Modeling total resistance and form resistance of movable bed channels via experimental data and a kernel-based approach

Seyed Mahdi Saghebian, Kiyoumars Roushangar, V. S. Ozgur Kirca and Roghayeh Ghasempour

ABSTRACT

An accurate prediction of roughness coefficient in alluvial channels is of substantial importance for river management. In this study, the total and form resistance in alluvial channels with dune bedform were assessed using experimental data. First, the data of experiments carried out at the Hydraulic Laboratory of University of Tabriz was used to investigate the impact of hydraulic and sediment parameters on roughness coefficient. Then, these data were combined with other laboratory data, and the total and bedform resistance were modeled via a Gaussian Process Regression (GPR) approach. For models, developing different input combinations were considered based on flow and sediment characteristics. The obtained results from the experiments showed that the Reynolds number has a better correlation with flow resistance in comparison with other hydraulic parameters. It was found that the roughness variations due to bedform are almost between 40 and 80% of the total roughness coefficient. Also, the obtained results proved the capability of the GPR method in the modeling process. It was found that the model which took the advantages of both flow and sediment characteristics performed better compared to the other models. The sensitivity analysis results showed that the Reynolds number has the most significant impact in the prediction process.

Key words | bedform, experimental data, GPR, roughness coefficient

Seyed Mahdi Saghebian (corresponding author)
Department of Civil Engineering, Ahar Branch,
Islamic Azad University,
Ahar,
Iran
E-mail: *smsaghebian@iau-ahar.ac.ir*

Kiyoumars Roushangar
Roghayeh Ghasempour
Department of Water Resource Engineering,
Faculty of Civil Engineering,
University of Tabriz,
Tabriz,
Iran

Kiyoumars Roushangar
Center of Excellence in Hydroinformatics,
University of Tabriz,
Tabriz,
Iran

V. S. Ozgur Kirca
Division of Hydraulics, Department of Civil Engineering,
Technical University,
Istanbul,
Turkey

INTRODUCTION

Assessment of flow resistance is not a trivial matter, due to the multitude of factors influencing roughness (e.g. bed material, bedforms, cross-sectional and plan form variability, vegetation etc.). Flow resistance in alluvial channels can be due to two roughnesses: (1) the grain (or friction) roughness, which in turn depends on size of the bed grain, and (2) the shape roughness, which depends on the shape and dimensions of the bedform as well as the depth of flow (Rouse 1965; Morvan *et al.* 2008). According to Kazemipour & Apelt (1983) and Talebbeydokhti *et al.* (2006), almost 90% of the total base flow resistance may be caused by form resistance; therefore, the form roughness should not be overlooked.

doi: 10.2166/hydro.2020.094

Total roughness coefficient can be developed in the form of a linear separation concept. The linear separation of the Manning roughness coefficient is expressed in two parts: (i) grain resistance (skin roughness), (ii) form resistance (shape roughness). Thus, the total bed roughness coefficient can be expressed as: $n = f\,(n', n'')$, where n' is the skin resistance and n'' is the form resistance that is due to bedform drag (form drag) or roughness bedform.

Also, the following equations are used for calculating the roughness coefficient in channels with bedforms (Meyer-Peter & Müller 1948; Einstein 1950). The approach follows the momentum concept assuming that the bed shear τ_0 can also be separated linearly as:

$$\tau_0 = \tau_0' + \tau_0'' \tag{1}$$

Here, τ_0 is the total shear stress, τ_0' is the shear stress due to the grain, and τ_0'' is the shear stress due to the bedform. The Shields parameter (non-dimensional shear stress) can also be decomposed into two accordingly:

$$\theta = \frac{\tau_0}{\gamma D_{50}(s-1)} = \theta' + \theta'' \tag{2}$$

where θ is the total dimensionless shear stress, θ' is the dimensionless shear stress due to the grain, θ'' is the dimensionless shear stress due to the bedform, s is the specific gravity, D_{50} is the median grain size of the bed material and γ is the specific weight. According to Engelund & Hansen (1967) θ'can be calculated as follows:

$$\theta' = \frac{D'S_f}{D_{50}(s-1)} \qquad \frac{V}{\sqrt{gD's_f}} = 6 + 2.5\ln\left(\frac{D'}{k_s}\right) \tag{3}$$

Here, D' is thickness of boundary layer over the bedform, S_f is the energy gradient, k_s is the Nikuradse equivalent sand roughness, g is gravitational acceleration and V is flow velocity. In order to decompose the Manning roughness coefficients into the grain (n') and bedform (n'') parts, the following procedure is followed. From conservation of momentum in a steady uniform open channel flow, the bed shear stress (total resistance) can be expressed as:

$$\tau_0 = \gamma\, RS_f \tag{4}$$

where R is the hydraulic radius. The Manning equation written, again, for a steady uniform open channel flow case is:

$$V = \frac{1}{n}R^{2/3}\,S_f^{1/2} \tag{5}$$

Now, combining Equations (1), (4) and (5), one can decompose the Manning roughness coefficient as:

$$\tau_0' + \tau_0'' = \gamma R(S_f' + S_f'') = \frac{V^2\gamma}{R^{1/3}}(n'^2 + n''^2) \tag{6}$$

The complexities and uncertainties of bedform configurations in alluvial channels continue to pose challenges for engineers, which stem from a variety of bedform shapes (i.e. plane bed, ripples, dunes, and antidunes) that arise under different flow conditions. The development and vanishing of the bedforms change the flow velocity distribution and flow resistance (Chien & Wan 1999). A variety of analytical and semi-empirical approaches have been developed to predict the roughness coefficient in alluvial channels. Gilbert (1914) performed some of the earliest experiments on alluvial channels, and found that the resistance coefficient varied with bedforms. Kennedy (1963) investigated the mechanics of formation of bedform, and the effect of bedform on the hydraulic resistance. Engel & Lau (1980) found that dune length to depth ratio (or dune steepness) has a significant effect on the friction factor when the dunes are steep. Karim (1999) proposed a new method for predicting relative dune height in a sand-bed stream based on the concept of relating energy loss due to form drag to the head loss across a sudden expansion in open channel flows. Heydari et al. (2014) proved that by increasing the Shields number, the ratio of Manning's roughness coefficient related to dune bedforms and the total Manning's roughness coefficient increased with a logarithmic trend. However, the existing equations rely on a limited database, untested model assumptions, and a general lack of field data, and they do not show the same results under variable flow conditions. These issues cause uncertainty in the prediction of flow resistance phenomenon; therefore, it is critical to utilize methods which are capable of predicting roughness coefficient within the channels with dune bedforms under varied hydraulic conditions.

