

## Modeling open channel flow resistance with dune bedform via heuristic and nonlinear approaches

Kiyoumars Roushangar, Mohammad Taghi Alami and Seyed Mahdi Saghebian

### ABSTRACT

Flow resistance in open channels with dune bedform is a substantial issue due to the influence of dunes on the hydraulic roughness, which can affect the performance of hydraulic constructions. There are a number of nonlinear approaches that have been developed to predict the roughness coefficient in alluvial channels, such as developed equations based on the mean velocity or shear stresses. However, due to the multitude of factors influencing roughness, establishing an accurate determination of the roughness coefficient is difficult. This study applies gene expression programming (GEP) and nonlinear approaches to predict the Manning's coefficient in dune bedform channels. Four different experimental data series were used for modeling. In order to develop the models, three scenarios with different input combinations were considered: scenario 1 considers only flow characteristics, scenario 2 considers flow and bedform characteristics, and scenario 3 considers flow and sediment characteristics. The results proved that GEP is capable of predicting the Manning's coefficient. It was found that for estimation of the roughness coefficient in dune bedform channels, scenario 3 performed more successfully than others. Sensitivity analysis showed that the Reynolds number plays a key role in the modeling process. Comparisons between GEP models and existing equations indicated that GEP models yield better results.

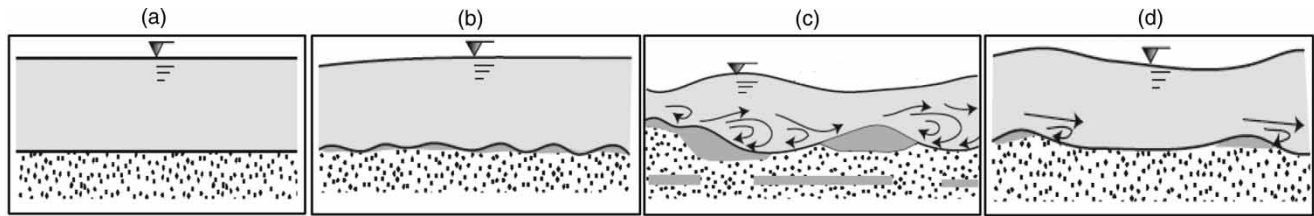
**Key words** | dune bedform, flow characteristics, GEP, nonlinear equations, roughness coefficient

Kiyoumars Roushangar (corresponding author)  
Mohammad Taghi Alami  
Seyed Mahdi Saghebian  
Division of Hydraulics, Department of Civil  
Engineering,  
University of Tabriz,  
Tabriz,  
Iran  
E-mail: kroshangar@yahoo.com

### INTRODUCTION

Determination of the flow resistance (i.e. roughness coefficient) in open channel hydraulics is a substantial and intransitive issue in the design and operation of hydraulic structures, the calculation of water depth, estimation of flow velocity and an accurate characterization of energy losses. Assessment of the roughness coefficient is not a trivial matter given the multitude of factors influencing roughness (e.g. bed material, bedforms, cross-sectional and plan form variability, vegetation, etc.). The complexities and uncertainties of bedform configurations in alluvial channels continue to pose challenges for engineers, which stem from a variety of bedform shapes that arise under different flow conditions. Starting from a plane bed without sediment

transport, ripples, dunes and washed-out dunes develop in large experimental flumes as the flow intensity increases in magnitude over a bed of loose sand particles (Figure 1). Meyer-Peter & Mueller (1948), Einstein & Barbarossa (1952), Taylor & Brooks (1962), Raudkivi (1967), Richardson & Simons (1967), Smith (1968), Van Rijn (1984), Karim (1995), Yang *et al.* (2005), and Van der Mark *et al.* (2008) have all presented expressions for the total roughness coefficient due to bedform roughness. Therefore, a variety of analytical and semi-empirical approaches with both simple and complex structures have been developed to predict the roughness coefficient. Engel & Lau (1980) found that dune length to depth ratio has a significant effect on the



**Figure 1** | Types of bedforms in alluvial channels (Lower regime): (a) Plane bed, (b) Ripple, (c) Dune, (d) Washed-out Dune (Julien 2010).

roughness coefficient 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 the Manning's coefficient related to dune bedforms and the total Manning's 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 do not show the same results under variable flow conditions. These issues cause uncertainty in the prediction of roughness coefficient; therefore, it is extremely critical to utilize methods which are capable of predicting roughness coefficient for the channels with bedform under varied hydraulic conditions.

On the other hand, artificial intelligence tools (e.g. artificial neural networks (ANNs), neuro-fuzzy models (NF), genetic expression programming (GEP) and support vector machine (SVM)) deal with the representation and generalization of datasets using data learning techniques. These methods are remarkable forecasting tools which in the past decade have been implemented in various fields of civil engineering, hydraulic and hydrological studies (e.g. Samui & Dixon 2011; Roushangar *et al.* 2014a, 2016a, 2016b; Sun *et al.* 2014; Roushangar & Koosheh 2015; Nourani *et al.* 2016). Such studies include the prediction of total bed load (Chang *et al.* 2012), prediction of suspended sediment concentration (Kisi *et al.* 2013), prediction of sediment transport in circular channels (Roushangar & Ghasempour 2017), modeling sediment transport (Bhattacharya *et al.* 2007), modeling the rainfall-runoff (Nourani *et al.* 2016), and prediction of total bed material load (Roushangar *et al.* 2014a).

The GEP as an artificial intelligence approach uses fundamental principles of the genetic algorithms (GA) and

genetic programming (GP) and mimics the biological evolution to make a calculation algorithm for predicting a specified phenomenon. The GEP approach has been applied in modeling various components of water resource systems including: developing stage-discharge curves, predicting energy dissipation over spillways (Roushangar *et al.* 2014b), determining the rainfall-runoff process (Kisi *et al.* 2013), modeling bridge pier scour (Azamathulla *et al.* 2010), and examining suspended sediment modeling (Aytek & Kisi 2008).

Computation of the river stage and flow velocity relies on the determination of bedform roughness; therefore, knowledge of the bedform is very important. Predicting the roughness coefficient is a complex phenomenon due to the nonlinearity and uncertainties of the process. Considering the reviewed literature, there is a lack of research on the comprehensive study of predicting the roughness coefficient in alluvial channels with dune bedforms using artificial intelligence. Therefore, the present research used GEP as a heuristic approach to assess the total Manning's coefficient (which includes both grain friction and form resistance) in channels with dune bedform, and also to determine the effective parameters in the estimation of the roughness coefficient in open channels with dune bedform. This method can be successfully applied to predict any variable of interest where the interrelationships among the relevant variables are poorly understood, finding the size and shape of the ultimate solution is difficult, and conventional mathematical analysis methods do not (or cannot) provide analytical solutions (according to Banzhaf *et al.* 1998). The simulations were performed using four different data sets obtained from experimental studies in different laboratories (one original experiment was performed by the author and three other experiments by other researchers). Modeling for prediction

of the Manning's coefficient was carried out under three scenarios and the impacts of flow, bedform, and sediment characteristics were assessed. Then, the outcome of GEP was compared with the results obtained from nonlinear equations. In addition to this analysis, since the knowledge about bedform geometry can provide useful information for hydraulic engineering research and also for investigating the capability of the GEP method in predicting a new output variable, the relative dune height ( $H/y$ ) and dune steepness ( $H/L$ ) were simulated.

## MATERIAL AND METHODS

### Nonlinear approaches

#### Gene expression programming

GEP as a nonlinear approach was developed by Ferreria (2001), using fundamental principles of the GA and GP. GEP is a heuristic search and optimization technique that

mimics the biological evolution to make a computer program to predict a specified phenomenon. The problems are encoded in linear chromosomes with fixed-length as a computer program (Ferreria 2001) which are then expressed or translated into expression trees (ETs). One strength of the GEP approach is its unique, multigenic nature, which allows the evolution of more complex programs composed of several subprograms. A GEP algorithm begins by selecting the five elements: the function set, terminal set, fitness function, control parameters, and stopping condition. Figure 2 shows the overall process of GEP algorithm. According to Figure 2, the process begins with the random generation of the chromosomes of the initial population. Then the chromosomes are expressed and the fitness of each individual is evaluated. The individuals are then selected according to fitness to reproduce with modification, leaving progeny with new traits. The individuals of this new generation are, in turn, subjected to the same developmental process: expression of the genomes, confrontation of the selection environment, and reproduction with modification. The process is repeated for a certain number of generations or until

