859–868.

Jopling, A. V. & Forbes, D. L. 1979 Flume study of silt transportation and deposition. *Geografiska Annaler: Series A, Physical Geography* **61** (1–2), 67–85.

Jorissen, A. L. 1938 Étude expérimentale du transport solide des cours d'eau. *Rev. Univ. Mines* **14**, 269–282.

Julien, P. Y. 2010 *Erosion and Sedimentation*. Cambridge University Press, New York, USA, p. 181.

Karami, H., Ardeshir, A., Saneie, M. & Salamatian, S. A. 2012 Prediction of time variation of scour depth around spur dikes using neural networks. *Journal of Hydroinformatics* **14** (1), 180–191.

Karim, F. 1995 Bed configuration and hydraulic resistance in alluvial-channel flows. *Journal of Hydraulic Engineering* **121** (1), 15–25.

Karim, F. 1999 Bed-form geometry in sand-bed flows. *Journal of Hydraulic Engineering* **125** (12), 1253–1261.

Keijzer, M. & Babovic, V. 2002 Declarative and preferential bias in GP-based scientific discovery. *Genetic Programming and Evolvable Machines* **3** (1), 41–79.

Kennedy, J. F. 1969 The formation of sediment ripples, dunes, and antidunes. *Annual Review of Fluid Mechanics* **1** (1), 147–168.

Kuss, M. 2006 *Gaussian Process Models for Robust Regression, Classification, and Reinforcement Learning*. PhD thesis, Technischen Universität, Darmstadt, Germany.

Lau, Y. L. 1988 Hydraulic resistance of ripples. *Journal of Hydraulic Engineering* **114** (10), 1275–1282.

Laursen, E. M. 1958 The total sediment load of streams. *Journal of the Hydraulics Division* **84** (1), 1–36.

Li, X., Sha, J., Li, Y. M. & Wang, Z. L. 2018 Comparison of hybrid models for daily streamflow prediction in a forested basin. *Journal of Hydroinformatics* **20** (1), 191–205.

Mantz, P. A. 1983 Semi-empirical correlations for fine and coarse cohesionless sediment transport. *Proceedings of the Institution of Civil Engineers* **75** (1), 1–33.

Meyer-Peter, E. & Müller, R. 1948 Formulas for bed-load transport. In: *Proceedings of 2nd Meeting of the International Association Hydraulics Research*, Delft, The Netherlands, pp. 39–64.

Mutter, D. G. 1971 *A Flume Study of Alluvial Bed Configurations*. PhD thesis, Department of Civil Engineering, University of Alberta, Edmonton, Canada.

Nomicos, G. N. 1956 Effects of sediment load on the velocity field and friction factor of turbulent flow in an open channel. Thesis, California Institute of Technology, Pasadena.

Nordin, C. F. 1976 *Flume Studies with Fine and Coarse Sands (No. 76-762)*. US Geological Survey, Reston, VA, USA.

Pratt, C. J. 1970 *Summary of Experimental Data for Flume Tests Over 0.49 mm Sand*. Department of Civil Engineering, University of Southampton, Southampton, UK.

Qaderi, K., Maddahi, M. R., Rahimpour, M. & Masoumi Shahr-babak, M. 2017 Investigating the capability of two hybrid intelligence methods to predict bedform dimensions of alluvial channels. *Water Science and Technology: Water Supply* **18** (5), 1706–1718.

Raudkivi, A. J. 1967 Analysis of resistance in fluvial channels. *Journal of the Hydraulics Division* **93** (HY5), 2084–2093.

Richardson, E. V. & Simons, D. B. 1967 Resistence to flow in sand channels. In: *Paper Presented at 12th Congress of the International Association of Hydraulic Research*, Fort Collins, CO, USA.

Rouse, H., Koloseus, H. J. & Davidian, J. 1963 The role of the Froude number in open-channel resistance. *Journal of Hydraulic Research* **1** (1), 14–19.

Roushangar, K. & Shahnazi, S. 2019 Bed load prediction in gravel-bed rivers using wavelet kernel extreme learning machine and meta-heuristic methods. *International Journal of Environmental Science and Technology* **16** (12), 8197–8208.

Roushangar, K. & Shahnazi, S. 2020 Determination of influential parameters for prediction of total sediment loads in mountain rivers using kernel-based approaches. *Journal of Mountain Science* **17** (2), 480–491.

