# Prediction of sediment transport rates in gravel-bed rivers using Gaussian process regression

Kiyoumars Roushangar and Saman Shahnazi

## ABSTRACT

Estimating sediment transport rate in rivers has high importance due to the difficulties and costs associated with its measurement, which has drawn the attention of experts in water engineering. In this study, Gaussian process regression (GPR) is applied to predict the sediment transport rate for 19 gravel-bed rivers in the United States. To compare the performance of GPR, the support vector machine (SVM) as a common type of kernel-based models was developed. Model inputs of sediment transport were prepared based on two scenarios: the first scenario considers only hydraulic characteristics and the second scenario was formed using hydraulic and sediment properties. Obtained results revealed that the GPR models present better performance compared to the SVM models and other empirical sediment transport formulas. Also, it was found that incorporating the second scenario as input led to better predictions. In addition, performing sensitivity analysis showed that the ratio of average velocity to shear flow velocity is the most effective parameter in predicting the sediment transport rate of gravel-bed rivers.

**Key words** | bed load, empirical methods, Gaussian process regression, sediment transport, support vector machine, total sediment load

**Kiyoumars Roushangar** (corresponding author)
**Saman Shahnazi**
Center of Excellence in Hydroinformatics,
University of Tabriz,
29 Bahman Ave., Tabriz,
Iran
E-mail: *kroshangar@yahoo.com*

## INTRODUCTION

For over half a century, there have been continuing efforts to enhance the understanding of sediment transport process. Many investigations proposed empirical relationships for predicting sediment transport of alluvial rivers (Yang *et al.* 2009). Despite an extensive application of empirical formulas, the prediction error of these formulas is reported to be extremely high (Barry *et al.* 2004; Bathurst 2007; Roushangar *et al.* 2014). The transport rate of sediment load which is carried by surface flow, has a major role in controlling river ecosystems and is one of the main parameters in design, implementation and operation of hydraulic structures, irrigation, transfer and treatment of water, watershed management, and flood control. Due to the importance of sediment transport phenomenon, extensive research during recent years has been done using artificial intelligence (AI) methods. Among them,

significant applications of artificial neural networks (ANN) method in sediment transport rate estimation have been reported in the literature (Bhattacharya *et al.* 2007; Doğan *et al.* 2007; Sasal *et al.* 2009; Yang *et al.* 2009; Kumar 2012). Remarkable performance of AI methods has been the motive of hydraulic and river engineers to develop more effective techniques with greater generalizability. In this way, Azamathulla *et al.* (2009) suggested the adaptive neuro-fuzzy inference system (ANFIS) method as a flexible and more optimum technique for predicting bed load. Azamathulla *et al.* (2010) conducted research on a case study sediment load prediction and demonstrated the encouraging performance of the support vector machine (SVM) for prediction of sediment load. Ghani & Azamathulla (2014) offered gene expression programming (GEP) for the development of functional relationship for total sediment

load in three Malaysian rivers. Having utilized ANFIS with GEP to model total bed material load of Qotur River, Roushangar et al. (2014) showed that the models based on stream power approach are more reliable than those that are based on shear stress approach. Okcu et al. (2016) applied polynomial best subset regression (PBSR) to a database containing both river and flume measurements and developed a new equation for predicting total sediment load. Kitsikoudis et al. (2014) found that ANN and ANFIS surpass the symbolic regression (SR) in terms of bed load prediction in gravel-bed rivers. Roushangar & Koosheh (2015) introduced a hybrid method based on support vector regression (SVR) coupled with genetic algorithm (GA) for quantification of bed load transport rate in three gravel-bed rivers. Their hybrid model enjoyed greater accuracy when it came to predict low transport rate. Sahraei et al. (2017) introduced a useful prediction method based on least square support vector regression (LSSVR) with particle swarm optimization (PSO) for the purpose of predicting total sediment load.

Gaussian process regression (GPR) is a newly developed learning approach that works based on the concept of kernel functions. GPR presents probabilistic models, which means that Gaussian process provides a reliability of responses to the given input data (Yuan et al. 2008). In addition, the GPR method is flexible as it has an ability to handle nonlinear problems and also non-parametric as it does not need parameter selection. Some previous studies have used GPR as a probabilistic stream flow forecaster (Sun et al. 2014; Zhu et al. 2018). In addition, promising application of GPR in forecasting daily seepage discharge of an earth dam (Roushangar et al. 2016), prediction of stream water temperature (Grbić et al. 2013), and prediction of urban water consumption (Roushangar & Alizadeh 2018) have been reported in the literature.

A detailed literature review demonstrated that although some research has been conducted on GPR, none considered prediction of sediment load as a method applicable to a wide range of flow and sediment characteristics. The present study aims to investigate the capability of GPR in predicting the bed load and total load of gravel-bed rivers. An extensive database compiled from 19 gravel-bed rivers (King et al. 2004) was used to feed the utilized

GPR models. Moreover, since the SVM is closely related to the employed GPR approach in terms of using kernel functions, the performance of the employed GPR approach was compared with SVM-based regression. Optimum input combination and the most important parameters in predicting sediment transport rate are determined using sensitivity analysis.

## MATERIALS AND METHODS

### Study area and used data

The present study covers 19 gravel-bed rivers, information for which was collected by the US Forest Service in cooperation with other agencies. This database has become a robust source for engineers and researchers working on sediment transport (Recking 2010; Schneider et al. 2015). This dataset includes bed load, suspended load, and hydraulic measurements of gravel-bed rivers, while the additional information regarding this dataset and details of the methods used to measure the various types of data are presented in King et al. (2004). Parallel measurements of suspended load and bed load from 19 streams within the Snake River basin with a range of discharge between 0.05 m³/s and 30 m³/s and varied hydraulic and sediment properties were selected. The notable point is, that on all sites, the diameters $d_{50}$ and $d_{90}$ of the surface material were larger than those of the subsurface material, indicating the presence of an armor layer which is the main characteristic of gravel-bed rivers. An armor layer establishes a stable boundary in low flows, but forms a complicated hydraulic condition in floods due to sudden scouring of finer sub-surface material (Wang & Liu 2009). Some characteristics of the selected rivers are presented in Table 1.

### Gaussian process regression

A Gaussian process (GP) is a set of random variables; any finite set of these random variables has a multivariate Gaussian distribution. Let $\chi \times \gamma$ represent the domains of inputs and outputs, respectively, from which n pairs $(x_i, y_i)$ are drawn independently and identically distributed. For regression, adjudge that $y \subseteq Re$; then, a GP on $\chi$ is specified