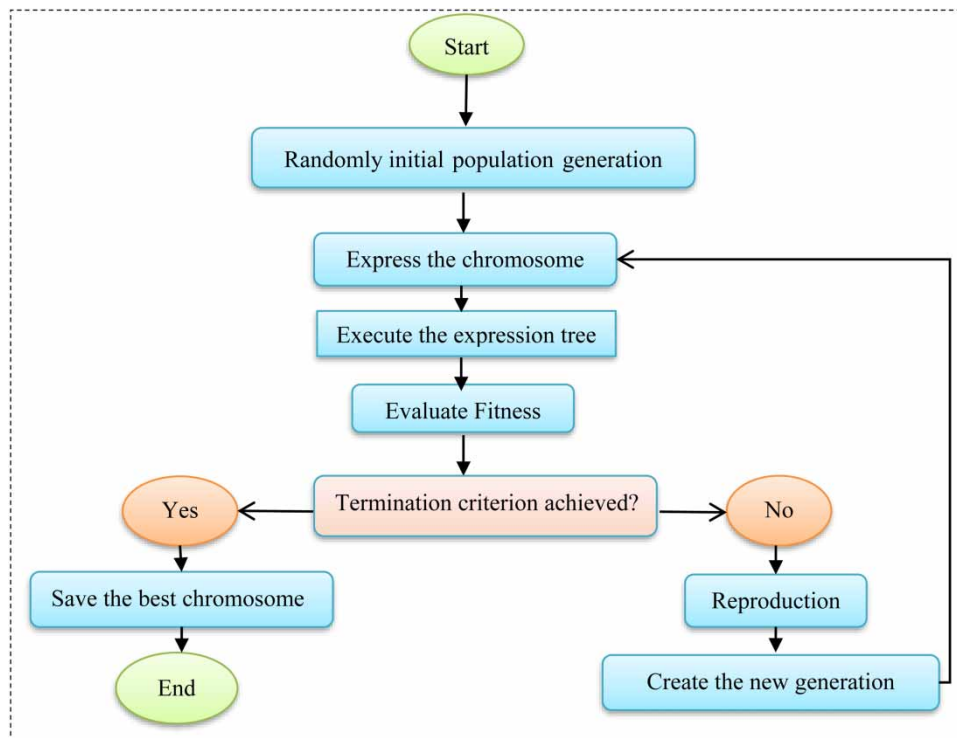


Figure 2 | The schematic view of GEP algorithm (from Ferreria 2001).

a solution has been found (Ferreria 2001). In fact, in GEP the first major step is to identify the set of terminals to be used in the individual computer programs. The major types of terminal sets contain the independent variables of the problem, the state variables of the system and the functions with no arguments. The second step is to determine the set of functions (+, -, ×, /,  $x^a$ , log(x), ln(x), sqrt(x), etc.). The third step is fitness measure, which identifies the method of evaluating how well a given program solves a particular problem. In the current study, root mean square error (RMSE) is taken as fitness function. The fourth step is the selection of certain parameters to control the runs. The control parameters contain the size of the population, the rate of crossover, etc. The last step is the determination of the criteria to terminate the run. In a subsequent step, a comparison between the predicted and actual values is completed. When the desired results are in accordance with the error criteria initially selected, the GEP process is terminated. If the desired error criteria could not be achieved, some chromosomes are chosen by a method called 'roulette wheel sampling', and the selected chromosomes are then mutated to obtain new chromosomes. After the desired fitness score is found, this process terminates and the chromosomes are then decoded for the best solution to the problem (Teodorescu & Sherwood 2008).

*GEP models utilized.* The quality of the data driven model does not depend only on the selected input features. It can also be related to the parameters of the model (GEP). In this study, GEP has been trained for the Manning's coefficient prediction in alluvial channel with dune bedforms. One benefit of genetic programming is the possibility of generating explicit formulae. In this study, four basic arithmetic operators (+, -, ×, /) and some basic mathematical functions ( $x^{1/2}$ ,  $x^2$ ,  $x^3$ ,  $x^{1/3}$ ,  $e^x$ ) were utilized as the GEP function in order to develop simple equations. The architecture of the chromosomes, including number of chromosomes (25-30-35), head size (7-8) and number of genes (3-4), were selected and different combinations of the mentioned parameters were tested. The model was run for a number of generations and was stopped when there was no significant change in the fitness function value and coefficient of correlation. It is observed that the model with number of chromosomes of 30, head size

of seven, and number of genes of three yielded better results. Addition and multiplication were also tested as linking functions and it was found that linking the sub-ETs by addition represented better fitness values. One of the important steps in preparing the GEP model is to choose the set of genetic operators. In the current study a combination of all genetic operators (recombination, mutation, transposition, and crossover) was used for this aim. Mutations can occur anywhere in the chromosome. However, the structural organization of chromosomes must remain intact. In GEP there are three kinds of recombination: one-point, two-point, and gene recombination. In all cases, two parent chromosomes are randomly chosen and paired to exchange some material between them. The transposable elements of GEP are fragments of the genome that can be activated and jump to another place in the chromosome. There are three kinds of transposable elements in the GEP model: short fragments with a function or terminal in the first position that transpose to the head of genes, except to the root (IS elements); short fragments with a function in the first position that transpose to the root of genes (root IS elements or RIS elements); and entire genes that transpose to the beginning of chromosomes (Ferreria 2001). Also, the inversion operator is restricted to the heads of genes. Here any sequence might be randomly selected and inverted. The certain rates of genetic operators which define a certain probability of a chromosome were determined based on a trial and error process. Each GEP model was evolved (trained model) until there was no significant change in the fitness function value, then the program was stopped (fixed model). Therefore, the GEP model optimized parameters were determined. The order of the data sets was selected in such a way that the training data set contains a representative sample of all the behavior in the data in order to obtain the model with higher accuracy. To find a good training set which can give a good accuracy both in training and testing sets, one method is the instance exchange (Bolat & Yildirim 2004). The process starts with a random selected training set. Parameters of the optimized GEP model are shown in Table 1. It should be noted that in this study, for all models, 75% of the entire dataset was used for training goals and the remaining 25% of the data were used for testing goals.

**Table 1** | Optimized parameters of GEP models used in this study

Description of parameter	Setting of parameter
Function set	+, +, -, ×, /, $x^{\frac{1}{2}}$ , $x^2$ , $x^3$ , $x^{\frac{1}{3}}$ , $e^x$
Chromosomes	30
Head size	7
Number of genes	3
Linking function	Addition
Fitness function error type	Root mean square error (RMSE)
Mutation rate	0.044
Inversion, IS and RIS transposition rate	0.1
One- and two-point recombination rate	0.3
Gene recombination and transposition rate	0.1

The same training set and the same testing set were used for all models.

### Nonlinear equations to alluvial channel resistance

There are different concepts and approaches that are used in the derivation and extraction process of roughness coefficient formulae, where some are simple and some are complex, such as linear and nonlinear approaches. Based on the linear superposition approach, the roughness coefficient is separated into the friction resistance and form resistance components. The components of the roughness coefficient vary according to the bed conditions (Yen 2002). In linear modeling of Manning's coefficient in movable beds, it is necessary to calculate both plane bed resistance and form resistance. This is a complex and time-consuming process. In the nonlinear approach, the resistance factor is not separated into grain roughness (friction resistance) and bedform roughness (form resistance), as in the linear superposition approach. Instead, it is kept as a single factor. Most of the available nonlinear approaches are developed based on dimensional analysis and statistical fitting of data to the parameters considered in the functional relationships. According to the reviewed literature, there are no comprehensive roughness coefficient formula specified only for alluvial channels with dune bedform. Due to this fact, the nonlinear approach is

selected based on the classification of Yen (2004) which can be divided into four groups: (a) those that consider the roughness coefficient as a dependent variable; (b) methods based on shear stresses; (c) equations based on the mean velocity; and (d) the energy approach derived from Bagnold's (1966) stream power concept for expressions of velocity that can be used to calculate resistance. Most of these methods implicitly assume the flow to be steady, reachwise uniform with equilibrium sediment transport. Furthermore, among the four groups, only group (a) considers the roughness coefficient explicitly, whereas the other three groups consider the roughness indirectly and implicitly.

Table 2 represents the used nonlinear equations in this study. The approach developed by Strickler (1923) considers only the median grain diameter ( $D_{50}$ ). The formula developed by Bruschin (1985) was designed by considering the grain and total Darcy–Weisbach roughness coefficient ( $f_0$  and  $f$ ), median grain diameter, the energy slope gradient ( $S_f$ ), and hydraulic radius ( $R$ ). Both approaches consider the  $D_{50}$  as the main parameter. Similar to Bruschin's formula, Karim (1995) developed a formula without using  $S_f$  valid for Froude numbers below 0.4. For values higher than 0.4, different formulae were used. On the other hand, the approach developed by Camacho & Yen (1989) took advantage of the Reynolds ( $Re$ ) and Froude numbers together with  $R$  and  $D_{50}$ . The formula varies for Froude numbers between 0.4 and 0.7 and between 0.7 and 1. Griffiths (1981) studied the hydraulic resistance in coarse gravel-bed rivers. Based on dimensional analysis and statistical justifications, he proposed relations for gravel rivers with rigid and movable beds using 136 field data. Yalin (1977) proposed that for subcritical flows, the roughness coefficient is a function of the grain Reynolds number, mobility number and grain size to depth ratio (see Table 2). Garde & Ranga Raju (1966) developed a roughness coefficient equation based on velocity. In this approach, the mean velocity was viewed as an implicit representation of the resistance. White et al. (1987) investigated the roughness coefficient in alluvial channels based on mixed momentum through an energy approach, resulting in a proposed equation for calculating the Darcy–Weisbach roughness coefficient. In this study, the output variable is parameter  $n$ , therefore, the equation