In recent years artificial intelligence approaches (e.g. Artificial Neural Networks (ANNs), Neuro-Fuzzy models (NF), Genetic Programming (GP), Gene Expression Programming (GEP), and Support Vector Machine (SVM), Gaussian Process Regression (GPR)) have been used for the assessment of the accuracy of complex hydraulic and hydrologic phenomena, such as prediction of groundwater levels (Amaranto et al. 2018), estimation of hydraulic jump energy dissipation in channels with rough elements (Roushangar & Ghasempour 2018), prediction of flow resistance in alluvial channels (Roushangar et al. 2018), prediction of pile group scour in waves (Ghazanfari-Hashemi et al. 2011), computing longitudinal dispersion coefficients in natural streams (Azamathulla & Wu 2011), real-time hydrologic forecasting (Yu et al. 2004), side weir discharge coefficient (Azamathulla et al. 2017), and prediction of non-cohesive sediment transport in circular channels (Roushangar & Ghasempour 2017). Machine learning, a branch of artificial intelligence, deals with the representation and generalization of physical phenomena using a data learning technique. Representation of data instances and functions evaluated on these instances are part of all machine learning systems. Generalization is the property that the system will perform well on unseen data instances; the conditions under which this can be guaranteed are a key object of study in the subfield of computational learning theory. There is a wide variety of machine learning tasks and successful applications. In general, the task of an ML algorithm can be described as follows: given a set of input variables and the associated output variable(s), the objective is learning a functional relationship for the input–output variables set.

In this study, first, the impact of hydraulic and sediment parameters on total and form roughness coefficient was assessed by using the experimental data carried out at the Hydraulic Laboratory of University of Tabriz. Then, these data were combined with several available data sets in the literature, and the capability of the GPR as a kernel-based approach was investigated for modeling roughness coefficient in channels with dune bedforms. The models were defined considering various input combinations alternatives, specifically based on hydraulic characteristics and sediment properties, in order to evaluate the most appropriate input combination for roughness coefficients modeling. Finally, a sensitivity analysis was performed to find the most significant parameters in the modeling processes.

MATERIALS AND METHODS

Gaussian process regression (GPR) as a kernel-based approach

Kernel-based approaches, such as GPR, are a relatively new and important method based on the different kernel types, which are based on statistical learning theory. Such models are capable of adapting themselves to predict any variable of interest via sufficient inputs. The training of these methods is fast and has high accuracy. GPRs can model non-linear decision boundaries, and there are many kernels to choose from. They are also fairly robust against overfitting, especially in high-dimensional space. However, the appropriate selection of kernel type is the most important step in the GPR due to its direct impact on the training and classification precision. In fact, these methods are memory intensive, trickier to tune due to the importance of picking the right kernel, and do not scale well to larger datasets. In these models the proper behavior of the system can be predicted, although its intrinsic structure and behavior cannot be characterized.

GPR models are based on the assumption that adjacent observations should convey information about each other. Gaussian processes are a way of specifying *a priori* directly over function space. This is a natural generalization of the Gaussian distribution, whose mean and covariance are a vector and matrix, respectively. The Gaussian distribution is over vectors, whereas the Gaussian process is over functions. Thus, due to prior knowledge about the data and functional dependencies, no validation process is required for generalization, and GP regression models are able to understand the predictive distribution corresponding to the test input (Rasmussen & William 2006). A GP is defined as a collection of random variables, any finite number of which has a joint multivariate Gaussian distribution. Considered input space $\chi = \Re^n$ of n-dimensional vectors to an output space $\gamma = \Re$ of real-valued targets, in which n pairs (x_i, y_i) are drawn independently and identically distributed. For regression, assume that $y \subseteq \Re$; then, a GP on γ is defined by a mean function $\mu : \chi \to \Re$ and a covariance function $k : \chi \times \chi \to \Re$.

In GP regression the main assumption is that y values can be calculated from $y = f(x) + \xi$, where $\xi \sim N(0, \sigma^2)$. In GP regression, for every input x there is an associated random variable $f(x)$, which is the value of the stochastic function f at that location. In this work, it is assumed that the observational error ξ is normal independent and identically distributed, with a mean value of zero ($\mu(x) = 0$), a variance of σ^2 and $f(x)$ drawn from the Gaussian process on χ specified by k. That is, $Y = (y_1, \ldots\ldots, y_n) \sim N(0, K + \sigma^2 I)$ where $K_{ij} = k(x_i, x_j)$, and I is the identity matrix. Because

$Y/X \sim \mathrm{N}(0, K + \sigma^2 I)$ is normal, so is the conditional distribution of test labels given the training and test data of $p(Y*/Y, X, X*)$. Then, one has $Y*/Y, X, X* \tilde{} \mathrm{N}(\mu, \Sigma)$, where:

$$\mu = K(X*, X)(K(X, X) + \sigma^2 I)^{-1} Y \tag{7}$$

$$\Sigma = K(X*, X*) - \sigma^2 I - K(X*, X)(K(X, X) + \sigma^2 I))^{-1} K(X, X*) \tag{8}$$

If there are N training data and $N*$ test data, then $K(X, X*)$ represents the $N \times N*$ matrix of covariances evaluated at all pairs of training and test data sets, and this is similarly true for the other values of $K(X, X)$, $K(X*, X)$ and $K(X*, X*)$; 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 specified covariance function is required to generate a positive semi-definite covariance matrix K, where $K_{ij} = k(x_i, x_j)$. The term of the kernel function used in Support Vector Machine (SVM) is equivalent to the covariance function used in GP regression. With the known values of kernel k and degree of noise σ^2, Equations (7) and (8) would be enough for inference. During the training process of GP regression models, one needs to choose a suitable covariance function as well as its parameters. In the case of GP regression with a fixed value of Gaussian noise, a GP model can be trained by applying Bayesian inference, i.e. maximizing the marginal likelihood. This leads to the minimization of the negative log-posterior:

$$p(\sigma^2, k) = \frac{1}{2} \mathrm{y}^T (K + \sigma^2 I)^{-1} \mathrm{y} + \frac{1}{2} \log|\mathrm{K} + \sigma^2 I| - \log p(\sigma^2) - \log p(k) \tag{9}$$

To find the hyperparameters, the partial derivative of Equation (8) can be obtained with respect to σ^2 and k, and minimization can be achieved by gradient descent. For more details about GP regression and different covariance functions, readers are referred to Kuss (2006). The optimal value of capacity constant (C), the size of error-intensive zone (ε), and kernel parameter (γ) in SVM and Gaussian noise in GPR are required due to their high impact on the accuracy of mentioned regression approaches. In this study, optimization of these parameters has been performed by a systematic grid search of the parameters using cross-validation on the training set. In this grid search a normal range of parameters settings are investigated. First, optimized values of C and ε for a specified γ were obtained and then γ was changed. Statistical parameters (R, DC, and MAPE) were used to find optimums. The values of C, ε and kernel parameter (γ) which lead to the highest R and DC and lowest MAPE, were selected as optimum amounts.