**Table 1** | Characteristics of the selected rivers

| Rivers | Drainage area (km$^2$) | Data for training | Data for testing | Total data | Slope m/m | $d_{50,sur}$ (mm) | Data of sampling | Range of discharge (m$^3$/s) |
|---|---|---|---|---|---|---|---|---|
| Big Wood River | 349.7 | 17 | 8 | 26 | 0.0091 | 119 | 1999–2000 | 9.6–30.8 |
| Bruneaul River | 989 | 18 | 9 | 27 | 0.0054 | 27 | 1998–2002 | 4.7–20.9 |
| Fourth Of July | 44.28 | 17 | 8 | 25 | 0.0202 | 51 | 1994–1995 | 0.2–3.8 |
| Herd Creek | 292.6 | 15 | 7 | 22 | 0.0077 | 67 | 1994–1995 | 0.5–8.1 |
| Jarbidge River | 79.25 | 18 | 9 | 26 | 0.0160 | 89 | 1998–2002 | 1.4–8 |
| Johns Creek | 293.1 | 14 | 7 | 22 | 0.207 | 199.2 | 1986–1995 | 0.97–26 |
| Little Slate Creek | 168.5 | 55 | 24 | 79 | 0.0268 | 98.1 | 1986–1997 | 0.52–15.7 |
| Lolo Creek | 107.7 | 28 | 13 | 41 | 0.0097 | 67 | 1980–1997 | 1.8–16.2 |
| Main Fork Red River | 129.3 | 77 | 33 | 110 | 0.0059 | 50.5 | 1986–1999 | 0.29–18.2 |
| Marsh Creek | 191.5 | 18 | 9 | 27 | 0.0060 | 57 | 1994–1995 | 3.36–23.2 |
| Rapid River | 279.5 | 50 | 22 | 72 | 0.0108 | 61.8 | 1986–2000 | 0.91–36.8 |
| South Fork Red River | 97.8 | 67 | 30 | 97 | 0.0146 | 105.7 | 1986–1999 | 0.2–11 |
| South Fork Salmon River | 853.6 | 35 | 16 | 51 | 0.0025 | 35 | 1985–1997 | 3.8–124.3 |
| Squaw Creek (USGS)[a] | 192 | 22 | 10 | 32 | 0.0100 | 46.6 | 1994–1995 | 0.4–7.5 |
| Thompson Creek | 58.1 | 16 | 8 | 24 | 0.0153 | 67.1 | 1994–1995 | 0.4–3.5 |
| Trapper Creek | 22.2 | 60 | 27 | 87 | 0.0414 | 86.1 | 1985–1997 | 0.05–2.8 |
| Hawley Creek | 104.8 | 45 | 20 | 65 | 0.0233 | 40 | 1990–1996 | 0.27–2.6 |
| Salmon River near Obsidian | 243.9 | 14 | 6 | 19 | 0.0066 | 61.8 | 1990 | 11.44–20.9 |
| Squaw Creek (USFS)[b] | 37.6 | 26 | 12 | 38 | 0.0240 | 23 | 1990–1996 | 0.18–1.5 |

[a]USGS: United States Geological Survey.
[b]USFS: United States Forest Service.

by a mean function $\mu: \chi \to Re$. In GPR calculations of output variable, y is defined by $y = f(x) + \xi$, where $\xi \sim N(0, \sigma^2)$. The symbol $\sim$ in statistics means sampling for. In GPR, 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 study, it is assumed that the observational error $\xi$ is normally independent and identically distributed, with a mean value of zero ($\mu(x) = 0$), a variance of $\sigma^2$ and $f(x)$ drawn from the Gaussian process on $\chi$ specified by k. That is:

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

where $K_{ij} = K(x_i, x_j)$, and $I$ is the identity matrix. For a given vector of the test data $X^*$, the predictive distribution of the corresponding output $Y_*/(X, Y)$, $X_* \sim N(\mu, \sum)$ is

Gaussian, where:

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

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

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 dataset, and this is similarly true for the other value of $K(X, X)$, $K(X_*, X)$, $X$ and $K(X_*, X_*)$; here, $X$ and $Y$ are the vector of the training data and training data labels $y_i$.

A specified covariance function is essential to produce a positive semi-definite covariance matrix K, where $K_{ij} = K(x_i, x_j)$. The term kernel function used in SVM is

synonymous with the covariance function applied in GPR. With the known kernel function and degree of noise $\sigma^2$, Equations (1) and (2) would be enough for derivation. The user needs to tune covariance function and its parameters and the degree of noise suitably during the training process of GPR models. In the case of GPR with a fixed value of Gaussian noise, a GP model could be trained by applying Bayesian inference, i.e., by maximizing the marginal likelihood. This leads to the minimization of the negative log-posterior:

$$p(\sigma^2, k) = \frac{1}{2} Y^T (K + \sigma^2 I)^{-1} Y + \frac{1}{2} log[K + \sigma^2 I] \\ - logp(\sigma^2) - logp(k) \tag{3}$$

To acquire the hyperparameters, the partial derivative of Equation (3) can be obtained with respect to $\sigma^2$ and k, and minimization can be obtained by gradient descent. For more information about GPR and different covariance functions, readers are referred to Kuss (2006).

## Support vector machine

Many studies have been carried out in various fields of engineering by using SVM. Therefore, only a brief summary of the employed SVM model is presented here. It is assumed that for dataset $\{x_i, y_i\}$, SVM equations founded on Vapnik theory (Vapnik 1998) approximate the function as:

$$f(x) = W\varphi(x) + b \tag{4}$$

where $\varphi(x)$ represents a nonlinear function in feature of input $x$, $W$ vector is known as the weight factor, $b$ is known as the bias. These coefficients are predicted by minimizing regularized risk function as shown below:

$$R_{SVM_s}(C) = \frac{1}{2} \|w\|^2 + C \frac{1}{2} \sum_{i=1}^{n} L_\varepsilon(t_i, y_i) \tag{5}$$

where

$$L_\varepsilon(t_i, y_i) = \begin{cases} 0 & |t_i - y_i| \leq \varepsilon \\ |t_i - y_i| - \varepsilon & \textbf{otherwise} \end{cases} \tag{6}$$

The constant C is the cost factor, $\frac{1}{2}\|w\|^2$ stands for the regularization term, $\varepsilon$ is the radius of the tube within which the regression function must lie, $n$ is the number of elements and $L_\varepsilon(t_i, y_i)$ denotes the loss function in which $y_i$ is forecasted value and $t_i$ stands for desired value in period $i$. The parameters $w$ and $b$ are estimated by minimization process of the regularized risk function after introducing positive slack variables $\xi_i$ and $\xi_i^*$ that express upper and lower excess deviation.

$$\textbf{Minimize } R_{SVM_s}(w, \xi^*, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} (\xi_i, \xi_i^*) \tag{7}$$

$$t_i - W_i\varphi(X_i) - b \leq \varepsilon + \xi_i, \; W_i\varphi(X_i) + b - t_i \leq \varepsilon + \xi_i^*, \\ \xi_i + \xi_i^* \geq 0$$

Equation (4) can be solved by introducing Lagrange multiplier and optimality constraints, therefore obtaining a general form of function given by:

$$f(x) = \sum_{i=1}^{n} (\beta_i - \beta_i^*) K(x_i, x_j) + b \tag{8}$$

where $\beta_i$ and $\beta_i^*$ are Lagrange multipliers, $K(x_i, x_j) = \varphi(x_i)\varphi(x_j)$ and the term $K(x_i, x_j)$ refers to the kernel function, which is an inner product of two vectors $x_i$ and $x_j$ in the feature space $\varphi(x_i)$ and $\varphi(x_j)$, respectively. Kernel functions map data into a high dimension feature space so that linear machine computational power can be increased. Kernel functions also allow the extension of linear hypotheses into nonlinear which can indirectly be achieved.

It is observed that the defined equation in SVM (Equation (4)) is similar to Gaussian process formulation for regression. In fact, GPR is inspired by SVM's structure and formulation and both approaches introduce two different but equivalent perspectives for regression by application of a function $f(.)$ directly to the input data points. Put differently, in the GPR case, data were generated with Gaussian white noise around the function $f$, but in the case of SVM, ε-insensitive error function can be considered as a non-Gaussian likelihood or noise model.