**Table 2** | Semi-empirical equations used in this study (from Yen 2004)

Researcher	Dependent variable	Need for a prior knowledge on bed	Data used	Condition
<i>(a) Based on resistance</i>				
Bruschin (1985)	$n = \frac{D_{50}}{12.8} \times \left[ \frac{R \times s_f}{D_{50}} \right]^{1/7.3}$	No	Laboratory and field	$D_{50}$ in millimeter
Karim (1995)	$n = 0.037 D_{50}^{0.126} \left[ \frac{f}{f_0} \right]^{0.465}$	No	Laboratory and field	$D_{50}$ in meter
Camacho & Yen (1989)	$n = \frac{R^{\frac{1}{6}}}{\sqrt{g}} \times 3.78 T^{0.175} / Re^{0.19}$	No	Laboratory and field	$Fr < 0.4$
	$n = \frac{R^{\frac{1}{6}}}{\sqrt{g}} \times 0.0081 \left( \frac{D_{50}}{R} \right)^{0.125} \times \frac{Re^{0.05}}{Fr^{0.88}}$			$0.4 < Fr < 0.7$
	$n = \frac{R^{\frac{1}{6}}}{\sqrt{g}} \times 0.078 \left( \frac{D_{50}}{R} \right)^{0.166} / Fr^{0.444}$			$0.7 < Fr < 1$
Strickler (from Yen 2004)	$n = 0.0474 D_{50}^{1/6}$	No	Laboratory and field	$D_{50}$ in meter
Griffiths (1981)	$\frac{1}{\sqrt{f}} = 2.21 \left( \frac{V}{\sqrt{g D_{50}}} \right)^{0.34}$	Yes (moving bed or not)	Field	$Fr < 1$
<i>(b) Based on shear stresses</i>				
Yalin (1977)	$\frac{8}{\sqrt{f}} = f \left( \frac{V \times D_{50}}{v}, \frac{V^2}{(\Delta \rho_s / \rho) g D_{50}}, \frac{D_{50}}{y} \right)$	No	Laboratory	As graphic results
<i>(c) Based on the mean velocity</i>				
Garde & RangaRaju (from Yen 2004)	$\frac{V}{\sqrt{(\Delta \rho_s / \rho) g D_{50}}} = K \left( \frac{R}{D_{50}} \right)^{0.67} \left( \frac{S}{\Delta \rho_s / \rho} \right)^{0.5}$	Yes	Laboratory and field	$K = 6$ (for dune bedform)
<i>(d) Based on mixed momentum and energy</i>				
White et al. (1987)	$f = 8 \left( \sqrt{(g D_{50}) \times (s_f)} / (V) \right)^2$	No	Laboratory	$Fr < 1$

R: hydraulic radius,  $D_{50}$ : mean grain diameter,  $s_f$ : Energy slope gradient,  $f_0$ ,  $f$ : grain and total Darcy–Weisbach roughness coefficient,  $T^*$ : Camacho parameter,  $g$ : acceleration due to gravity,  $Fr = v/(gy)^{0.5}$ : Froude number,  $Re = (\rho v R/\mu)$ : Reynolds number,  $v$ : flow velocity,  $n$ : Manning's coefficient,  $y$ : flow depth,  $\Delta \rho_s = \rho_s - \rho$ ,  $\rho_s$  and  $\rho$  are sediment and flow density, respectively.

$\sqrt{(f/8)} = (n\sqrt{g})/(R^{1/6})$  was used to convert the dependent variables.

## Dune bedform

From extensive laboratory experiments at Colorado State University by Simons & Richardson (1963, 1966), a classification of several bedforms types was performed in terms of governing non-dimensional parameters. Flat bed, or plane bed, refers to a bed surface without bedforms. Ripples are small bedforms with wave heights less than a few cm ( $\sim 3.048$  cm) and only seen in cases of hydraulically smooth bed conditions. Ripple shapes vary from nearly triangular to almost sinusoidal. Dunes are much larger than ripples and are out of phase with the water

surface waves. Dunes are often triangular with relatively gentle upstream slopes and downstream slopes approaching the angle of repose of the bed material (Julien 2010).

## Dune geometry

The geometry of dune bedforms, such as their height and resulting decrease of water depth for navigation, can be a concern. Also their roughness might increase water levels which increase flood risk. The geometry of dune bedform refers to the representative dune height ( $H$ ) and length of dunes ( $L$ ) as a function of the average flow depth ( $y$ ), median bed particle diameter ( $D_{50}$ ), and other flow parameters, such as the grain shear Reynolds number. Figure 3 shows the dune bedform geometry.



Figure 3 | Dune bedform geometry (Roushangar 2010).

### Flow resistance

The laws of conservation of energy and momentum must be considered for hydraulic resistive forces in the calculation of open channel hydraulics. Steady and uniform flow conditions are needed for the driving and resisting forces to be balanced, where flow is not accelerating nor decelerating, so the average channel cross-section, slope, and velocity are assumed to be constant under constant discharge conditions. In natural streams, velocity or discharge must often be predicted or calculated using other flow parameters, most commonly hydraulic radius, energy slope (or some approximation), and an estimate of channel roughness. Even though natural streams do not strictly comply with uniform flow assumptions, uniform flow conditions are often assumed to simplify velocity and discharge computations. The roughness coefficient in river engineering applications is commonly calculated using one of the three equations: Manning, Chézy, or Darcy–Weisbach (Yen 2002).

### Experimental data used in the study

In this paper, the following four kinds of data corresponding to the roughness coefficient in open channels with dune bedform selected from published literature are used:

(I) 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.

- (II) As a part of the research programme of the Water Resources Division of the US Geological Survey, a project was organized at Colorado State University between 1956 and 1961 to determine the effects of the size of bed material, temperature of flow, and the fine sediment in the flow of the hydraulic and transport variables. The investigations for each set cover flow phenomena ranging from a plane bed with no sediment movement to violent anti-dunes (Guy *et al.* 1966).
- (III) The US Army Corps of Engineers Waterways Experiment Station in 1935 organized several experiments to research sediment transport.
- (IV) Roushangar (2010) studied the roughness coefficient in alluvial channels. New original experiments were carried out in the hydraulic laboratory of Caen University using a 5 m long, 0.15 m wide, and 0.25 m high rectangular flume. Natural quartz sand with a specific gravity of 2.65 and uniform average diameters of 0.15 and 0.4 mm were used as sediment particles in the experiments. By changing the flow depth and discharge, different average velocities, Froude numbers, dune height, and other hydraulic parameters were measured or calculated. The ranges of selected parameters used in the tests are listed in Table 3, where  $b$  represents the channel width. It should be noted that the measured values of the Manning's coefficient are calculated from Manning's famous equation which is simple in form and the formula is well proven by much practical experience.

### Performance criteria

The main statistical criteria used to evaluate the performance of the different proposed models include the root mean

**Table 3** | Details of the utilized experimental data

Parameters	Researchers			
	Williams (1970)	Guy et al. (1966)	WSA (1935)	Roushangar (2010)
y (mm)	87.1–222	91.4–405	65.5–208	71–145
b (mm)	76.2–1,118	6,092,438	705,736	150
D <sub>50</sub> (mm)	1.35	0.19–0.95	0.18–0.47	0.15, 0.4
Fr	0.34–0.84	0.25–0.65	0.3–0.72	0.21–0.40
Re	11,932–101,920	46,800–255,500	19,061–66,432	24,192–45,869
n	0.0091–0.0201	0.015–0.038	0.0127–0.0249	0.0203–0.0252
f	0.019–0.059	0.031–0.163	0.032–0.108	0.039–0.063
Data number	89	114	61	54

square error (RMSE), mean absolute percentage error (MAPE), correlation coefficient (R), Nash–Sutcliffe factor (NS or DC). 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, the RMSE describes the average difference between predicted and measured values corresponding to the predicted values, and the MAPE indicates the absolute differences between estimated and observed values as depicted in Equation (1). These measurements are not oversensitive to extreme values (outliers), but are rather sensitive to additive and proportional differences between model predictions and observations. Therefore, correlation-based measures (e.g. DC and R statistic) can indicate that a model is a good predictor (Legates & McCabe 1999). Evidently, a high value for DC and R (up to one) and a small value for RMSE and MAPE indicate a high efficiency of the model.

$$DC = 1 - \frac{\sum_{i=1}^N (l_o - l_p)^2}{\sum_{i=1}^N (l_o - \bar{l}_p)^2},$$

$$R = \frac{\sum_{i=1}^N (l_o - \bar{l}_o) \times (l_p - \bar{l}_p)}{\sqrt{\sum_{i=1}^N (l_o - \bar{l}_o)^2 \times \sum_{i=1}^N (l_p - \bar{l}_p)^2}},$$

$$RMSE = \sqrt{\sum_{i=1}^N \frac{(l_o - l_p)^2}{N}}, \quad MAPE = \frac{1}{\bar{l}_o} \frac{\sum_{i=1}^n |l_o - l_p|}{N} \times 100 \quad (1)$$

where  $l_o$ ,  $l_p$ ,  $\bar{l}_o$ ,  $\bar{l}_p$ , and  $N$  respectively are: the measured value, predicted value, mean of measured values, mean of

predicted values and number of data samples. The input and output variables were normalized before utilizing as model input in order to avoid the zero and minus predictions and to ensure that all variables receive equal attention during the calibration of the model. This will also increase the training speed and capability of the model. The following equation was used to normalize the utilized data in this study:

$$x_n = \alpha + \beta \times \left( \frac{x - x_{\min}}{x_{\max} - x_{\min}} \right) \quad (2)$$

where  $x_n$ ,  $x$ ,  $x_{\max}$ ,  $x_{\min}$  respectively are: the normalized value of variable  $x$ , the original value, the maximum and minimum of variable  $x$ . The  $\alpha$  and  $\beta$  parameters are fixed values that may differ depending between which numbers the normalization range rests, and can be determined based on trial and error process. The normalization range of parameters considered in this study is [0.1–1] as suggested by Dawson & Wilby (1998). Therefore, the values of  $\alpha$  and  $\beta$  are 0.1 and 0.9 respectively. The used normalization only rescales the data to the range of 0.1 and 1, which after modeling could be de-normalized to real values.