Data collection

Experimental setup

In order to study the variation of roughness coefficients in open channels with dune bedform, several dune bedform experiments were performed in a 10 m long, 0.5 and 1 m wide, and 0.8 m high rectangular flume at the hydraulic laboratory of Tabriz University (Saghebian 2018). The flume had glass walls and a metal floor. Sediment particles used in the experiments were sand with specific gravity of 2.65 and uniform average diameters of 0.15 and 0.27 mm. Water flow was supplied by a pump, re-circulating between the upstream and downstream. In these experiments, discharge rate was controlled by a valve in the discharge pipe of the pump, and sediment was re-circulated together with water. The original flume had ratchet screw jacks for adjusting the slope of the flume. In this research, the flume slope was variable from 0 to 0.5%. To measure the water depth, a point gauge was used with accuracy of 0.1 mm. The point gauge was able to move along the length and width of the channel and measure the bedform height and water depth in the entire channel. By changing the flow depth and discharge, the average velocities, Froude numbers, dune height, and other parameters were calculated. In this study, at first several experiments were performed in the state of channel without bedform. Then, the friction coefficients of the channel's walls and bed were determined. To determine the form resistance, the effective roughness coefficient was extracted using the composite channel roughness equations. Finally, the walls roughness coefficients were subtracted from effective roughness coefficient and form resistance was obtained. Figure 1 shows a view from the channel, and pebbles

Figure 1 | A view of the channel.

used in downstream of the channel for reducing the turbulence of the flow.

Together with these experiments, the data sets of laboratory experiments of roughness coefficient in open channels with dune bedforms carried out by Guy *et al.* (1966), Williams (1970) and Roushangar (2010) were used. The ranges of various parameters used in the experiments are listed in Table 1. The used variables in this table are: channel width (b), mean grain diameter (D_{50}), Flow depth (y), Froude number ($Fr = V/[g \times y]^{1/2}$) in which V is flow velocity and g is gravitational acceleration, and Reynolds number ($Re = VR/v$) in which R is hydraulic radius and v is kinematic viscosity. Williams (1970) organized several experiments that were made in channels with different widths and water depths in laboratories in Washington, DC. Sediment transport rates, grain size, water depth, and channel width were measured. Furthermore, water discharge, mean velocity, slope (energy gradient), and bedform characteristics were considered as the dependent variables. Guy *et al.* (1966) studied the effects of the bed material size, flow temperature, and the fine sediment within the flow on the hydraulic and transport variables at Colorado State University. The investigations for each set covered flow phenomena ranging from a plane bed with no sediment movement to violent anti-dunes. Roushangar (2010) organized several dune bedforms experiments that were made in a 5 m long, 0.5 m wide, and 0.25 m high rectangular flume in the hydraulic laboratory of Caen University. Natural quartz sand was used as sediment particles in the experiments.

Performance criteria

Evaluating the performance of a model is commonly carried out using different statistical indexes. In this study the performance of the GPR models was evaluated using three statistical indexes: Determination Coefficient (DC), Correlation Coefficient (CC), and Mean Absolute Percentage

Table 1 | Details of the utilized experimental data

	Parameters					
Researcher	b (mm)	D_{50} (mm)	Fr	Re	y (mm)	No. of data
Williams (1970)	76.2–1,118	1.35	0.34–0.84	11,932–101,920	87.1–222	89
Guy *et al.* (1966)	609–2,438	0.19–0.93	0.25–0.65	46,800–255,500	91.4–405	114
Roushangar (2010)	150	0.15–0.4	0.21–0.40	24,192–45,869	71–145	54
Saghebian (2018)	500, 1,000	0.15, 0.27	0.19–0.49	23,561–47,238	190–370	65

Error (MAPE). The DC measures the fraction of the variance in the data explained by the model. The R provides information for the linear dependence between measured and predicted values, and the MAPE indicates the absolute differences between estimated and observed values as depicted in Equation (10):

$$CC = \frac{\sum_{i=1}^{N}(l_o - \overline{l_o}) \times (l_p - \overline{l_p})}{\sqrt{\sum_{i=1}^{N}(l_o - \overline{l_o})^2 \times (l_p - \overline{l_p})^2}},$$
$$DC = 1 - \frac{\sum_{i=1}^{N}(l_o - l_p)^2}{\sum_{i=1}^{N}(l_o - \overline{l_p})^2}, \tag{10}$$
$$MAPE = \frac{1}{\overline{l_o}} \frac{\sum_{i=1}^{n}|l_o - l_p|}{N} \times 100$$

where l_o, l_p, $\overline{l_o}$, $\overline{l_p}$, N respectively represent: the measured values, predicted values, mean measured values, mean predicted values and number of data samples.

SIMULATION AND MODEL DEVELOPMENT

Input variables

The appropriate selection of input parameters is an important step in the modeling process. The total and bedform resistance can be expressed as a function of different sets of input variables. In this study, the total and bedform Manning roughness coefficients (n and n'') are used to quantify the resistance, and these two coefficients are expressed through a set of dimensionless variables. To investigate the impacts of different parameters on roughness coefficient, two states were considered for developing the models. In the first state, parameters of the flow condition were selected as the model input: n and $n'' = f\ (Re,\ Fr,\ y/b)$. In the second state, flow, bedform, as well as sediment properties were considered as input combinations:

$$n \text{ and } n'' = f(Re, Fr, R/D_{50}, y/L, y/H, Vy/[g \times (s-1)D_{50}^3]^{0.5}, V/[g(s-1)D_{50}]^{0.5})$$

where R is the hydraulic radius, L and H are bedform length and depth, respectively, and $Vy/[g \times (s-1)\ D_{50}^3]^{0.5}$ and $V/[g(s-1)\ D_{50}]^{0.5}$ are the relative discharge and modified (densiometric) Froude number, respectively. Table 2 shows the developed GPR models in the study along with the input parameters used for each model. It should be noted that 75% of data were used for training and 25% of data were used for validating or testing the models. The order of the data sets was selected in a way such that the training data set contains a representative sample of all the behavior in the data in order to obtain a model with higher accuracy. One method for finding a good training set, which can give good accuracy both in training and testing sets, is an instance exchange which starts with a random selected training set (Bolat & Yildirim 2004).