Among the various kernel functions presented in Table 2, the RBF kernel is reported to perform better than other kernel functions and was used in the presented

**Table 2** | Kernel functions

| Kernel type | Function | Kernel parameter |
|---|---|---|
| Linear | $\mathbf{K}(\mathbf{x_i}, \mathbf{x_j}) = (\mathbf{x_i}, \mathbf{x_j})$ | – |
| Polynomial | $\mathbf{K}(\mathbf{x_i}, \mathbf{x_j}) = ((\mathbf{x_i}, \mathbf{x_j}) + 1)^{\mathbf{d}}$ | d |
| RBF | $\mathbf{K}(\mathbf{x_i}, \mathbf{x_j}) = \mathbf{exp}(- \parallel \mathbf{x_i}, \mathbf{x_j} \parallel^2 / 2\gamma^2)$ | $\gamma$ |
| Sigmoid | $\mathbf{K}(\mathbf{x_i}, \mathbf{x_j}) = \mathbf{tanh}(-\boldsymbol{\alpha}(\mathbf{x_i}, \mathbf{x_j}) + \mathbf{c})$ | $\boldsymbol{\alpha}$, c |

study (Pal *et al.* 2014; Azamathulla *et al.* 2016; Komasi *et al.* 2018).

## Experimental setup

In the present study, the prediction process was codified in MATLAB® environment. Regression learning toolbox, which is available for free, was partly used for prediction of sediment transport rate. Setting of the optimal values of capacity constant (C), the size of error-intensive zone ($\varepsilon$), Gaussian noise and, most importantly, the kernel parameter ($\gamma$) are some of the issues which considerably affect the proposed modeling process. The optimum values of Gaussian noise and kernel parameter ($\gamma$) were obtained after trial-and-error process. Furthermore, optimization of parameters C and $\varepsilon$ has been carried out by a systematic grid search of the parameters using cross-validation on the training dimensionless measures.

Typically, importation of data in non-normalized form reduces the speed and accuracy of the network and leads to undesirable results. In this study, all input variables were normalized in the range of 0.1–0.9 by the following equation:

$$\mathbf{x_{norm}} = 0.1 + 0.9 \times \left( \frac{\mathbf{x_i} - \mathbf{x_{imin}}}{\mathbf{x_{imax}} - \mathbf{x_{imin}}} \right) \qquad (9)$$

where $\mathbf{x_{norm}}$, $\mathbf{x_i}$, $\mathbf{x_{imax}}$, $\mathbf{x_{imin}}$, respectively, are: the normalized value of variable $\mathbf{x_i}$, the original value, the maximum and minimum of variable $\mathbf{x_i}$. Selecting input variables is the most important step of modeling through all machine learning methods and can affect the accuracy of the results. Considering that the data compilation includes datasets coming from various streams and rivers, for all cases, 75% of data from each river was divided for training the model and the remaining 25% was used for test purposes. As a

result, there are 612 measurements for training and 278 measurements for testing. Determination of optimal input parameters for the machine learning approaches was made by means of trial-and-error procedure through a set of dimensionless variables. In this study, in order to determine appropriate inputs, different dimensionless parameters were defined as follows (Sasal *et al.* 2009; Okcu *et al.* 2016):

$$\mathbf{Fr}, \theta, \mathbf{R}_h / \mathbf{d}_{50}, \mathbf{y}/\mathbf{B}, \mathbf{V}/\mathbf{U}_*, \mathbf{S}_0, \mathbf{D}_*, \mathbf{T}, \mathbf{V}^3 / \mathbf{g}.\mathbf{y}.\boldsymbol{\omega}_s$$

where $\boldsymbol{R}_h$ is the hydraulic radius, $\boldsymbol{d}_{50}$ the median bed material particle diameter, $y$ the average flow depth, $B$ the width of channel, $V$ the average flow velocity, $\boldsymbol{U}_*$ the shear velocity. The mentioned dimensionless parameters are independent variables for section-average sediment transport to be explored with the employed GPR and SVM methods: $\boldsymbol{Fr}$ is the Froude number ($V/\sqrt{g.y}$), $\boldsymbol{\theta}$ stands for particle mobility parameter ($\theta = \mathbf{U}_*^2 / (\mathbf{G_s} - 1)\mathbf{g}\mathbf{d}_{50}$) where $G_s$ is the specific sediment density and $g$ the acceleration due to gravity, $\mathbf{R_h}/\mathbf{d}_{50}$ is representative of the channel roughness and flow resistance, $\boldsymbol{y}/\boldsymbol{B}$ refers to the dimensionless width; $\boldsymbol{V}/\boldsymbol{U}_*$ is the ratio of average velocity to shear velocity, $\boldsymbol{S}_0$ is the bed slope of the channel, $\mathbf{D}_*$ is dimensionless particle parameter which is defined as:

$$\mathbf{D}_* = \mathbf{d}_{50} \left[ \frac{(\boldsymbol{G_s} - 1)\boldsymbol{g}}{\boldsymbol{v}^2} \right]^{\frac{1}{3}} \qquad (10)$$

where $\boldsymbol{v}$ is the kinematic viscosity. Transport stage parameter $T$ is defined as:

$$T = \frac{\theta' - \theta_{cr}}{\theta_{cr}} \qquad (11)$$

$$C' = 18 log \left( \frac{4y}{\boldsymbol{d}_{90}} \right) \qquad (12)$$

$$\theta' = \frac{V^2}{(\boldsymbol{G_s} - 1)\boldsymbol{D}_{50}\boldsymbol{C'}^2} \qquad (13)$$

where $\theta'$ refers to mobility parameter (van den Berg & van Gelder 1993), $\theta_{cr}$ denotes Shields' critical shear stress, $C'$ is the Chézy's coefficient, and $\boldsymbol{d}_{90}$ is the characteristic grain

size, in which 90% of particles show smaller sizes. $V^3/g.y.\omega_s$ is an extracted non-dimensional parameter from the Molinas and Wu formula for predicting total load which is based on the gravitational power theory of Velikanov and $\omega_s$ denotes particle fall velocity in water (Molinas & Wu 2001).

Different models were defined using the mentioned parameters and after trial-and-error procedure, the best models were selected. Table 3 sums up the selected input combination for modeling bed load and total load in two scenarios based on flow conditions and sediment properties.

## Empirical approaches

Due to the plethora of equations that have been developed for sediment transport rate, the results of Khorram & Ergil's (2010a, 2010b) studies were used in order to choose empirical approaches. They utilized above 2,000 laboratory and 700 field data for investigating the efficiency of 75 different formulas in order to predict bed load and total sediment load and introduced the most appropriate equations for sand and gravel-bed rivers separately. The selected formulas are presented in Tables 4 and 5.

## Performance criteria

In this study, correlation coefficient (R), Nash–Sutcliffe efficiency (NSE), root mean square error (RMSE) and relative error (RE), as depicted in Equations (14)–(17), were used

as statistical parameters for evaluating performance of the GPR and SVM models. The larger values of the NSE and R and smaller one of RMSE indicate the higher accuracy of the model.

$$R = \frac{\sum_{i=1}^{N} (X_i - \bar{X}) \times (Y_i - \bar{Y})}{\sum_{i=1}^{N} (X_i - \bar{X})^2 \times (Y_i - \bar{Y})^2} \tag{14}$$

$$RMSE = \sqrt{\sum_{i=1}^{N} \frac{(X_i - Y_i)^2}{N}} \tag{15}$$

$$NSE = 1 - \frac{\sum_{i=1}^{N} (X_i - Y_i)^2}{\sum_{i=1}^{N} (X_i - \bar{X})^2} \tag{16}$$

$$RE = \frac{\sum_{i=1}^{N} |(X_i - Y_i)/X_i|}{N} \tag{17}$$

where N represents the number of data, $X_i$ is the observed value, $Y_i$ is the predicted value, $\bar{X}$ and $\bar{Y}$ stand for the mean values of the observed and predicted values.