## SIMULATION AND MODEL DEVELOPMENT

### Input variables for simulating the Manning's coefficient

Selection of various parameters as input combinations can affect the accuracy of the results throughout the modeling



process. Thus, selection of appropriate parameters and input combinations plays a key role during modeling. Therefore, in order to determine the output variables (Manning's coefficient) under three different scenarios, a set of dimensionless variables as input combinations were attempted to develop the process. The scenarios considered in this study are as follows.

(I) *Models based on flow characteristics; Scenario 1*

Initially, the models were defined according to flow characteristics. The roughness coefficient ( $n$ ) based on flow characteristics could be described as a function of dimensionless variables as follows:

$$n = f\left(Re, Fr, \frac{y}{b}\right) \quad (3)$$

(II) *Models based on flow and bedform characteristics; Scenario 2*

In order to consider both flow and bedform features the in modeling process, the models were defined according to the flow and bedform characteristics. The roughness coefficient based on these characteristics could be described as a function of the dimensionless variables as follows:

$$n = f\left(Re, Fr, \frac{H}{L}, \frac{L}{y}, \frac{H}{y}\right) \quad (4)$$

where  $H/L$ ,  $L/y$  and  $H/y$  are bedform characteristics. In Equation (4),  $H$  and  $L$  are dune height and dune length respectively.

(III) *Models based on flow and sediment characteristics; Scenario 3*

To consider the flow and sediment features in roughness coefficient estimation, the flow and sediment characteristics were employed in the establishment of input combination. The roughness coefficient based on the characteristics could be described as a function of dimensionless variables as follows:

$$n = f\left(Re, Fr, (R/D_{50}), \frac{V}{\sqrt{g(s-1)D_{50}}}, \frac{Vy}{\sqrt{g(s-1)D_{50}^3}}\right) \quad (5)$$

where  $V$  and  $s$  stand for flow velocity and specific gravity. The last two parameters of Equation (5) are included as independent variables of the model since they include most of the dimensional sediment and flow variables (except viscosity). Table 4 shows the GEP developed models for simulating the Manning's coefficient. The Froude number deals with the relationship between gravity and inertial forces, whereas the Reynolds number deals with the relationship between frictional and inertial forces. The similarity requirements posed by the Froude and Reynolds numbers can typically not be satisfied simultaneously. Therefore, the input combination of Reynolds and Froude numbers were not used during this study.

**Input variables for simulating the relative dune height**

So far several studies have been carried out in the field of bedforms geometry. In many practical applications, it was mentioned that it is desirable and sometimes necessary to predict the ratio of dune height to flow depth. It is an important factor in modeling, since it can provide useful information for hydraulic engineering research. Therefore, different input combinations were used here to investigate the relative dune height. For predicting this parameter, some of the developed models for the Manning's coefficient, which represented higher accuracy, were selected. The input parameters in this state are as:

$$Re, \frac{y}{b}, \frac{R}{D_{50}}, \frac{V}{\sqrt{g(s-1)D_{50}}}, \frac{Vy}{\sqrt{g(s-1)D_{50}^3}}$$

Figure 4 demonstrates the model combinations and modeling methodology in this study.

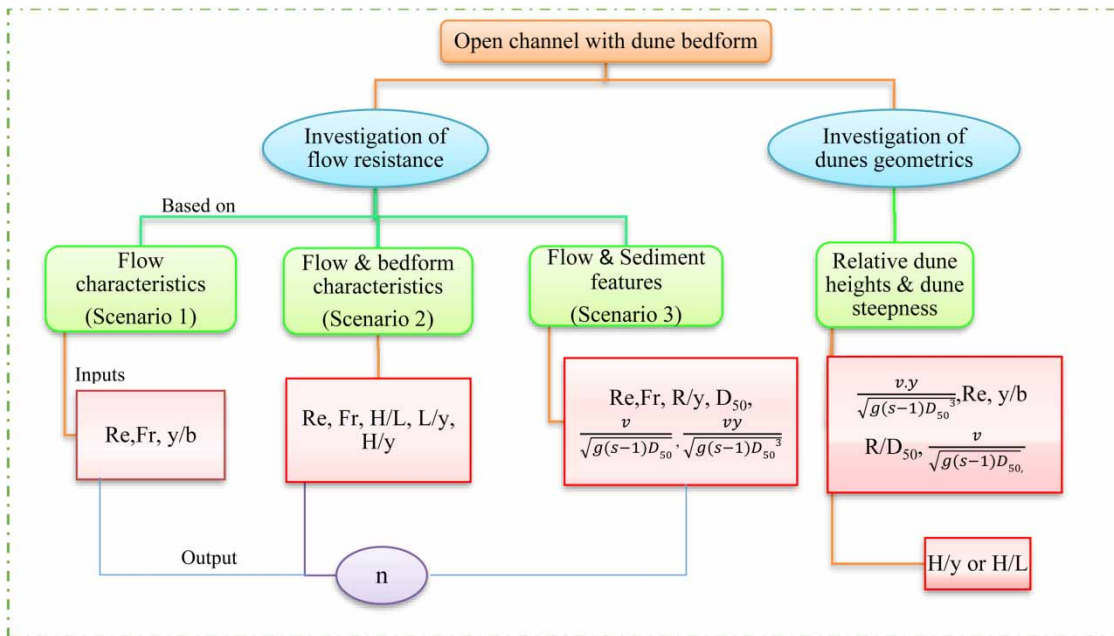
## RESULTS AND DISCUSSION

### Models developed based on flow characteristics; Scenario 1

In the first scenario, models were defined according to the flow characteristics. Table 5 sums up the performance evaluation indices of the models with flow characteristics. According to Table 5, the first three combinations were included with only one variable. Results revealed that the parameter  $Re$  led to the best results amongst the single parameter

**Table 4** | GEP developed models for predicting Manning's *n*

Scenario	Inputs	Scenario	Inputs
1	Re	3	Re
Models based on flow characteristics	Fr	Models based on flow and sediment characteristics	Fr
	y/b		R/D <sub>50</sub>
	Re, y/b		$\frac{V}{\sqrt{g(s-1)D_{50}}}$
	Fr, y/b		$\frac{Vy}{\sqrt{g(s-1)D_{50}^3}}$
2	Re		Fr, R/D <sub>50</sub>
Models based on flow and bedform characteristics	Fr		$\frac{V}{\sqrt{g(s-1)D_{50}}}, R/D_{50}$
	Re, L/y		$\frac{V}{\sqrt{g(s-1)D_{50}}}, Re$
	Re, y/H		$\frac{Vy}{\sqrt{g(s-1)D_{50}^3}}, Re$
	Re, L/H		$\frac{V}{\sqrt{g(s-1)D_{50}}}, R/D_{50}$
	Re, y/H, L/H		$Re, R/D_{50}, \frac{Vy}{\sqrt{g(s-1)D_{50}^3}}$
			$R/D_{50}, \frac{V}{\sqrt{g(s-1)D_{50}}}, \frac{Vy}{\sqrt{g(s-1)D_{50}^3}}$
	Fr, y/H, L/H, L/y		$Re, \frac{V}{\sqrt{g(s-1)D_{50}}}, \frac{Vy}{\sqrt{g(s-1)D_{50}^3}}, R/D_{50}$



**Figure 4** | Schematic view of scenarios considered in this study.

**Table 5** | Performance evaluation of models with flow characteristics for test sets

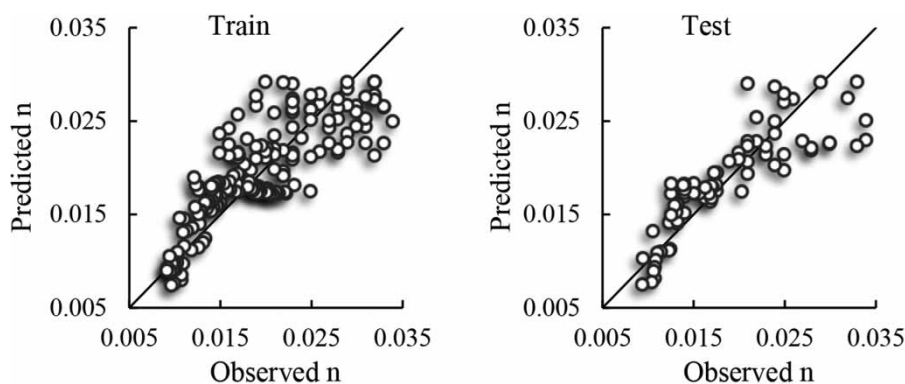
Model	Performance criteria for test series			
	R	DC	RMSE	MAPE
Re	0.8025	0.5304	0.0049	19.06
Fr	0.5384	0.1285	0.0067	26.37
y/b	0.7695	0.3154	0.0053	22.90
Re, y/b	0.8329	0.6926	0.0047	15.82
Fr, y/b	0.6523	0.2586	0.0054	20.21
Re > 80,000	0.4452	0.2386	0.0060	24.32
Re < 80,000	0.8432	0.7153	0.0041	13.21

models for the Manning's coefficient. The next two combinations were defined with two variables. The combination including  $Re$  and  $y/b$  proved to be the best double parameter GEP model with flow characteristics for predicting the roughness coefficient. According to the results of Table 5, it could be seen that using the Reynolds number as the only input parameter yielded an acceptable accuracy, while using only the Froude number as input did not lead to a desired prediction performance. Adding parameter  $y/b$  to input parameters caused a fair increase in model efficiency. It could be deduced that  $Re$  and  $y/b$  were the most efficient parameters in modeling the Manning's coefficient based on flow characteristics. Figure 5 presents the scatter plot of the observed data versus model results that uses these two flow characteristics as input. According to Figure 5, the lower values are predicted more accurately compared to the higher values of the Manning's coefficient. 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. Results are represented in Table 5 (last two rows) and Figure 6. These findings demonstrate that when the Reynolds number is less than 80,000, the model tends to be more accurate in predicting the Manning's coefficient. Figure 6 highlights the better model performance for Reynolds numbers less than 80,000 ( $Re < 80,000$  category), in comparison to the higher range of the Reynolds number ( $Re > 80,000$ ).