Table 2 | GPR developed models

Hydraulic properties		Hydraulic and sediment properties (HS) and hydraulic and bedform geometry (HB)			
Model	**Inputs**	**Model**	**Inputs**		
H1	Re	HS1	R/D_{50}	HS6	$V/[g(s-1)D_{50}]^{0.5}, R/D_{50}$
H2	Fr	HS2	$V/[g(s-1)D_{50}]^{0.5}$	HS7	$V/[g(s-1)D_{50}]^{0.5}, Re$
H3	y/b	HS3	$Vy/[g(s-1)D_{50}^3]^{0.5}$	HS8	$Vy/[g(s-1)D_{50}^3]^{0.5}, Re$
H4	$Re, y/b$	HS4	$Re, R/D_{50}$	HS9	$Vy/[g(s-1)D_{50}^3]^{0.5}, R/D_{50}$
H5	$Fr, y/b$	HS5	$Fr, R/D_{50}$	HS10	$Re, Vy/[g(s-1)D_{50}{}^3]^{0.5}, R/D_{50}$
		HB1	$Re, y/L$		
		HB2	$Re, y/H$		
		HB3	$Re, y/L, L/H$		

RESULTS AND DISCUSSION

Results of the experimental study of total and bedform friction factor

In this section, the variations of the total and bedform Manning roughness coefficients with hydraulic parameters only (*Fr* and *Re*), and both hydraulic and sediment parameters (θ, $Vy/[g(s\text{–}1)\ D_{50}^3]^{0.5}$) were investigated through the experimental data only. The results are shown in Figure 2. It should be noted that for calculating the grain (skin) roughness and bedform roughness coefficients the shear stresses of τ' and τ'' were used, which corresponds to the skin and bedform resistances, respectively. At first, skin roughness (n') was calculated for the plane bed based on calculating θ' and τ'. Then, θ'' was calculated using Equation (1). Finally, the following equation was used for calculating the n'' parameter:

$$\tau'' = \frac{\theta''}{\rho\left(\frac{\rho_s}{\rho}-1\right)gD_{50}} \qquad \tau'' = \rho g R S_f'' \qquad n'' = \frac{R^{\frac{2}{3}} S_f''^{1/2}}{V} \tag{11}$$

It should be noted that at first *D*' in Equation (3) was solved by trial and error, followed by the determination of θ' given in this equation. Then θ'' could be found from Equation (1), since it is straightforward to find θ. Finally, Equation (11) is used to solve the n'' value.

According to Figure 2, it can be seen that the correlation of the n with the Reynolds number is slightly better than the Froude number. For n'', there is no desired correlation between this parameter and *Re* and *Fr* and, with increasing the Froude number which leads to dune bedform variation, the form friction coefficient decreased. It seems that the total and bedform roughness coefficients could not only depend on hydraulic parameters. According to Figure 2, the Shields parameter (θ) shows the best correlation with total and form resistance among the studied parameters, and with increasing the Shields parameter values, the values of n and n'' increase (due to the direct relationship between the energy line slope and the flow resistance coefficient). Also, it could be seen that n and n'' did not show a significant correlation with $Vy/[g(s\text{–}1)\ D_{50}^3]^{0.5}$. This issue indicates that there is an uncertainty in using the hydraulic and sediment parameters as the only input variables in the roughness coefficient estimating process.

To investigate the impact of the form resistance on the total roughness coefficient, the variations of the skin resistance, form resistance and total resistance versus hydraulic and sediment parameters are presented in Figure 3.

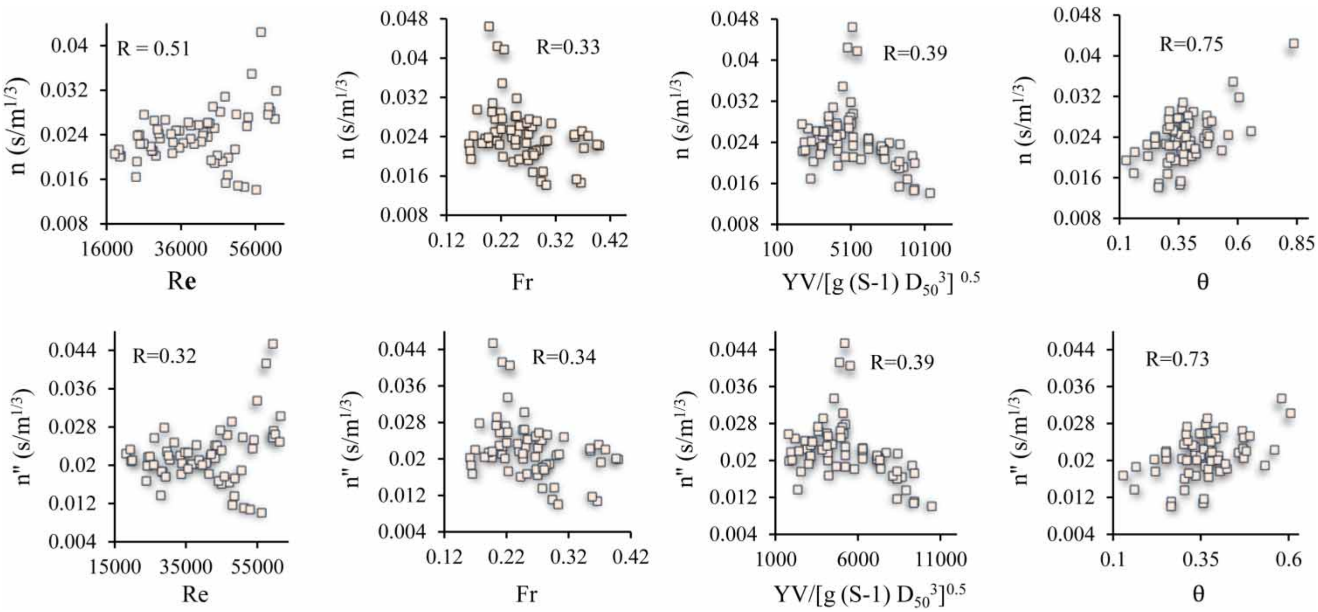

Figure 2 | The variation of the *n* and *n″* versus hydraulic and sediment parameters.