## RESULTS AND DISCUSSION

Owing to constraints of available information about sediment transport and also fewer variables of flow characteristics, scenario 1 was defined based on flow conditions. In scenario 2, for predicting bed load and total load, six models were developed based on flow conditions

**Table 3** | Input models

| Scenario 1 | | Scenario 2 | | | |
|---|---|---|---|---|---|
| **Parameters of flow conditions** | | **Parameters of flow conditions and sediment properties** | | | |
| **All states** | | **Bed load** | | **Total load** | |
| **Models** | **Input variables** | **Models** | **Input variables** | **Models** | **Input variables** |
| (I) | $\left(Fr, \dfrac{y}{B}, \dfrac{V}{U_*}\right)$ | BL(I) | $\left(Fr, \dfrac{V}{U_*}, D_*, T\right)$ | TL(I) | $\left(Fr, \dfrac{V}{U_*}, D_*, T\right)$ |
| (II) | $\left(Re, \dfrac{y}{B}, \dfrac{V}{U_*}\right)$ | BL(II) | $\left(Fr, \dfrac{V}{U_*}, \dfrac{R_h}{d_{50}}, \theta\right)$ | TL(II) | $\left(Fr, \dfrac{V}{U_*}, \theta, \dfrac{V^3}{g.y.\omega_s}\right)$ |
| (III) | $\left(Fr, \dfrac{y}{B}, \dfrac{V}{U_*}, S_0\right)$ | BL(III) | $\left(Fr, \dfrac{V}{U_*}, \dfrac{R_h}{d_{50}}, \theta, D_*, T\right)$ | TL(III) | $\left(Fr, \dfrac{V}{U_*}, \theta, \dfrac{V^3}{g.y.\omega_s}, D_*, T\right)$ |
| (IV) | $\left(Re, \dfrac{y}{B}, \dfrac{V}{U_*}, S_0\right)$ | | | | |

**Table 4** | Selected empirical formulas for predicting bed load from Khorram & Ergil (2010b)

| Formula | Name | Approach |
|---|---|---|
| $q_b = \dfrac{W^* U_*^3 \rho_s}{(G_s - 1)g} \cdot \phi_{50} = \dfrac{\theta}{\theta_{cr}} \cdot \theta = \dfrac{U_*^2}{(G_s - 1)g d_{50sub}}$ $$W^* = \begin{cases} 11.2\left(1 - \dfrac{0.822}{\phi_{50}}\right)^{4.5} \phi_{50} > 1.65 \\ 0.0025\exp[14.2(\phi_{50} - 1) - (\phi_{50} - 1)^2]\ 0.95 \leq \phi_{50} \leq 1.65 \\ 0.0025\phi_{50}^{14.2}\ \phi_{50} < 0.95 \end{cases}$$ | Parker *et al.* (1982); Pitlick *et al.* (2009) | Deterministic equal mobility method |
| $q_b = \dfrac{W_g^* U_*^3 \rho_s}{(G_s - 1)g}$ $$W_g^* = \begin{cases} 11.2\left(1 - 0.846\dfrac{\tau_{cr}}{\tau_0}\right)^{4.5} \tau_0 > \tau_{cr} \\ 0.0025\left(\dfrac{\tau_0}{\tau_{cr}}\right) \tau_0 \leq \tau_{cr} \end{cases}$$ $\theta = \dfrac{U_*^2}{(G_s - 1)g d_{50sub}} \cdot \phi = \dfrac{\theta}{\theta_{cr}}$ | Wilcock (2001); Pitlick *et al.* (2009) | Deterministic equal mobility method |
| $q_b = \gamma_s R_h V \times \left\{\left[0.667\left(\dfrac{d_{50}}{R_h}\right)^{\frac{2}{3}} + 0.14\right] - 0.778\left(\dfrac{d_{50}}{R_h}\right)^{\frac{2}{3}}\right\}^3$ | Rottner (1959); Yang (1996) | Regression method |
| $\Phi_b = 0.05\left(\dfrac{V}{U_*}\right)^2 \theta^{\frac{5}{2}}$ $\Phi_b = \dfrac{q_b}{\sqrt{(G_s - 1)g d_s^3}}$ | Engelund & Hansen (1967) | Regression method |

and sediment properties. The values of performance criteria obtained from GPR and SVM models based on scenario 1 are presented in Table 6. From the obtained results of statistical parameters (RMSE, R, NSE, and RE), it is observed that models based on scenario 1 have not been accurate enough for predicting bed load transport rate, while the estimated and observed values of scenario 1 for total load are in good agreement and it can be stated that it is possible to achieve a good approximation of total sediment load by using only hydraulic characteristics. It seems that the inability of the first scenario in predicting the bed load transport rate is due to the fact that the bed load transport is more affected by characteristics of bed layer such as median diameter of particles and using hydraulic characteristics as input parameters caused insufficient results. Conversely, suspended load (which is more affected by flow conditions) is more important in studied gravel-bed

rivers and constitutes almost 75% of the total sediment load on average. Therefore, it seems different effective hydraulic parameters are the effective factors for prediction of total sediment load. In scenario 1, the most accurate estimations correspond to model (II), in which the input parameters are: $Re$, $y/B$, and $V/U_*$. In scenario 2, different combinations of input variables were developed after a trial-and-error process according to flow conditions and sediment properties. Comparing the results between the two scenarios demonstrates the superiority of scenario 2 in quantification of bed load and total load transport rate. The results of the employed methods for models based on scenario 2 are listed in Table 7. The best models for predicting bed load and total load were, respectively, BL(II) and TL(II) with input parameters of $Fr$, $V/U_*$, $R_h/d_{50}$, and $\theta$ for bed load and $Fr$, $V/U_*$, $\theta$, and $V^3/g.y.\omega_s$ for total sediment load. According to the results presented in Table 7,