### Models developed based on flow and bedform characteristics; Scenario 2

In order to consider both flow and bedform features in modeling processes, the models were defined based on flow and bedform characteristics. Considering the Manning's coefficient simulations with bedform characteristics, the GEP was fed with single, double and triple input variables. The first input combination with only  $Re$  as an input variable led to the best results for the prediction of the roughness coefficients. Moreover, adding  $H/L$ ,  $H/y$  and  $L/H$  ( $H$  and  $L$  respectively are the dune's height and length) to the Reynolds number ( $Re$ ) had less impact on the results of Manning's coefficient predictions. The next three input datasets consisted of a double input combination. Among the mentioned input combinations, the one including  $Re$  and  $L/y$  ended with a better outcome. The last four datasets with more than two input combinations did not perform as well as the single input ( $Re$ ). Results of models in the verification step for the Manning's coefficient with bedform datasets are presented in Table 6. Based on the outcome, the developed models

**Figure 5** | Scatter plots of predicted Manning's coefficient vs. observed values using non-normalized data; flow characteristics.

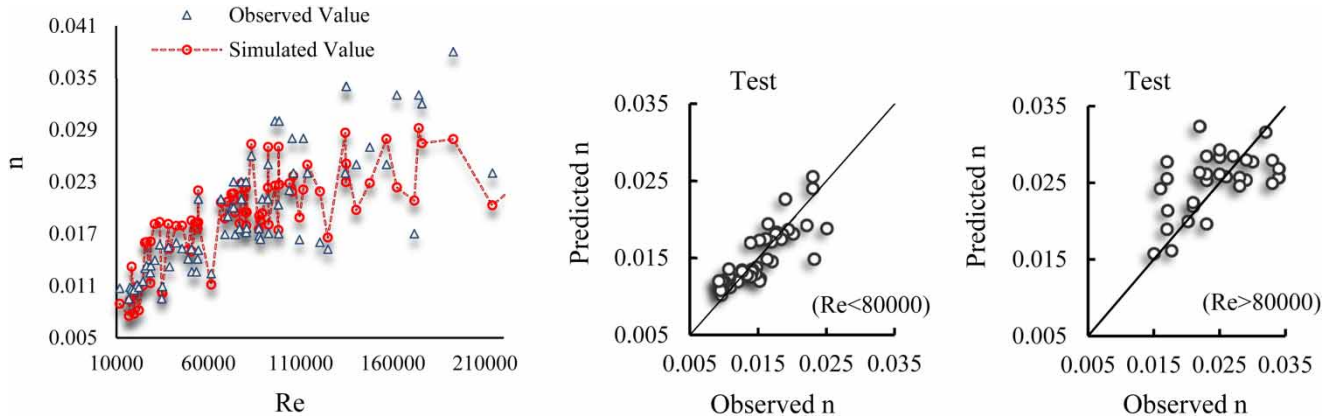


Figure 6 | Scatter plots of predicted Manning's coefficient vs. observed values for different ranges of Re.

Table 6 | Evaluation of predicted Manning's coefficient with bedform characteristics for test sets

Model	Performance criteria for test series			
	R	DC	RMSE	MAPE
Re	0.8025	0.5304	0.0049	19.06
Fr	0.5384	0.1285	0.0067	28.37
Re, L/y	0.8009	0.5220	0.0053	21.01
Re, y/H	0.7986	0.5111	0.0059	23.05
Re, L/H	0.7987	0.5128	0.0057	22.46
Re, y/H, L/H	0.8010	0.5201	0.0054	21.69
Re, y/H, L/H, L/y	0.7992	0.5185	0.0055	22.11
Fr, y/H, L/H, L/y	0.6530	0.2390	0.0065	25.37
Re, L/y, L/H	0.7992	0.5282	0.0052	20.87

with flow and bedform characteristics performed relatively weakly in comparison to developed models with only flow characteristics. Figure 7 presents the scatterplot obtained

from developed models using flow and bedform characteristics for the best performing model (the first model).

### Developed models with flow and sediment characteristics; Scenario 3

Developed models based on flow characteristics along with modeling via flow and bedform characteristics demonstrated that the flow characteristics were capable of more precise predictions so far. Moreover, to consider the flow and sediment features in roughness coefficient estimation, flow and sediment characteristics were employed in the establishment of input combination. To measure the significance of the sediment features in the modeling of flow resistance, sensitivity analysis was performed for predicting the variables based on the performance of models. Results in Table 7 indicate that the input combination including Re,  $R/D_{50}$  and  $(Vy/\sqrt{g(s-1)D_{50}^3})$  led to the best consequence

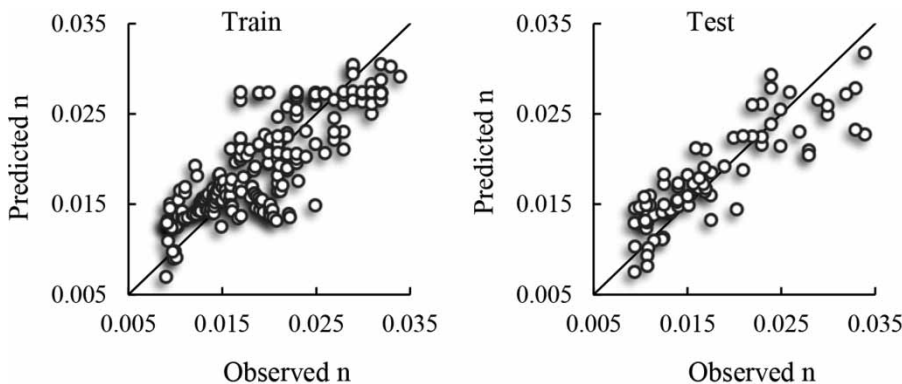


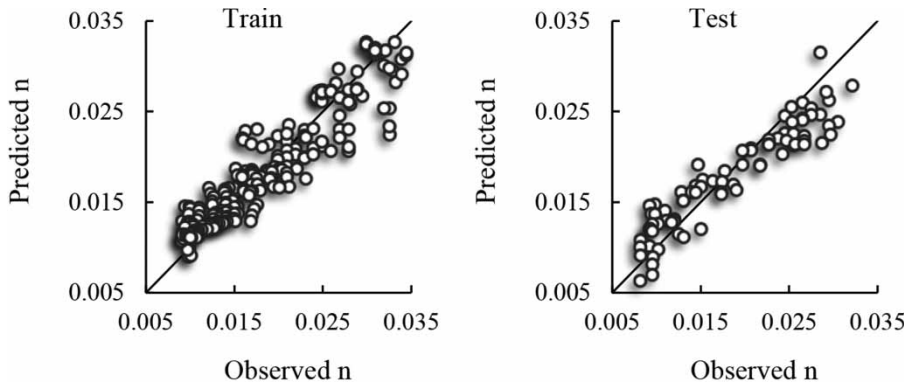
Figure 7 | Scatter plot of predicted Manning's coefficient vs. observed values; flow and bedform characteristics.

**Table 7** | Evaluation of predicted Manning’s coefficient with sediment feature characteristics for test sets

Model	Performance criteria				Model	Performance criteria			
	R	DC	RMSE	MAPE		R	DC	RMSE	MAPE
Re	0.8025	0.5304	0.0049	19.06	$\frac{V}{\sqrt{g(s-1)D_{50}}}$ , Re	0.8227	0.6383	0.0037	17.02
Fr	0.5384	0.1285	0.0067	26.37	$\frac{Vy}{\sqrt{g(s-1)D_{50}^3}}$ , Re	0.8287	0.6388	0.0036	16.06
R/D <sub>50</sub>	0.7595	0.5456	0.0047	22.36	$\frac{Vy}{\sqrt{g(s-1)D_{50}^3}}$ , R/D <sub>50</sub>	0.8253	0.5485	0.0046	22.29
$\frac{V}{\sqrt{g(s-1)D_{50}}}$	0.7410	0.3176	0.0055	24.84	Re, R/D <sub>50</sub> , $\frac{V}{\sqrt{g(s-1)D_{50}}}$	0.8642	0.6392	0.0039	18.30
$\frac{Vy}{\sqrt{g(s-1)D_{50}^3}}$	0.7725	0.4105	0.0052	21.34	Re, R/D <sub>50</sub> , $\frac{Vy}{\sqrt{g(s-1)D_{50}^3}}$	0.8662	0.7425	0.0035	13.87
Re, R/D <sub>50</sub>	0.8492	0.6342	0.0038	18.32	R/D <sub>50</sub> , $\frac{V}{\sqrt{g(s-1)D_{50}}}$ , $\frac{Vy}{\sqrt{g(s-1)D_{50}^3}}$	0.8130	0.5331	0.0048	22.40
Fr, R/D <sub>50</sub>	0.6541	0.4339	0.0051	21.85	Re, $\frac{V}{\sqrt{g(s-1)D_{50}}}$ , $\frac{Vy}{\sqrt{g(s-1)D_{50}^3}}$ , R/D <sub>50</sub>	0.8659	0.6304	0.0041	18.27
$\frac{V}{\sqrt{g(s-1)D_{50}}}$ , R/D <sub>50</sub>	0.8144	0.5569	0.0044	22.11					

for the prediction of the Manning’s coefficient. In other words, it could be inferred that the ratio of hydraulic radius to the sediment diameter and Reynolds number as the most effective input combination led to the best results in predicting the Manning’s coefficient. Therefore, both hydraulic and sediment characteristics have a significant effect on the prediction process. From the results it is clear that adding  $(Vy/\sqrt{g(s-1)D_{50}^3})$  (relative discharge) to Re and R/D<sub>50</sub> caused an improvement in the results of the

Manning’s coefficient. Looking into the data presented in Table 7, using Re and R/D<sub>50</sub> as the individual input variables shows almost equal performance as the three combination. However, the use of R/D<sub>50</sub> seems to be a more practical option since velocity measurements in the field situations are typically very difficult. The scatter plot of the observed and predicted Manning’s coefficient for the best input combination with sediment characteristics is shown in Figure 8.