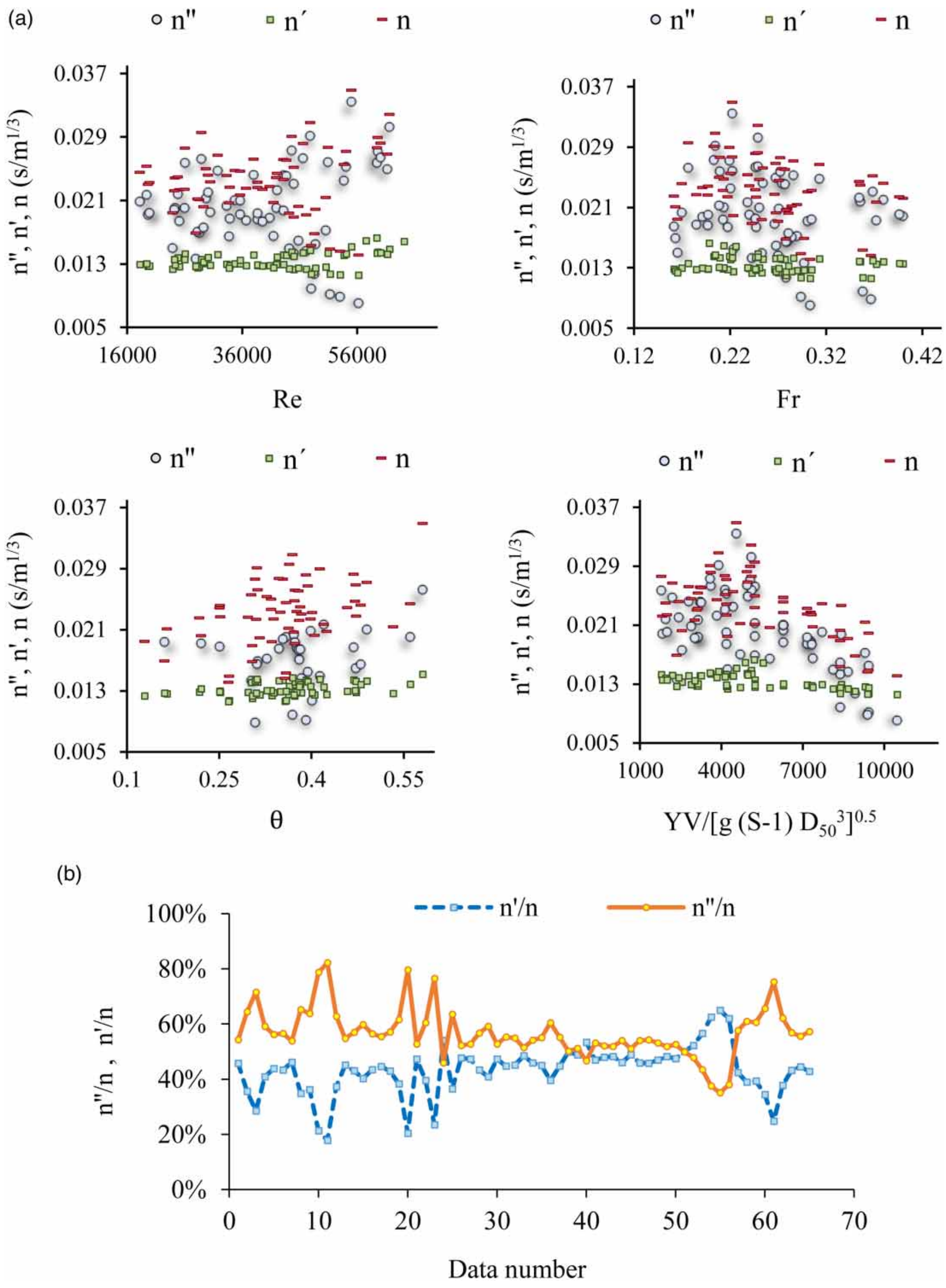

Figure 3 | (a) The variation of the n and n', n'' versus hydraulic and sediment parameters, and (b): the percentage of n' and n'' parameters variations.

According to Figure 3(a), it can be seen that the roughness variations due to the form resistance are almost in agreement with the variations of the total Manning roughness coefficient, and it can be concluded that the performance of the n'' is similar to n. From the results, it was found that the skin resistance variations are in a small range, meaning that bedform resistance dominates the flow. This issue also could be deduced from Figure 3(b). According to this figure, the roughness variations due to the bedform are almost between 40–80% of the total roughness coefficient, while the n' variation percentage is between 20 and 60%.

Modeling based on GPR approach

GPR models development

In the design of the GPR model, it is necessary to select the appropriate type of kernel function. A number of kernels are discussed in the literature. In this study, for determining the best performance of GPR and selecting the best kernel function, the model SH10 was predicted using various kernels. Figure 4 indicates the results of statistical parameters of different kernels for this model. According to the results, using the kernel function of Pearson led to better prediction accuracy. Pearson kernels were used as a core tool of GPR which was applied for the rest of the models.

The impact of hydraulic characteristics on total and bedform roughness coefficients

For evaluating the impact of flow features on total and bedform roughness coefficient, several models using only hydraulic characteristics as input variables were developed. Then, the developed models were trained and tested using the GPR. The results of the developed models are listed in Table 3 and shown in Figure 5. According to the obtained

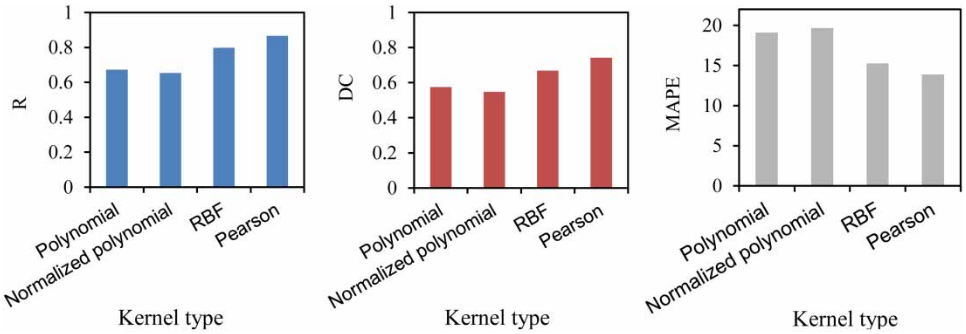

Figure 4 | Statistics parameters via GPR kernel function types for testing set of model SH10.