**Table 5** | Selected empirical formulas for predicting total load from Khorram & Ergil (2010a)

| Formula | Name | Approach |
|---|---|---|
| $$\frac{Q_t}{Q} = C\frac{d_{50}}{R_h}\left(\frac{V}{U_*}\right)^n\left(\frac{F_g}{A} - 1\right)^m$$ $$F_g = \frac{U_*^n}{\sqrt{\left[gD_*\left(\frac{\rho_s}{\rho} - 1\right)\right]}}\left[\frac{V}{32\log(10R_h/D_*)}\right]^{1-n}$$ $$D_* > 60 \begin{cases} A = 0.17 \\ C = 0.025 \\ m = 1.78 \\ n = 0 \end{cases} D_* \leq 1 \begin{cases} A = 0.17 \\ C = 0.025 \\ m - 1.78 \\ n = 1 \end{cases}$$ $$1 < D_* < 60 \begin{cases} A = \dfrac{0.23}{\sqrt{D_*}} + 0.14 \\ \log C = 2.86\log D_* - (\log D_*)^2 - 3.53 \\ m = \dfrac{9.66}{D_*} + 1.34 \\ n = 1 - 0.56\log D_* \end{cases}$$ | Ackers & White (1973); Yang (2006) | Energy balance concept |
| $$\Phi_A = \frac{\left(\dfrac{q_s}{q_t}\right)VR_h}{\sqrt{(G_s - 1)gd_{50}^3}} \cdot \Phi_A = 10.39(\Psi_A)^{-2.52}$$ $$\Psi_A = \frac{(G_s - 1)d_{50}}{S_f R_h}$$ | Graf & Acaroglu (1968) | Shear intensity |
| $$\frac{q_t}{\sqrt{(G_s - 1)gd_{50}^3}} = 0.00139\left(\frac{V}{\sqrt{(G_s - 1)gd_{50}}}\right)^{2.97}\left(\frac{U^*}{\omega_s}\right)^{1.47}$$ | Karim (1998) | Regression method |
| $$\frac{q_t}{\sqrt{(G_s - 1)gd_{50}^3}} = \begin{cases} 0.072078\dfrac{T^{0.893}}{D_*^{0.353}}\left(\dfrac{R_h}{d_{50}}\right)^{0.486} & T > 2.22 \\ 0.000182\dfrac{T^{0.13}}{D_*^{0.673}}\left(\dfrac{R_h}{d_{50}}\right)^{1.16} & T \leq 2.22 \end{cases}$$ | Bhattacharya et al. (2007) | Regression analysis via machine learning |

$q_b$, $q_s$ and $q_t$: bed, suspended and total load transport rate per unit width ($L^2T^{-1}$).
Q, $Q_t$: discharge of water and total sediment load ($L^3T^{-1}$).
$F_g$: sediment mobility parameter (–).
$U_*$: shear velocity ($LT^{-1}$).
$G_s$: sediment specific gravity (–).
$\rho$ and $\rho_s$: density of water and sediment ($ML^{-3}$).
$\gamma$ and $\gamma_s$: specific weight of water and sediment ($ML^{-2}T^{-2}$).
$S_f$: energy slope (m/m).
$\tau_0$ and $\tau_{cr}$: shear and critical shear stress at the bed ($ML^{-1}T^{-2}$).
$R_h$: hydraulic radius (L).
V: average velocity ($LT^{-1}$).
$d_{50}$: particle median size; 50% of the sample is finer (L).
$d_{50sub}$: median particle size for subsurface bed zone (L).
$d_s$: sediment particle diameter (L).
$D_*$: dimensionless particle parameter (–).
$T_*$: transport stage parameter (–).
$\omega_s$: fall velocity of sediment particles ($LT^{-1}$).
g: acceleration due to gravity ($LT^{-2}$).
$\Phi_b$: dimensionless intensity of the bed load transport rate (–).
$\Psi_A$: shear intensity parameter (–).
$\Phi_A$: transport parameter (–).

**Table 6** | Performance criteria for applied models based on scenario 1

| Input models | Method | Performance criteria | | | | | | | |
| | | Train | | | | Test | | | |
| | | R | NSE | RMSE (t/day) | RE (%) | R | NSE | RMSE (t/day) | RE (%) |
|---|---|---|---|---|---|---|---|---|---|
| **Bed load** | | | | | | | | | |
| (I) | SVM | 0.796 | 0.575 | 0.042 | 5.31 | 0.781 | 0.564 | 0.046 | 8.07 |
| | GPR | 0.824 | 0.676 | 0.037 | 5.92 | 0.814 | 0.663 | 0.040 | 8.23 |
| (II) | SVM | 0.809 | 0.615 | 0.040 | 5.88 | 0.779 | 0.565 | 0.045 | 7.99 |
| | GPR | 0.842 | 0.705 | 0.035 | 5.53 | 0.819 | 0.672 | 0.039 | 6.48 |
| (III) | SVM | 0.792 | 0.590 | 0.041 | 4.62 | 0.765 | 0.570 | 0.045 | 8.39 |
| | GPR | 0.829 | 0.685 | 0.036 | 5.90 | 0.814 | 0.660 | 0.040 | 8.32 |
| (IV) | SVM | 0.807 | 0.620 | 0.040 | 5.55 | 0.788 | 0.601 | 0.044 | 8.54 |
| | GPR | 0.841 | 0.706 | 0.035 | 5.35 | 0.830 | 0.688 | 0.038 | 6.55 |
| **Total load** | | | | | | | | | |
| (I) | SVM | 0.916 | 0.794 | 0.026 | 6.35 | 0.876 | 0.713 | 0.037 | 9.39 |
| | GPR | 0.921 | 0.845 | 0.022 | 5.61 | 0.903 | 0.775 | 0.032 | 8.54 |
| (II) | SVM | 0.948 | 0.895 | 0.018 | 7.27 | 0.919 | 0.843 | 0.027 | 9.09 |
| | GPR | 0.947 | 0.893 | 0.018 | 4.50 | 0.927 | 0.850 | 0.026 | 5.48 |
| (III) | SVM | 0.914 | 0.784 | 0.027 | 4.61 | 0.910 | 0.775 | 0.029 | 7.34 |
| | GPR | 0.934 | 0.871 | 0.020 | 5.17 | 0.899 | 0.749 | 0.034 | 8.89 |
| (IV) | SVM | 0.934 | 0.868 | 0.021 | 8.98 | 0.912 | 0.827 | 0.028 | 11.77 |
| | GPR | 0.942 | 0.884 | 0.019 | 4.79 | 0.915 | 0.820 | 0.029 | 6.20 |

**Table 7** | Performance criteria for applied models based on scenario 2

| Input models | Method | Performance criteria | | | | | | | |
| | | Train | | | | Test | | | |
| | | R | NSE | RMSE (t/day) | RE (%) | R | NSE | RMSE (t/day) | RE (%) |
|---|---|---|---|---|---|---|---|---|---|
| **Bed load** | | | | | | | | | |
| BL(I) | SVM | 0.912 | 0.823 | 0.027 | 8.88 | 0.887 | 0.786 | 0.032 | 12.43 |
| | GPR | 0.941 | 0.883 | 0.022 | 4.79 | 0.934 | 0.870 | 0.025 | 7.24 |
| BL(II) | SVM | 0.916 | 0.810 | 0.028 | 7.21 | 0.898 | 0.806 | 0.030 | 10.14 |
| | GPR | 0.926 | 0.854 | 0.024 | 5.78 | 0.916 | 0.831 | 0.028 | 8 |
| BL(III) | SVM | 0.916 | 0.835 | 0.026 | 10.06 | 0.874 | 0.738 | 0.035 | 15.36 |
| | GPR | 0.914 | 0.831 | 0.026 | 6.04 | 0.899 | 0.800 | 0.030 | 9.15 |
| **Total load** | | | | | | | | | |
| TL(I) | SVM | 0.953 | 0.908 | 0.017 | 9.66 | 0.888 | 0.775 | 0.032 | 15.68 |
| | GPR | 0.965 | 0.928 | 0.015 | 4.77 | 0.930 | 0.865 | 0.025 | 8.21 |
| TL(II) | SVM | 0.955 | 0.892 | 0.019 | 6 | 0.932 | 0.858 | 0.026 | 9.07 |
| | GPR | 0.968 | 0.935 | 0.014 | 4.15 | 0.948 | 0.894 | 0.022 | 6.56 |
| TL(III) | SVM | 0.946 | 0.888 | 0.019 | 13.16 | 0.891 | 0.778 | 0.032 | 17.31 |
| | GPR | 0.977 | 0.954 | 0.012 | 4.39 | 0.941 | 0.882 | 0.023 | 9.44 |
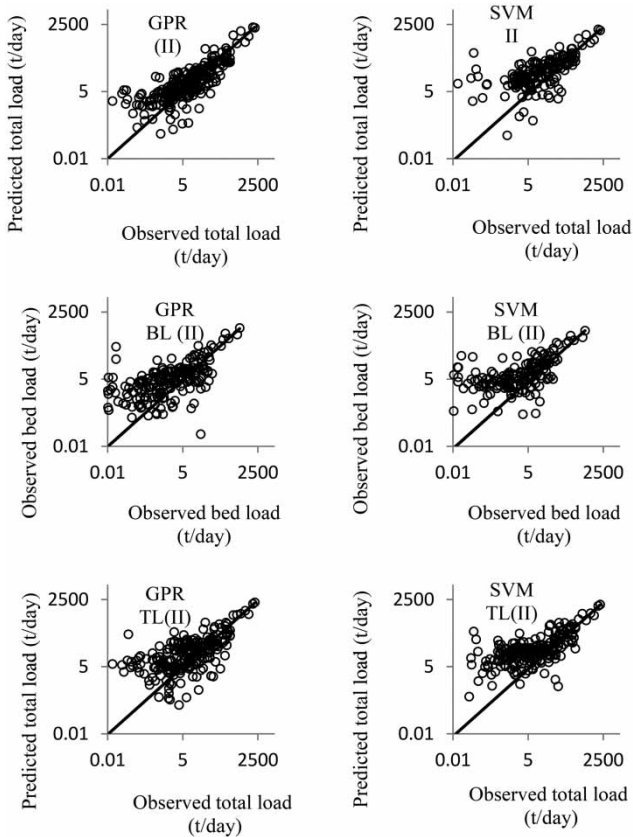
**Figure 1** | Scatter plots of observed and predicted sediment rate obtained from the best models of each scenario.