**Figure 8** | Scatter plot of predicted Manning’s coefficient vs. observed values; flow and sediment feature characteristics.

**Table 8** | Verification statistics of relative dune height for test series

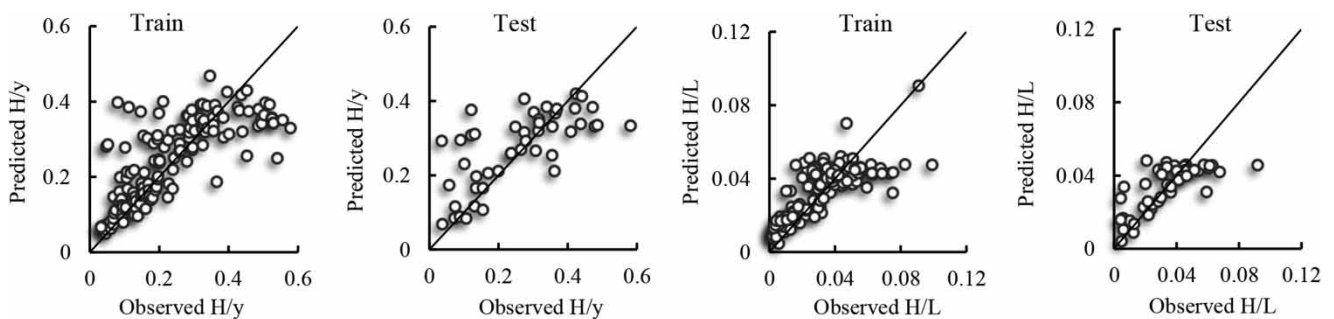
Input parameters	Output parameter: H/y Performance criteria				Output parameter: H/L Performance criteria			
	R	DC	RMSE	MAPE	R	DC	RMSE	MAPE
Re	0.7002	0.2732	0.1515	39.51	0.6166	0.3525	0.0239	13.11
Re, R/D <sub>50</sub>	0.6985	0.2679	0.1618	42.48	0.4299	0.3125	0.0244	15.31
Re, y/b	0.7025	0.3995	0.1414	35.83	0.6112	0.3585	0.0235	12.98
Re, y/b, $\frac{v}{\sqrt{g(s-1)D_{50}}}$	0.7105	0.4704	0.1307	31.23	0.7216	0.4215	0.0226	10.07
Re, y/b, R/D <sub>50</sub>	0.7015	0.3574	0.1472	37.84	0.5109	0.1854	0.0262	17.42
Re, y/b, R/D <sub>50</sub> , $\frac{v}{\sqrt{g(s-1)D_{50}}}$	0.7063	0.4648	0.1314	32.84	0.7201	0.3254	0.0241	14.19
Re, y/b, R/D <sub>50</sub> , $\frac{v}{\sqrt{g(s-1)D_{50}}}$ , $\frac{v \cdot y}{\sqrt{g(s-1)D_{50}^3}}$	0.7065	0.4525	0.1364	33.18	0.7195	0.3214	0.0242	14.84

### Prediction of relative dune height (H/y and H/L)

So far, several different studies have been implemented in the field of predicting the roughness coefficient. In many practical applications, it was mentioned that it is desirable and sometimes necessary to predict the ratio of dune height to flow depth (H/y). For this purpose, in the present study different parameters were used as input combinations to investigate their role in the modeling of relative dune height. Results of models are presented in Table 8. According to the listed performances in Table 8, it was observed that the combination of Re, y/b, and  $(V/\sqrt{g(s-1)D_{50}})$  led to the best outcome for prediction of H/y (with the highest R and DC and the lowest RMSE and MAPE). Furthermore, these three parameters were also dominant

parameters for predicting the ratio of dune height to its length (H/L).

By comparing the various input combinations, it can be concluded that adding  $(R/D_{50})$  to other parameters does not cause an increase in the accuracy of the results. According to the results, it can be inferred from Table 8 that the use of  $V/\sqrt{g(s-1)D_{50}}$  as input parameter caused an increment in models efficiency. Finally, according to the results, it is almost impossible to draw a sensible relation between H/y and each one of the input parameters solely. Resultantly, it can be assumed that the relative dune height has a suitable correlation with a combination of Re, (y/b) and  $(V/\sqrt{g(s-1)D_{50}})$ . In Figure 9, the scatter plot of the simulated relative dune height vs. the observed values for this best performing model for dune height prediction is

**Figure 9** | Scatter plot of simulated relative dune height vs. observed values using non-normalized data.

**Table 9** | The developed equations by GEP (n values are normalized  $n \rightarrow n_{norm}^a$ )

Condition	Formula	Consideration
Manning's coefficient		
Scenario 1	$n = d_0 \times d_1 \times [(d_1 - c_0) \times c_0 d_1] + d_1 \times \frac{d_1}{d_1^3 + \sqrt[3]{d_1}} + e^{(d_0 - c_1 \times d_0) - (d_1^2 + d_1)}$ $n = \left( \frac{c_0}{2(c_0 - d_0) - (d_0 d_1)} \right) + \frac{(d_1 + c_0) - (d_1 - d_0)}{c_0^3 - (c_0 - d_0)} + \frac{c_1}{c_0^2 - d_1}$ $n = \frac{(d_0 - c_1)/(c_0 - d_0)}{(c_1 - d_1)(c_0 - d_1)} + \frac{c_0^2}{(c_0 - d_0)} \times \left( \frac{c_0}{c_1} + (d_1 + c_0) \right) + \frac{c_0 - 2c_1}{d_0 d_1 + c_1}$	$d_0 = y/b, d_1 = Re$ $c_0 = -0.4156, c_1 = 7.98$ (Re < 80000) $d_0 = y/b_e, d_1 = Re, c_0 = -9.98$ $c_1 = 6.42$ (Re > 80000) $d_0 = Re, d_1 = y/b$ $c_0 = 5.9, c_1 = -5.57$
Scenario 2	$n = \left( (d_0^2 - 2d_0) \frac{d_0^2}{c_0 d_0} \right)^3 + d_0^{12} (d_0^2 - \sqrt{d_0}) + d_0^{5/4}$	$d_0 = Re, c_0 = 1.413$
Scenario 3	$n = (d_0^6 \times d_1^{0.3})^2 + \left( \sqrt[3]{d_1 - d_2^2} \right) + d_2 \times d_0^5$	$d_1 = Re, d_0 = R/D_{50},$ $d_2 = \frac{Vy}{\sqrt{g(s-1)D_{50}^3}}$
Dune height	$\frac{H}{y} = (d_2^3 - d_0) + \frac{c_1^3 \left( \frac{d_1}{d_2} \right)}{2d_2^2 \times c_0} + d_0(d_0 - d_2) - d_0 \times d_2^4$ $\frac{H}{L} = d_2 \left( d_4 + (d_4 d_3)^3 (c_0 + d_1) \right) + d_1 + (\sqrt{c_1} \times d_2 - d_1) d_1$ $+ (d_0 d_3 - d_2^2)(d_1 - d_2 + \sqrt{d_0})$	$d_2 = \frac{V}{\sqrt{gD_{50}}}, c_0 = -5.053,$ $d_0 = y/b, d_1 = Re, c_1 = 0.2088$ $d_2 = R_b, c_1 = 1.478, c_0 = -3.29$ $d_1 = R/D_{50}, d_0 = y/b$ $d_4 = \frac{V}{\sqrt{g^* D_{50}(s-1)}}, d_3 = Re$

<sup>a</sup>Here the Manning's coefficient is given in the normalized form with respect to the used data.  $x_n = \alpha + \beta \times \left( \frac{x - x_{min}}{x_{max} - x_{min}} \right)$  where  $\alpha$  and  $\beta$  are 0.1 and 0.9 respectively.

presented. It can be seen that the model incidentally underestimates the high values of relative dune height. It is noteworthy to mention that other models listed in Table 8 yielded similar results for the high H/y values.

### General assessment of model performance

The key point of gene expression programming is that it is able to give the explicit expression of the relationship between the variables. It should be noted that in each run of the GEP method a unique formula will be obtained. Therefore, every model was run several times using different settings of GEP parameters and the best models (with minimum error) were selected. The mathematical expressions of the GEP best models for all cases are presented in Table 9. In this table, the  $d$  parameters represent the input parameters and  $c$  parameters are constant values which are determined by GEP. In this study, two constants per gene were selected for each model. Looking into Table 9, the equation of scenario 3 model is not only the best performing equation, but also the simplest relation compared to the other equations. Furthermore, it does not include any constants.