Table 3 | Performance evaluation of models with flow characteristics

Model	Output	Performance criteria: Train CC	DC	MAPE	Test CC	DC	MAPE
H1	n''	0.683	0.508	18.560	0.664	0.513	20.173
	n	0.877	0.675	17.405	0.827	0.546	18.679
H2	n''	0.368	0.131	27.142	0.315	0.118	30.910
	n	0.576	0.198	24.088	0.555	0.132	25.843
H3	n''	0.611	0.355	22.025	0.573	0.323	25.788
	n	0.817	0.427	19.933	0.793	0.325	22.442
H4	n''	0.738	0.631	16.185	0.716	0.614	18.561
	n	0.869	0.757	14.445	0.858	0.713	15.504
H5	n''	0.658	0.475	23.122	0.589	0.423	24.391
	n	0.797	0.533	18.463	0.672	0.466	19.806
$Re < 80{,}000$	n''	0.745	0.634	15.08	0.721	0.618	19.23
	n	0.971	0.758	12.08	0.867	0.716	14.21
$Re > 80{,}000$	n''	0.504	0.311	25.18	0.412	0.208	33.84
	n	0.533	0.319	23.05	0.446	0.244	29.58

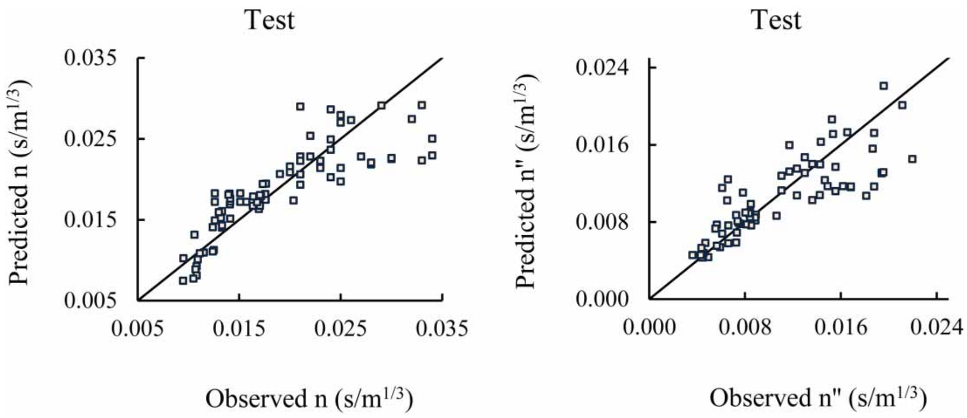

Figure 5 | Scatter plots of predicted total and dune bedform roughness coefficients vs. observed values; based on flow characteristics.

results, in both n and n' modeling, the model H4 with input parameters of Re and y/b shows better performance than the others from the RMSE, CC, and DC viewpoints (i.e. highest CC and DC and lowest RMSE). It seems that using parameter y/b as the input parameter caused an increase in model efficiency. Also, it could be seen that using Reynolds number as the only input parameter yielded to the desired accuracy compared to the case where Froude number was used as the only input. Figure 5 represents the scatter plot of the observed and predicted total and bedform roughness coefficients for the best input combination. Additionally, based on trial and error, data sets were divided into two groups in terms of the Reynolds number ($Re < 80{,}000$ and $Re > 80{,}000$). Then, the best input combination (Re, y/b) was rerun for both data categories. The results are represented in Table 3 (last two rows). These findings demonstrate that when the Reynolds numbers is less than 80,000, the model tends to be more accurate in predicting the Manning's coefficient. In higher ranges of the Reynolds number ($Re > 80{,}000$) the model accuracy decreased.

The impact of hydraulic, bedform geometry, and sediment characteristics on total and bedform roughness coefficients

Flow and sediment characteristics were employed in the establishment of input combination to evaluate the significance of the bedform and sediment features for the modeling of the total and bedform roughness coefficients. The results in Table 4 reveal that the model HS10 including Re, R/D_{50} and $Vy/[g \times (s{-}1)D_{50}^3]^{0.5}$ as inputs parameters yielded better predictions. According to the results, it could be inferred that the ratio of hydraulic radius to the sediment diameter and Reynolds number led to the desired results in predicting n and n'' roughness coefficients. Therefore, both hydraulic and sediment characteristics have a significant effect on the prediction process. It could be seen that adding parameters $Vy/[g \times (s{-}1)D_{50}^3]^{0.5}$ and R/D_{50} to Re caused an increase in models efficiency. Based on the results, it was found that the developed models with flow and bedform characteristics performed relatively weakly in comparison with developed models with only flow characteristics. Comparison between Tables 3 and 4 indicates that the applied methods for developed models based on flow and sediment properties yielded better predictions compared to using only flow parameters. Figure 6 presents the scatter plot of the observed and n and n'' roughness coefficients for the best input combination.