using four inputs ensures the best performance, and an increased number of inputs did not have any effect on improving the accuracy of the employed methods. According to NSE values in predicting the total load, when comparing TL(I) and TL(II), introducing particle mobility parameter $\theta$ and $V^3/g.y.\omega_s$ instead of $T$ and $D_*$ improves the accuracy of the models in NSE = 0.894 (GPR) and NSE = 0.858 (SVM) for the test set. Furthermore, in the case of bed load, considering $R_h/d_{50}$ and $\theta$ and also omitting $D_*$ and $T$ leads to better outcomes in NSE = 0.831 (GPR) and NSE = 0.806 (SVM). However, according to the performance statistics which are presented in Table 7, GPR shows more flexibility and provides better prediction capability for both BL(I) and BL(II) models. Results of models BL(I) and TL(I) revealed that considering the combination of $T$ and $D_*$ with influential parameters of $Fr$ and $V/U_*$ produces relatively accurate prediction for sediment transport rate in gravel-bed rivers. Therefore, it can be

assumed that the dimensionless shear stress and dimensionless median particle size are effective parameters in prediction of sediment transport rates. From the kernel-based methods utilized, a detailed comparison of the overall performance shows that prediction of GPR is reasonably better than SVM in the case of sediment transport rate. The scatter plots of the model predictions for the test set including 278 points are depicted in Figure 1. Due to the high dispersion of data in low sediment transport rate and in order to compare the obtained results in a better way, the scatter plots are shown on logarithmic scale.

Figure 2 illustrates NSE values of different $\gamma$ values of the employed GPR and SVM models (fed with the BL(II) and TL(II) as the best input combinations). In the case of RBF kernel, $\gamma$ indicates the optimal width of kernel function. From the figure it can be seen that the NSE values fluctuate with varying $\gamma$ values. Considering the SVM approach, small values of $\gamma$ lead to the risk of overfitting (as a result of ignoring most of the support vectors). Conversely, GPR provides better performance with smaller $\gamma$ values and is less threatened by the danger of overfitting. Moreover, in contrast to the SVM method, a clear smooth change of NSE values with variation of $\gamma$ values can be seen in utilization of the GPR approach.

### Results of empirical equations

The results of selected empirical equations in comparison to GPR and SVM for predicting bed load and total load are demonstrated in Figure 3. Based on the value of RMSE, it can be clearly seen that none of the proposed equations are sufficiently precise. The important point about empirical methods is that the existing equations were developed in special laboratories with specific flow conditions and sediment particle features, therefore, these equations show acceptable results in particular conditions, but their applicability to field data with various hydraulic conditions is questionable. However, developing an equation that quantifies bed load and total load rate for all streams, seems to be impossible.

### Sensitivity analysis

In this step, sensitivity analysis is used to investigate the effect of different parameters on the sediment transport process.
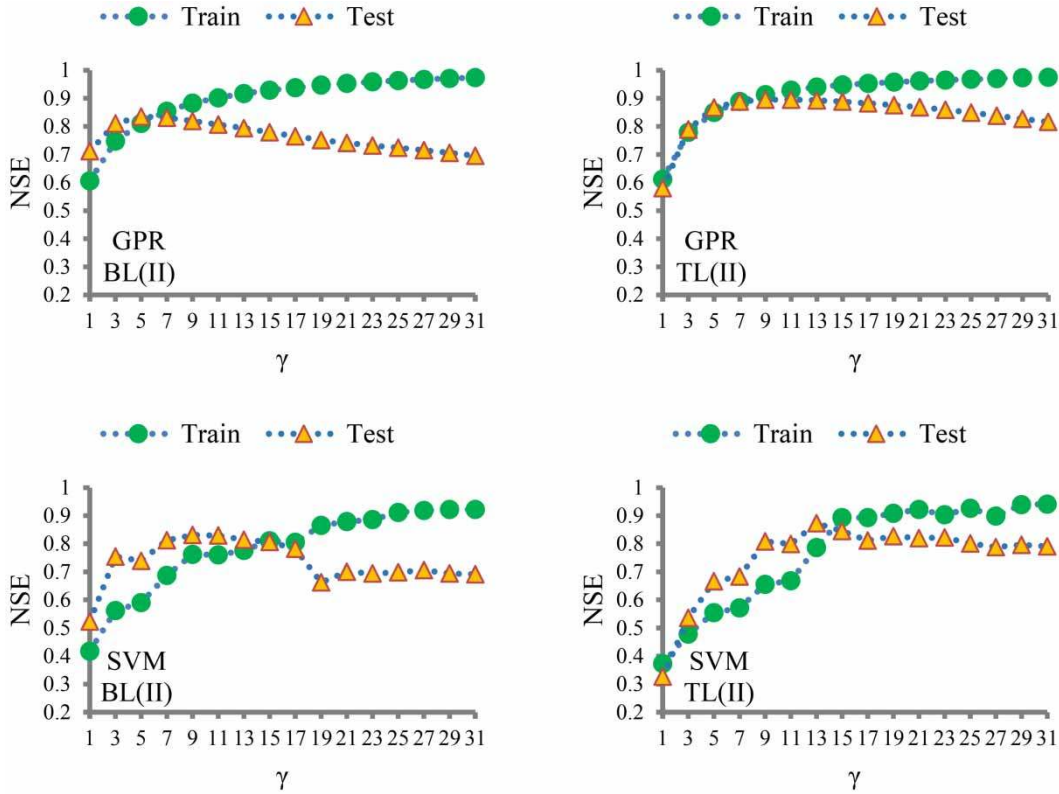
**Figure 2** | Variation of NSE vs $\gamma$ values for BL(II) and TL(II) models.
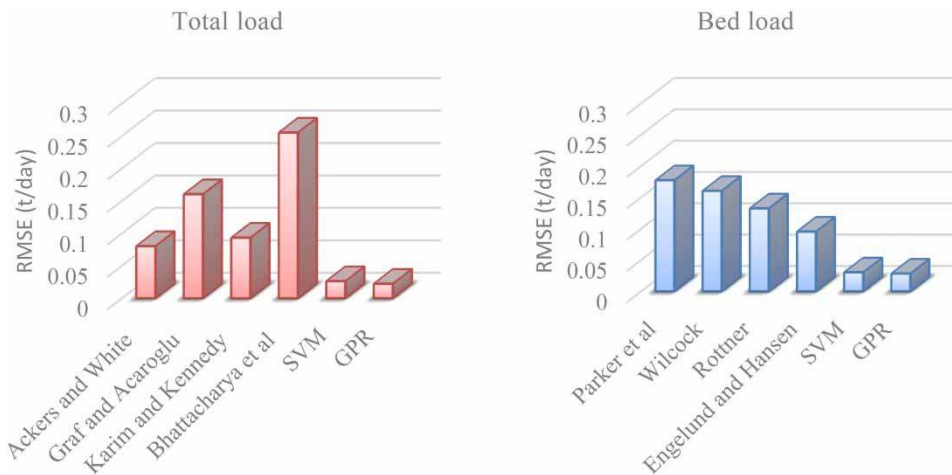