### Sensitivity analysis

To investigate the impacts of different parameters of the GEP-best performing models on the Manning's coefficient, sensitivity analysis was performed. The significance of each parameter was evaluated by eliminating the parameter. Table 10 reveals that in the prediction of the Manning's coefficient with flow characteristics, the Reynolds number is the most efficient parameter. This also applies to the prediction of bedform and sediment feature characteristics given that the best performing model has  $Re$  as the only input parameter. Finally, in predicting the relative dune height for both dimensionless forms  $H/y$  and  $H/L$ ,  $V/\sqrt{g(s-1)D_{50}}$  is the dominant parameter.

### Comparison of nonlinear equations with GEP-best performing model

The experimental data presented in Table 2 was used to evaluate the applicability of several existing nonlinear equations for the Manning's coefficient in the channels with dune bedform. The overall performance of each

**Table 10** | Relative significance of each of input parameter of the GEP-best models for test sets

GEP-best model	Eliminated variable	Performance criteria			
		R	DC	RMSE	MAPE
Scenario 1					
$Re, y/b$	$Re$	0.7695	0.3154	0.0058	22.9
	$y/b$	0.8025	0.5304	0.0049	19.06
Scenario 3					
$Re, R/D_{50}, \frac{Vy}{\sqrt{g(s-1)D_{50}^3}}$	$Re$	0.8253	0.5485	0.0046	22.19
	$R/D_{50}$	0.8287	0.6388	0.0036	16.06
	$\frac{Vy}{\sqrt{g(s-1)D_{50}^3}}$	0.8492	0.6342	0.0038	18.32
Relative dune height H/L					
$Re, y/b, \frac{V}{\sqrt{g(s-1)D_{50}}}$	$Re$	0.7116	0.4021	0.1381	32.88
	$y/b$	0.6821	0.3924	0.1408	33.42
	$\frac{V}{\sqrt{g(s-1)D_{50}}}$	0.6112	0.3585	0.1414	35.83
Relative dune height H/y					
$Re, y/b, \frac{V}{\sqrt{g(s-1)D_{50}}}$	$\frac{V}{\sqrt{g(s-1)D_{50}}}$	0.7025	0.3095	0.0235	12.98
	$y/b$	0.7930	0.3589	0.0230	11.48
	$Re$	0.8179	0.4042	0.0228	11.08



equation is presented graphically in a plot of observed Manning's coefficient against their computed values. Four evaluation criteria (R, DC, RMSE, and MAPE) were used as indications of the accuracy of the equations. Then a comparison was performed among the GEP-best performing

model and the equations. The nonlinear equations were divided in four groups: those that considered the roughness coefficient as dependent variable, methods based on shear stresses, equations based on the mean velocity, and the energy approach. The results of the comparisons are plotted

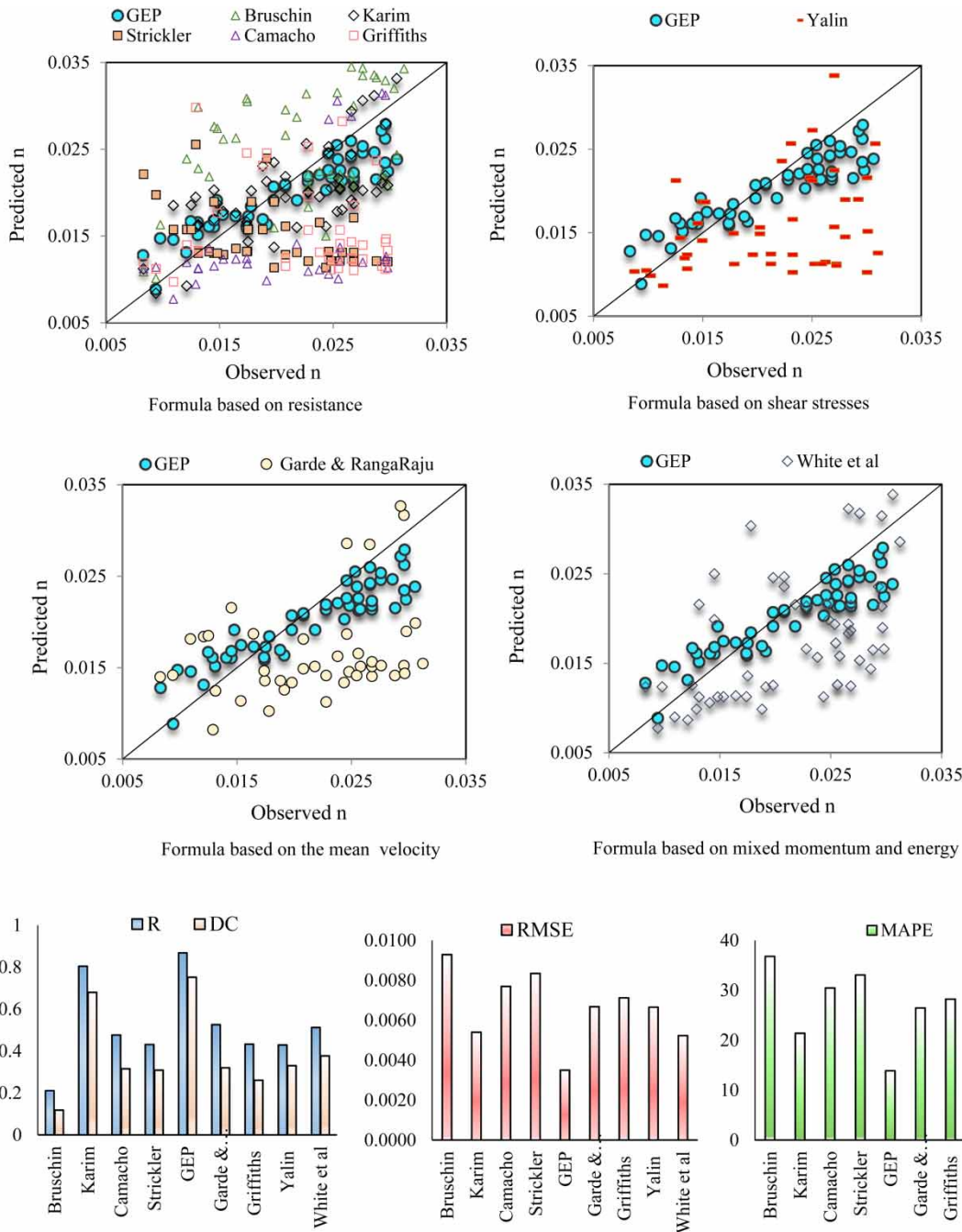


Figure 10 | Comparison between GEP method and semi-empirical equations of Manning's coefficient.

in Figure 10. From the obtained results among nonlinear equations, Karim's formula that considered the roughness coefficient as a dependent variable predicted the roughness coefficient better in comparison to the others. It should be noted that in the Karim equation  $D_{50}$ ,  $f$  and  $f_0$  parameters were used as input variables. The parameters  $f$  and  $f_0$  were taken from the measurements (Table 3), therefore, this equation presented good performance in comparison to the other equations. However, in most cases, the outcome of the equations did not provide reasonable results. On the other hand, the obtained results by GEP-best performing model were close to the measured data and were in good agreement. It should be noted that the existing equations are developed based on special flow conditions, therefore, their application is limited to those special cases of their development and they did not show uniform results under different conditions. The above-mentioned issue can be seen in Figure 10, which shows that the obtained results from equations differ from each other, and have less accuracy than GEP model which, due to its multigenic nature, allows the evolution of more complex programs composed of several subprograms. Based on Figure 10, GEP had the highest R and DC and the lowest RMSE and MAPE. This confirms the workability of GEP as an efficient approach in the modeling of the Manning's coefficient in channels with dune bedform.

## CONCLUSIONS

In the present study, the performance of nonlinear approaches as semi-empirical equations and GEP model were evaluated in order to predict the Manning's coefficient in open channels with dune bedform. In order to develop GEP models, three scenarios were considered and the impact of different modeling for the estimation of the roughness coefficient was evaluated. Then, the outputs from the proposed GEP methods were compared to the nonlinear approaches. Results verified the capability of GEP in predicting the Manning's coefficient and it was found that for prediction of the roughness coefficient in channels with dune bedform, the third scenario that took the advantages of flow and sediment characteristics performed more successfully than the others. The obtained results revealed the

accurate prediction via the GEP approach in comparison with nonlinear equations. According to the outcome, when predicting the Manning's coefficient by using flow characteristics, including the ratio of flow depth to channel width ( $y/b$ ) together with the Reynolds number ( $Re$ ), there was an improvement in the results compared to the other parameters ( $Fr$ ,  $Re$  and  $y/b$ ) when they were used as the individual input variables. Also, it was deduced that the prediction of the Manning's coefficient for Reynolds values of less than 80,000 was more precise than in those with higher Reynolds values. It was found that in the prediction of the Manning's coefficient using flow and bedform characteristics together, the single input of the Reynolds number was the highest influential parameter and plays a key role in the process. Regarding the Manning's coefficient with flow and sediment characteristics,  $Re$ ,  $R/D_{50}$  and  $Vy/\sqrt{g(s-1)D_{50}^3}$  played the most influential parameter role. Also, the results showed that using  $Re$  and  $R/D_{50}$  as the individual input variables shows almost equal performance as the three combination. Since velocity measurements in field situations are typically very difficult, the use of  $R/D_{50}$  seems to be a more practical option. In the estimation of the relative dune height ( $H/y$  or  $H/L$ ), using the  $Re$ ,  $y/b$  and  $(v/\sqrt{g(s-1)D_{50}^3})$  produced improved GEP results. According to sensitivity analysis, the Reynolds number ( $Re$ ) played the most important role in predicting the Manning's coefficient in channels with dune bedform. It should, however, be noted that the GEP is a data-driven model and the GEP-based models are data sensitive, so further studies using data ranges out of this study and preferably field data should be used to find the merits of the models to predict roughness coefficient in the real condition of flow.