Sensitivity analysis

To investigate the impacts of different parameters of the GPR-best models on total and bedform roughness coefficients, a sensitivity analysis was performed. In order to evaluate the effect of each independent parameter, the model was run with all input parameters and then, one of the input parameters was eliminated and the GPR model was re-run. The MAPE performance criteria was used as indication of the significance of each parameter. The obtained results are shown in Figure 7. Based on obtained results and according to Roushangar et al. (2018), it can be

Table 4 | Performance evaluation of models with flow and sediment characteristics

Model	Output	Performance criteria: Train CC	Train DC	Train MAPE	Test CC	Test DC	Test MAPE
HS1	n''	0.551	0.523	24.441	0.533	0.507	29.066
	n	0.795	0.582	22.423	0.782	0.562	26.913
HS2	n''	0.438	0.412	22.131	0.424	0.399	33.386
	n	0.852	0.495	20.492	0.796	0.423	30.913
HS3	n''	0.466	0.447	25.237	0.451	0.433	29.251
	n	0.767	0.583	23.153	0.766	0.561	27.084
HS4	n''	0.518	0.481	22.016	0.501	0.471	26.266
	n	0.853	0.662	20.198	0.829	0.574	24.320
HS5	n''	0.630	0.597	26.327	0.609	0.573	30.361
	n	0.703	0.538	24.153	0.674	0.447	28.112
HS6	n''	0.712	0.558	23.277	0.704	0.546	27.912
	n	0.810	0.635	21.355	0.799	0.565	25.844
HS7	n''	0.757	0.589	20.212	0.552	0.521	23.414
	n	0.827	0.709	18.543	0.812	0.657	21.680
HS8	n''	0.670	0.636	19.261	0.655	0.629	16.998
	n	0.816	0.743	17.671	0.754	0.668	15.739
HS9	n''	0.751	0.712	18.224	0.741	0.681	19.369
	n	0.866	0.766	16.719	0.821	0.658	17.934
HS10	n''	0.852	0.728	14.318	0.784	0.715	15.360
	n	0.881	0.817	12.671	0.864	0.765	13.593
HB1	n''	0.803	0.515	24.871	0.765	0.493	26.43
	n	0.848	0.545	19.770	0.808	0.522	21.01
HB2	n''	0.776	0.504	26.461	0.739	0.482	28.12
	n	0.838	0.534	21.690	0.798	0.511	23.05
HB3	n''	0.804	0.531	22.706	0.766	0.508	24.13
	n	0.839	0.552	19.639	0.799	0.528	20.87

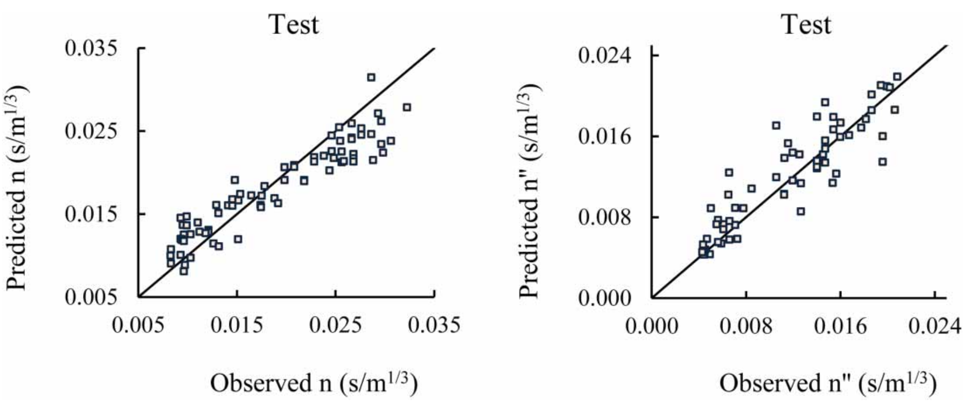

Figure 6 | Scatter plots of predicted dune bedform roughness coefficient vs. observed values; based on flow and sediment characteristics.

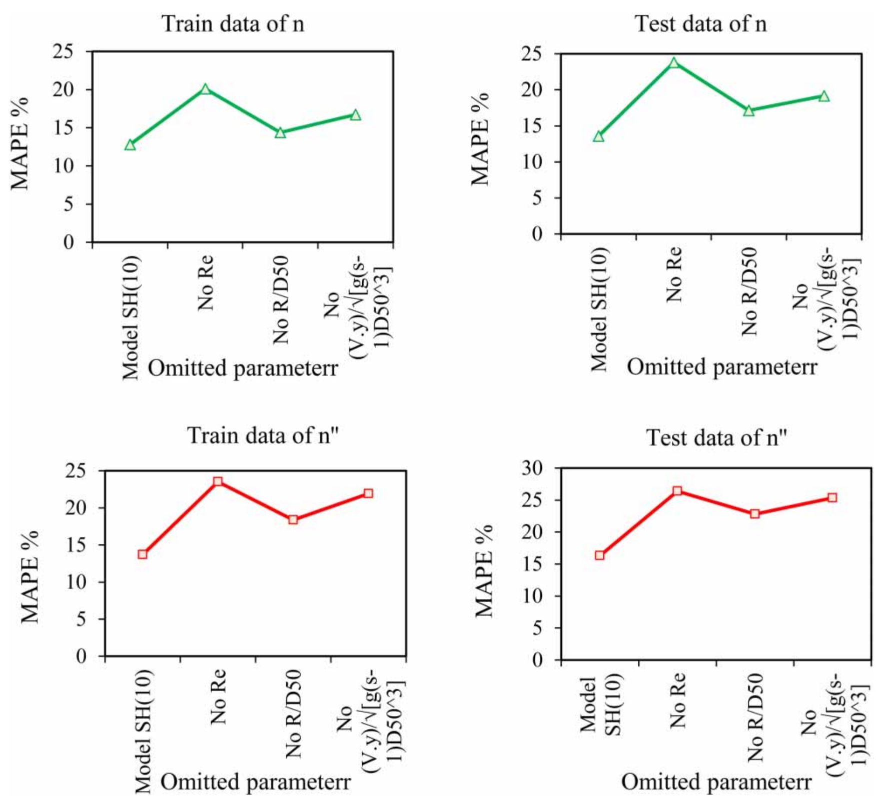

Figure 7 | Comparison of statistical parameters obtained from sensitivity analysis.

seen that *Re* and $Vy/[g\times(s-1)D_{50}^3]^{0.5}$ had the key role in the modeling process.