**Figure 3** | Results of empirical equations in terms of RMSE values.

The superior models with four inputs for bed load and total load were selected and the importance of each parameter was evaluated by eliminating them. According to the results of sensitivity analysis which are presented in Table 8, it can

be clearly seen that the ratio of average to shear velocity $(V/U_*)$ has the most significant effect in quantification of sediment transport rate in gravel-bed rivers, which represents the flow resistance in open channels. Furthermore, the Froude

**Table 8** | Results of sensitivity analysis

| | | | Performance criteria | | | | | | | |
| | | | Train | | | | Test | | | |
| Best model | Eliminated variable | Method | R | NSE | RMSE (t/day) | RE (%) | R | NSE | RMSE (t/day) | RE (%) |
|---|---|---|---|---|---|---|---|---|---|---|
| **BL(II)** | Fr | SVM | 0.833 | 0.658 | 0.038 | 5.89 | 0.804 | 0.641 | 0.041 | 7.83 |
| | | GPR | 0.889 | 0.789 | 0.029 | 5.14 | 0.874 | 0.706 | 0.037 | 7.69 |
| | $\frac{V}{U_*}$ | SVM | 0.446 | 0.173 | 0.059 | 21.84 | 0.233 | 0.002 | 0.069 | 24.48 |
| | | GPR | 0.530 | 0.278 | 0.055 | 13.70 | 0.400 | 0.154 | 0.064 | 15.52 |
| | $\frac{R_h}{d_{50}}$ | SVM | 0.896 | 0.760 | 0.031 | 6.37 | 0.877 | 0.737 | 0.035 | 9.57 |
| | | GPR | 0.933 | 0.869 | 0.023 | 4.89 | 0.900 | 0.807 | 0.030 | 6.59 |
| | $\theta$ | SVM | 0.903 | 0.792 | 0.029 | 10.31 | 0.884 | 0.778 | 0.032 | 13.60 |
| | | GPR | 0.921 | 0.845 | 0.025 | 6.14 | 0.900 | 0.781 | 0.032 | 8.71 |
| TL(II) | Fr | SVM | 0.879 | 0.727 | 0.030 | 3.17 | 0.847 | 0.603 | 0.043 | 7.56 |
| | | GPR | 0.962 | 0.924 | 0.016 | 4 | 0.947 | 0.882 | 0.023 | 6.04 |
| | $\frac{V}{U_*}$ | SVM | 0.607 | 0.205 | 0.051 | 8.32% | 0.273 | 0.027 | 0.068 | 9.75 |
| | | GPR | 0.801 | 0.640 | 0.034 | 9.55 | 0.332 | 0.064 | 0.067 | 13.87 |
| | $\theta$ | SVM | 0.951 | 0.897 | 0.018 | 13.35 | 0.863 | 0.738 | 0.035 | 16.92 |
| | | GPR | 0.970 | 0.940 | 0.014 | 4.34 | 0.951 | 0.901 | 0.021 | 7.26 |
| | $\frac{V^3}{gy\omega_s}$ | SVM | 0.966 | 0.932 | 0.015 | 9.09 | 0.891 | 0.782 | 0.032 | 12.57 |
| | | GPR | 0.970 | 0.940 | 0.014 | 4.5 | 0.947 | 0.887 | 0.023 | 7.48 |

number is also a common parameter in relative modeling and plays an important role in predicting bed and total sediment load. Meantime, results of performed sensitivity analysis (Table 8) show that for prediction of total sediment load, elimination of $\theta$ from the list of input variables leads to better generalization ability of the GPR approach (R = 0.951, NSE = 0.901, RMSE = 0.021, and RE = 7.26%). From the analysis, it can be inferred that the employed GPR approach with three inputs is able to successfully predict total load transport in a great variety of gravel-bed rivers.

## CONCLUSION

In this study, datasets of 19 gravel-bed streams and rivers located in the State of Idaho, USA, were used to exhibit the functionality of machine learning methods in predicting sediment transport rate. Different combinations of non-dimensional parameters based on two scenarios were developed in order to attain the purpose and obtained results were compared with empirical approaches. The obtained results of the employed GPR and SVM methods demonstrate a great performance over empirical formulas. In predicting bed load transport rate, the obtained results reveal that the second scenario based on flow conditions

and sediment properties is more accurate, while in the case of total sediment load, both scenarios lead to good outcomes. It was found that the inclusion of Fr, V/U*, R/D50, and $\theta$ inputs resulted in the best performance accuracy for prediction of bed load transport rate, and conversely, in prediction of total sediment load, using Fr, V/U*, $\theta$, and $V^3/g.y.\omega_s$ yielded the best results. Performing sensitivity analysis demonstrates the significant effect of V/U* in sediment transport rate of gravel-bed rivers. The results found that the employed kernel-based approach represented by the GPR model was quite accurate in respect to prediction of sediment transport rates of gravel-bed rivers and performed better than the common SVM method. In addition, the Froude number plays an important role and is a common parameter in all superior models. However, the GPR and SVM are data-driven models and the results presented here are data sensitive, so further studies should be done using data from different rivers worldwide to evaluate the effectiveness of the recommended models.

## REFERENCES

Ackers, P. & White, W. R. 1973 Sediment transport: new approach and analysis. *Journal of the Hydraulics Division* **99** (11), 2041–2060.

Azamathulla, H. M., Chang, C. K., Ghani, A. A., Ariffin, J., Zakaria, N. A. & Hasan, Z. A. 2009 An ANFIS-based approach for predicting the bed load for moderately sized rivers. *Journal of Hydro-Environment Research* **3** (1), 35–44.

Azamathulla, H. M., Ghani, A. A., Chang, C. K., Hasan, Z. A. & Zakaria, N. A. 2010 Machine learning approach to predict sediment load–a case study. *CLEAN–Soil, Air, Water* **38** (10), 969–976.

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

Barry, J. J., Buffington, J. M. & King, J. G. 2004 A general power equation for predicting bed load transport rates in gravel bed rivers. *Water Resources Research* **40** (10), W1041.