## REFERENCES

- Aytek, A. & Kisi, O. 2008 A genetic programming approach to suspended sediment modeling. *J. Hydrol.* **351**, 288–298.
- Azamathulla, M., Ghani, A., Zakaria, A. & Guven, A. 2010 Genetic programming to predict bridge pier scour. *J. Hydraul. Eng. ASCE* **136**, 165–169.
- Bagnold, R. A. 1966 *An Approach to the Sediment Transport Problem From General Physics*. US Geological Survey, Professional Paper 422-I, Washington, DC, USA, 37 pp.

- Banzhaf, W., Nordin, P., Keller, P. E. & Francone, F. D. 1998 *Genetic Programming*. Morgan Kaufmann, San Francisco, CA, p. 512.
- Bhattacharya, B., Price, R. K. & Solomatine, D. P. 2007 *Machine Learning Approach to Modeling Sediment Transport*. *J. Hydraul. Eng.* **133** (4), DOI: 10.1061/(ASCE)0733-9429 (2007)133:4(440).
- Bolat, B. & Yildirim, T. 2004 A data selection method for probabilistic neural networks. *J. Electr. Electron. Eng.* **4** (2), 1137–1140.
- Bruschin, J. 1985 *Discussion on Brownlie (1983): flow depth in sand-bed channels*. *J. Hydraul. Eng. ASCE* **111**, 736–739.
- Camacho, R. & Yen, B. C. 1989 Nonlinear Resistance Relationships for Alluvial channels. In: *Proceedings of the International Conference on Channel Flow and Catchment Runoff* (B. C. Yen, ed.). University of Virginia, Charlottesville, pp. 392–399.
- Chang, C. K., Azmathulla, H. M., Zakaria, N. A. & Ghani, A. A. 2012 *Appraisal of soft computing techniques in prediction of total bed material load in tropical rivers*. *J. Earth Syst. Sci.* **12** (1), 125–133.
- Dawson, W. C. & Wilby, R. 1998 *An artificial neural network approach to rainfall-runoff modelling*. *Hydrol. Sci. J.* **43** (1), 47–66.
- Einstein, H. A. & Barbarossa, N. L. 1952 River channel roughness. *Trans. ASCE* **117**, 1121–1132.
- Engel, P. & Lau, Y. L. 1980 *Friction factor for two-dimensional dune roughness*. *J. Hydraul. Res.* **18** (3), 213–225.
- Ferreria, C. 2001 *Gene expression programming: a new adaptive algorithm for solving problems*. *Complex Syst.* **13** (2), 87–129.
- Garde, R. J. & Ranga Raju, K. G. 1966 Resistance relationships for alluvial channel flow. *J. Hydraul. Eng.-ASCE*. **92** (4), 77–100.
- Griffiths, G. A. 1981 Flow resistance in coarse gravel bed rivers. *J. Hydraul. Div. ASCE HY* **107**, 899–918.
- Guy, H. P., Simons, D. B. & Richardson, E. V. 1966 *Summary of Alluvial Channel Data from Flume Experiments*. US Government Printing Office. 1956–61 (No. 462-1).
- Heydari, H., Zarrati, A. R. & Karimae Tabarestani, M. 2014 *Bed form characteristics in a live bed alluvial channel*. *Sci. Iran. Trans. A Civil Eng.* **21** (6), 1773–1780.
- Julien, P. Y. 2010 *Erosion and Sedimentation*. Cambridge University Press, New York, p. 181.
- Karim, F. 1995 *Bed configuration and hydraulic resistance in alluvial-channel flows*. *ASCE J. Hydraul. Eng.* **121** (1), 15–25.
- Karim, F. 1999 *Bed-form geometry in sand-bed flows*. *ASCE J. Hydraul. Eng.* **125** (12), 1253–1261.
- Kisi, O., Shiri, J. & Tombul, M. 2013 *Modeling rain fall-runoff process using soft computing techniques*. *Comput. Geosci.* **51**, 108–117.
- Legates, D. R. & McCabe, G. J. 1999 *Evaluating the use of 'goodness of fit' measures in hydrologic and hydroclimatic model validation*. *Water Resour. Res.* **35** (1), 233–241.
- Meyer-Peter, E. & Mueller, R. 1948 *Formulas for Bed Load Transport*. In: *IAHSR 2nd Meeting*, Stockholm, appendix 2. IAHR Research, 2nd meeting, No. 2.
- Nourani, V., Alizadeh, F. & Roushangar, K. 2016 *Evaluation of a two-stage SVM and spatial statistics methods for modeling monthly river suspended sediment load*. *Water Resour. Manag.* **30**, 393–407.
- Raudkivi, A. J. 1967 *Analysis of resistance in fluvial channels*. *Proc. ASCE* **93**, 2084–2093.
- Richardson, E. V. & Simons, D. B. 1967 *Resistance to Flow in Sand Channels*. In: *Proc. 12th Congress IAHR*, Vol. 1. Association of Hydraulic Research, Fort Collins, CO.
- 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* **2**, 537–551.
- Roushangar, K. & Koosheh, A. 2015 *Evaluation of GA-SVR method for modeling bedload transport in gravel-bed rivers*. *J. Hydrol.* **527**, 1142–1152.
- Roushangar, K., Vojoudi, F. & Shiri, J. 2014a *Modeling river total bed material load discharge using artificial intelligence approaches (based on conceptual inputs)*. *J. Hydrol.* **514**, 114–122.
- Roushangar, K., Akhgar, S., Salmasi, F. & Shiri, J. 2014b *Modeling energy dissipation over stepped spillways using machine learning approaches*. *J. Hydrol.* **508**, 254–265.
- Roushangar, K., Garekhani, S. & Alizadeh, F. 2016a *Forecasting daily seepage discharge of an earth dam using wavelet-Mutual information Gaussian process regression approaches*. *Geotech. Geol. Eng.* **5**, 1313–1326.
- Roushangar, K., Akhgar, S., Erfan, A. & Shiri, J. 2016b *Modeling scour depth downstream of grade-control structures using data driven and empirical approaches*. *J. Hydroinform.* **18** (6), 946–960.
- Samui, P. & Dixon, B. 2011 *Application of support vector machine and relevance vector machine to determine evaporative losses in reservoirs*. *Hydrol. Process.* **9**, 1361–1369.
- Simons, D. B. & Richardson, E. V. 1963 *Form of bed roughness in alluvial channels*. *Trans. ASCE* **128**, 284–323.
- Simons, D. B. & Richardson, E. V. 1966 *Resistance to Flow in Alluvial Channels*. Professional Paper 422-J. US Geological Survey, Washington, DC.
- Smith, K. V. C. 1968 *Alluvial channel resistance related to bedforms*. *Proc. ASCE* **94** (1), 59–70.
- Strickler, A. 1923 *mitteilungen des EidgenossischenAmtes fur Wasserwirtschaft* 16, Bern, Switzerland, Translated as "Contributions to the question of a velocity formula and roughness data for streams, channels and closed pipelines." by T. Roesgan and W. R. Brownie, Translation T-10, W. M. Keck Laboratory of Hydraulics and Water Resources, Calif. Inst. Tech., Pasadena, Calif. Jan. <http://arxiv.org/abs/physics/0507066>.
- Sun, A. Y., Wang, D. & Xu, X. 2014 *Monthly streamflow forecasting using Gaussian process regression*. *J. Hydrol.* **511**, 72–81.
- Taylor Jr., R. H. & Brooks, V. N. 1962 *Discussion of resistance to flow in alluvial channels*, by Simons, D. B. & Richardson, E. V. *Trans. ASCE* **127** (1), 982–992.

- Teodorescu, L. & Sherwood, D. 2008 [High energy physics event selection with gene expression programming](#). *Comput. Phys. Commun.* **178** (6), 409–419.
- Van der Mark, C. F., Blom, A. & Hulscher, S. J. 2008 [Quantification of variability in bedform geometry](#). *J. Geophys. Res.* **113**, 1–11.
- Van Rijn, L. C. 1984 [Sediment transport, part III: bedforms and alluvial roughness](#). *ASCE J. Hydraul. Eng.* **110** (12), 1733–1754.
- White, W. R., Paris, E., Bettess, R. & Wang, S. 1987 Frictional characteristics of alluvial streams in the lower and upper regimes. *Proc. Inst. Civil Eng. (London)* **83** (2), 685–700.
- Williams, G. P. 1970 *Flume Width and Water Depth Effects in Sediment Transport Experiments*. US Geological Survey, US Government Printing Office, Professional Paper 562-H.
- Yalin, M. S. 1977 *Mechanics of Sediment Transport*, 2nd edn. Pergamon, Oxford, UK.
- Yang, S. Q., Tan, S. K. & Lim, S. Y. 2005 [Flow resistance and bedform geometry in a wide alluvial channel](#). *Water Resour. Res.* **41** (9), 1–8.
- Yen, B. C. 2002 [Open channel flow resistance](#). *J. Hydraul. Eng.* **128** (1), 20–39.
- Yen, B. C. 2004 *Channel Flow Resistance, Centennial of Manning's Formula*. Water Resources Publications, Littleton, Colorado, pp. 1–135.

First received 25 January 2017; accepted in revised form 21 December 2017. Available online 24 January 2018