Roushangar, K., Saghebian, S. M. & Mouaze, D. 2017 Predicting characteristics of dune bedforms using PSO-LSSVM. *International Journal of Sediment Research* **32** (4), 515–526.

Roushangar, K., Alami, M. T. & Saghebian, S. M. 2018a Modeling open channel flow resistance with dune bedform via heuristic and nonlinear approaches. *Journal of Hydroinformatics* **20** (2), 356–375.

Roushangar, K., Alipour, S. M. & Mouaze, D. 2018b Linear and non-linear approaches to predict the Darcy-Weisbach friction factor of overland flow using the extreme learning machine approach. *International Journal of Sediment Research* **33** (4), 415–432.

Saghebian, S. M., Roushangar, K., Ozgur Kirca, V. S. & Ghasempour, R. 2020 Modeling total resistance and form resistance of movable bed channels via experimental data and a kernel-based approach. *Journal of Hydroinformatics* **22** (3), 528–540.

Shinohara, K. 1959 *On the Characteristics of Sand Waves Formed upon the Beds ofn Open Channels*, Vol. 7(25). Research Institute of Applied Mechanics, Kyushu University, Japan, pp. 15–45.

Singh, B. 1960 *Transport of Bed-Load in Channels with Special Reference to Gradient and Form*. PhD thesis, Imperial College London, London, UK.

Smith, K. V. 1968 Alluvial channel resistance related to bed form. *Journal of the Hydraulics Division* **94** (1), 59–70.

Straub, L. G. 1954 *Terminal Report on Transportation Characteristics – Missouri River Sediment*. MRD Sediment Series No. 4. University of Minnesota, Minneapolis.

Straub, L. G., Anderson, A. G. & Flammer, G. H. 1958 *Experiments on the Influence of Temperature on the Sediment Load (No. 10)*. US Army Engineer Division, Missouri River.

Tayfur, G. 2014 *Soft Computing in Water Resources Engineering: Artificial Neural Networks, Fuzzy Logic and Genetic Algorithms*. WIT Press, Boston, MA, USA.

Taylor, B. D. 1972 *Temperature Effects in Alluvial Streams*. PhD thesis, California Institute of Technology, Pasadena, CA, USA.

Taylor Jr, R. H. & Brooks, N. H. 1962 Discussion of resistance to flow in alluvial channels. *Transactions of the American Society of Civil Engineers* **127**, 982–992.

Ueno, T. 1981 On the regions of occurrence of ripples and dunes. In: *Proceedings of the 25th Japanese Conference on Hydraulics*. JSCE, pp. 93–98.

Ugarte, A. & Madrid, M. 1994 Roughness coefficient in mountain rivers. In: *Hydraulic Engineering* (Cotroneo, G. & Rumer, R. eds). American Society of Engineering, New York, USA, pp. 652–656.

United States Army Corps of Engineers 1935 *Studies of River Bed Materials and Their Movement with Special Reference to the Lower Mississippi River*. Waterways Experiment Station, Vicksburg, MS, USA. Paper 17, p. 161.

Van der Mark, C. F., Blom, A. & Hulscher, S. J. M. H. 2008 Quantification of variability in bedform geometry. *Journal of Geophysical Research: Earth Surface* **113** (F3), F03020.

Vanoni, V. A. & Brooks, N. H. 1957 *Laboratory Studies of the Roughness and Suspended Load of Alluvial Streams*. Sedimentation laboratory, California Institute of Technology, Pasadena.

Vanoni, V. A. & Hwang, L. S. 1967 Bed forms and friction in streams. *Journal of the Hydraulics Division* **93** (HY3), 121–144.

van Rijn, L. C. 1984 Sediment transport, part III: bed forms and alluvial roughness. *Journal of Hydraulic Engineering* **110** (12), 1733–1754.

Vapnik, V. 1998 *Statistical Learning Theory*. Wiley, New York, USA.

Yang, S. Q., Tan, S. K. & Lim, S. Y. 2005 Flow resistance and bed form geometry in a wide alluvial channel. *Water Resources Research* **41**, 9.

Zhang, G. H., Luo, R. T., Cao, Y., Shen, R. C. & Zhang, X. C. 2010 Impacts of sediment load on Manning coefficient in supercritical shallow flow on steep slopes. *Hydrological Processes* **24** (26), 3909–3914.