CONCLUSIONS

In the present study, the capability of the GPR model as a kernel-based approach was verified for predicting the total and bedform roughness coefficients in alluvial channels. In this regard, an experimental study was conducted at the Hydraulic Laboratory of University of Tabriz to investigate the impact of hydraulic and sediment parameters on total and bedform roughness coefficients. Then, these data were combined with other laboratory data sets and the capability of the GPR was assessed in modeling the roughness coefficients. Different input combinations based on flow and sediment characteristics were considered in order to develop the GPR models. The experimental study showed that the Reynolds number has a better correlation with *n* roughness coefficient in comparison with Froude number for the tested conditions in the experiments. It was found that the variation of Manning roughness coefficient due to the bedform were almost in agreement with the variations of the total Manning roughness coefficient, and the variations of skin resistance were in a small range. The roughness variations due to the bedform were almost between 40–80% of the total roughness coefficient. It was shown that the total and bedform roughness coefficients cannot be represented only by means of hydraulic parameters, without including the sediment characteristics. Also, the obtained results from the developed models proved the desired capability of the GPR method in the modeling process. The obtained results revealed that for predicting the roughness coefficient, the model which took the advantages of flow and sediment characteristics performed more successfully than the others. Regarding the total and bedform roughness coefficients with sediment feature characteristics, the model named HS10 with parameters Re, R/D_{50}, *and* $Vy/[g\times(s-1)D_{50}^3]^{0.5}$ was the most accurate model. These results showed that for predicting the total and bedform roughness coefficients under only hydraulic characteristics, including (y/b) as an input parameter significantly improved the efficiency of the models. According to the conducted sensitivity analysis, it was found that the Re and $Vy/[g \times (s-1)D_{50}^3]^{0.5}$ played the most important roles in predicting the total and bedform roughness coefficients in alluvial channels. However, it should be noted that the used method is a data-driven model and the GPR-based model is data sensitive, so further studies using data ranges out of this study and field data should be carried out to determine the merits of the model to estimate roughness coefficient in the real conditions of flow.

REFERENCES

Amaranto, A., Munoz-Arriola, F., Corzo, G., Solomatine, D. P. & Meyer, G. 2018 Semi-seasonal groundwater forecast using multiple data-driven models in an irrigated cropland. *J. Hydroinform.* **20** (6), 1227–1246.

Azamathulla, H. M. & Wu, F. C. 2011 Support vector machine approach for longitudinal dispersion coefficients in natural streams. *Appl. Soft Comput.* **11**, 2902–2905.

Azamathulla, H. M., Haghiabi, A. H. & Parsaie, A. 2017 Prediction of side weir discharge coefficient by support vector machine technique. *Water Sci. Technol. Water Supply* **16** (4), 1002–1016.

Bolat, B. & Yildirim, T. 2004 A data selection method for probabilistic neural networks. *J. Electr. Electron. Eng.* **4** (2), 1137–1140.

Chien, N. & Wan, Z. 1999 *Mechanics of Sediment Transport*. ASCE Press, Reston, VA, USA.

Einstein, H. A. 1950 *The Bed-Load Function for Sediment Transportation in Open Channel Flows*, Vol. 1026. US Department of Agriculture, Washington, DC.

Engel, P. & Lau, Y. L. 1980 Friction factor for two-dimensional dune roughness. *J. Hydraul. Res.* **18** (3), 213–225.

Engelund, F. & Hansen, E. 1967 *A Monograph on Sediment Transport in Alluvial Streams*. Technical University of Denmark 0stervoldgade 10, Copenhagen K.

Ghazanfari-Hashemi, S., Etemad-Shahidi, A., Kazeminezhad, M. H. & Mansoori, A. R. 2011 Prediction of pile group scour in waves using support vector machines and ANN. *J. Hydroinform.* **13** (4), 609–620.

Gilbert, G. K. 1914 *The Transport of Debris by Running Water*. Professional Paper 86, U.S. Geological Survey, US Government Printing Office, Washington, DC.

Guy, H. P., Simons, D. B. & Richardson, E. V. 1966 *Summary of alluvial channel data from flume experiments*. US Geological Survey. *Prof. Pap.* **462** (1), 11–96. US Government Printing Office.

Heydari, H., Zarrati, A. R. & Karimaee Tabarestani, M. 2014 Bed form characteristics in a live bed alluvial channel. *Scientia Iranica. Trans. A Civil Eng.* **21** (6), 1773–1780.

Karim, F. 1999 Bed-form geometry in sand-bed flows. *ASCE J. Hydraul. Eng.* **125** (12), 1253–1261.

Kazemipour, A. K. & Apelt, C. J. 1983 Effects of irregularity of form on energy losses in open channel flow. *Austr. Civil Eng. Trans.* **CE25**, 294–299.

Kennedy, J. F. 1963 The mechanics of dunes and antidunes in erodible_bed channels. *J. Fluid Mech.* **16**, 521–544.

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

Meyer-Peter, E. & Müller, R. 1948 Formulas for bed-load transport. In: *Proc. 2nd Meeting IAHR*, Stockholm, pp. 39–64.

Morvan, H., Knight, D., Wright, N., Tang, X. & Crossley, A. 2008 The concept of roughness in fluvial hydraulics and its formulation in 1D, 2D and 3D numerical simulation models. *J. Hydraul. Res.* **46** (2), 191–208.

Rasmussen, C. E. & Williams, C. K. I. 2006 *Gaussian Processes for Machine Learning*. The MIT Press, Cambridge, MA.

Rouse, H. 1965 Critical analysis of open-channel resistance. *J. Hydraul. Div. Am. Soc. Civ. Eng.* **91** (HY4), 1–25.

Roushangar, K. 2010 *Open Channel Flow Resistance*. PhD thesis, Department of Civil Engineering, University of Tabriz, Iran.

Roushangar, K. & Ghasempour, R. 2017 Prediction of non-cohesive sediment transport in circular channels in deposition and limit of deposition states using SVM. *Water Sci. Technol. Water Supply* **17** (2), 537–551.

Roushangar, K. & Ghasempour, R. 2018 Evaluation of the impact of channel geometry and rough elements arrangement in hydraulic jump energy dissipation via SVM. *J. Hydroinform.* **21** (1), 92–103.

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

Saghebian, S. M. 2018 *Study on Flow Resistance Coefficient Variation of Movable Beds*. PhD thesis, Department of Civil Engineering, University of Tabriz, Iran.

Talebbeydokhti, N., Hekmatzadeh, A. A. & Rakhshandehroo, G. R. 2006 Experimental modeling of dune bedform in a sand bed channels. *Iran. J. Sci. Technol.* **30** (4), 503–516.

Williams, G. P. 1970 *Flume Width and Water Depth Effects in Sediment Transport Experiments*. Professional Paper 562-H. U.S. Geological Survey, US Government Printing Office, Washington, DC.

Yu, X., Liong, S. Y. & Babovic, V. 2004 EC-SVM approach for real-time hydrologic forecasting. *J. Hydroinformatics* **6** (3), 209–23.

First received 6 May 2019; accepted in revised form 30 December 2019. Available online 17 February 2020