Bathurst, J. C. 2007 Effect of coarse surface layer on bed-load transport. *Journal of Hydraulic Engineering* **133** (11), 1192–1205.

Bhattacharya, B., Price, R. K. & Solomatine, D. P. 2007 Machine learning approach to modeling sediment transport. *Journal of Hydraulic Engineering* **133** (4), 440–450.

Doğan, E., Yüksel, İ & Kişi, Ö. 2007 Estimation of total sediment load concentration obtained by experimental study using artificial neural networks. *Environmental Fluid Mechanics* **7** (4), 271–288.

Engelund, F. & Hansen, E. 1967 *A Monograph on Sediment Transport in Alluvial Stream*. Teknisk Forlag, Copenhagen, Denmark.

Ghani, A. A. & Azamathulla, H. M. 2014 Development of GEP-based functional relationship for sediment transport in tropical rivers. *Neural Computing and Applications* **24** (2), 271–276.

Graf, W. H. & Acaroglu, E. R. 1968 Sediment transport in conveyance systems (Part 1)/A physical model for sediment transport in conveyance systems. *Hydrological Sciences Journal* **13** (2), 20–39.

Grbić, R., Kurtagić, D. & Slišković, D. 2013 Stream water temperature prediction based on Gaussian process regression. *Expert Systems with Applications* **40** (18), 7407–7414.

Karim, F. 1998 Bed material discharge prediction for nonuniform bed sediments. *Journal of Hydraulic Engineering* **124** (6), 597–604.

Khorram, S. & Ergil, M. 2010a A sensitivity analysis of total-load prediction parameters in standard sediment transport equations. *JAWRA Journal of the American Water Resources Association* **46** (6), 1091–1115.

Khorram, S. & Ergil, M. 2010b Most influential parameters for the bed-load sediment flux equations used in alluvial rivers. *JAWRA Journal of the American Water Resources Association* **46** (6), 1065–1090.

King, J. G., Emmett, W. W., Whiting, P. J., Kenworthy, R. P. & Barry, J. J. 2004 *Sediment Transport Data and Related Information for Selected Coarse-Bed Streams and Rivers in Idaho*. Gen. Tech. Rep. RMRS-GTR-131. US Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, CO, p. 26, 131.

Kitsikoudis, V., Sidiropoulos, E. & Hrissanthou, V. 2014 Machine learning utilization for bed load transport in gravel-bed rivers. *Water Resources Management* **28** (11), 3727–3743.

Komasi, M., Sharghi, S. & Safavi, H. R. 2018 Wavelet and cuckoo search-support vector machine conjugation for drought forecasting using Standardized Precipitation Index (case study: Urmia Lake. Iran). *Journal of Hydroinformatics* **20** (4), 975–988.

Kumar, B. 2012 Neural network prediction of bed material load transport. *Hydrological Sciences Journal* **57** (5), 956–966.

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

Molinas, A. & Wu, B. 2001 Transport of sediment in large sand-bed rivers. *Journal of Hydraulic Research* **39** (2), 135–146.

Okcu, D., Pektas, A. O. & Uyumaz, A. 2016 Creating a non-linear total sediment load formula using polynomial best subset regression model. *Journal of Hydrology* **539**, 662–673.

Pal, M., Singh, N. K. & Tiwari, N. K. 2014 Kernel methods for pier scour modeling using field data. *Journal of Hydroinformatics* **16** (4), 784–796.

Parker, G., Klingeman, P. C. & McLean, D. G. 1982 Bedload and size distribution in paved gravel-bed streams. *Journal of the Hydraulics Division* **108** (4), 544–571.

Pitlick, J., Cui, Y. & Wilcock, P. R. 2009 *Manual for Computing Bed Load Transport Using BAGS (Bedload Assessment for Gravel-bed Streams) Software*. US Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, CO.

Recking, A. 2010 A comparison between flume and field bed load transport data and consequences for surface-based bed load transport prediction. *Water Resources Research* **46**, 3.

Rottner, J. 1959 A formula for bed load transportation. *La Houille Blanche* **14** (3), 285–307.

Roushangar, K. & Alizadeh, F. 2018 Investigating effect of socio-economic and climatic variables in urban water consumption prediction via Gaussian process regression approach. *Water Science and Technology: Water Supply* **18** (1), 84–93.

Roushangar, K. & Koosheh, A. 2015 Evaluation of GA-SVR method for modeling bed load transport in gravel-bed rivers. *Journal of Hydrology* **527**, 1142–1152.

Roushangar, K., Mehrabani, F. V. & Shiri, J. 2014 Modeling river total bed material load discharge using artificial intelligence approaches (based on conceptual inputs). *Journal of Hydrology* **514**, 114–122.

Roushangar, K., Garekhani, S. & Alizadeh, F. 2016 Forecasting daily seepage discharge of an earth dam using wavelet–mutual information–Gaussian process regression approaches. *Geotechnical and Geological Engineering* **34** (5), 1313–1326.

Sahraei, S., Alizadeh, M. R., Talebbeydokhti, N. & Dehghani, M. 2017 Bed material load estimation in channels using machine learning and meta-heuristic methods. *Journal of Hydroinformatics* **20** (1), 100–116.

Sasal, M., Kashyap, S., Rennie, C. D. & Nistor, I. 2009 Artificial neural network for bedload estimation in alluvial rivers. *Journal of Hydraulic Research* **47** (2), 223–232.

Schneider, J. M., Rickenmann, D., Turowski, J. M., Bunte, K. & Kirchner, J. W. 2015 Applicability of bed load transport models for mixed-size sediments in steep streams considering macro-roughness. *Water Resources Research* **51** (7), 5260–5283.

Sun, A. Y., Wang, D. & Xu, X. 2014 Monthly streamflow forecasting using Gaussian process regression. *Journal of Hydrology* **511**, 72–81.

van den Berg, J. H. & van Gelder, A. 1993 A new bedform stability diagram, with emphasis on the transition of ripples to plane bed in flows over fine sand and silt. *Special Publications of the International Association of Sedimentologists* **17**, 11–21.

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

Wang, T. & Liu, X. 2009 The breakup of armor layer in a gravel-bed stream with no sediment supply. In: *Advances in Water Resources and Hydraulic Engineering* (C. Zhang & H. Tang, eds). Springer, Berlin, Heidelberg, pp. 919–923.

Wilcock, P. R. 2001 Toward a practical method for estimating sediment-transport rates in gravel-bed rivers. *Earth Surface Processes and Landforms* **26** (13), 1395–1408.

Yang, C. T. 1996 *Sediment Transport: Theory and Practice*. McGraw-Hill, Singapore.

Yang, C. T. 2006 *Erosion and Sedimentation Manual*. US Department of the Interior, Bureau of Reclamation, Denver, CO.

Yang, C. T., Marsooli, R. & Aalami, M. T. 2009 Evaluation of total load sediment transport formulas using ANN. *International Journal of Sediment Research* **24** (3), 274–286.

Yuan, J., Wang, K., Yu, T. & Fang, M. 2008 Reliable multi-objective optimization of high-speed WEDM process based on Gaussian process regression. *International Journal of Machine Tools and Manufacture* **48** (1), 47–60.

Zhu, S., Luo, X., Xu, Z. & Ye, L. 2018 Seasonal streamflow forecasts using mixture-kernel GPR and advanced methods of input variable selection. *Hydrology Research* **50** (1), 200–214.