J. Mt. Sci. (2016) 13(10): 1773-1788 e-mail: jms@imde.ac.cn http://jms.imde.ac.cn
DOI: 10.1007/s11629-015-3726-0

Local vs. cross station simulation of suspended sediment load in successive hydrometric stations: heuristic modeling approach

Kiyoumars ROUSHANGAR[1*] http://orcid.org/0000-0002-2620-5106; e-mail: kroshangar@yahoo.com

Shabnam HOSSEINZADEH[1] http://orcid.org/0000-0002-7177-1879; e-mail: shabnamhoseinzade@yahoo.com

Jalal SHIRI[2] http://orcid.org/0000-0002-5726-7924; e-mail: j_shiri2005@yahoo.com

* *Corresponding author*

1 Department of Civil Hydraulics Engineering, Faulty of Civil Engineering, University of Tabriz, Tabriz 51666, Iran

2 Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz 51666, Iran

Citation: Roushangar K, Hosseinzadeh SH, Shiri J (2016) Local vs. cross station simulation of suspended sediment load in successive hydrometric stations: heuristic modeling approach. Journal of Mountain Science 13(10). DOI: 10.1007/s11629-015-3726-0

Abstract: The present paper aims at modeling suspended sediment load (SSL) using heuristic data driven methodologies, e.g. Gene Expression Programming (GEP) and Support Vector Machine (SVM) in three successive hydrometric stations of Housatonic River in U.S. The simulations were carried out through local and cross-station data management scenarios to investigate the interrelations between the SSL values of upstream/downstream stations. The available scenarios were applied to predict SSL values using GEP to obtain the best models. Then, the best models were predicted by SVM approach and the obtained results were compared with those of GEP. The comparison of the results revealed that the SVM technique is more capable than the GEP for modeling the SSL through the both local and cross-station data management strategies. Besides, local application seems to be better than cross-station application for modeling SSL. Nevertheless, the cross-station application demonstrated to be a valid methodology for simulating SSL, which would be of interest for the stations with lack of observational data. Also, the prediction capability of conventional Sediment Rating Curve (SRC) method was compared with those of GEP and SVM techniques. The obtained results revealed the superiority of GEP and SVM-based models over the traditional SRC technique in the studied stations.

Keywords: Suspended sediment load; Successive hydrometric stations; Gene expression programming; Support vector machine

Received: 26 September 2015
Revised: 11 January 2016
Accepted: 20 March 2016

Introduction

The correct estimation of the volume of sediment being transported by streams is of great importance in water resources engineering, as it directly affects the planning, design and management of hydraulic structures. So far, numerous studies have been carried out to investigate the possible functional relationships between the Suspended Sediment Load (SSL) and flow characteristics, based on different concepts, leading to some empirical and semi-empirical SSL modeling equations. However, none of such equations have received universal acceptance (Aytek and Kisi 2008). The physically-based distributed models are based on the simplified partial differential equations of flow and sediment flux. Examples of such models are presented by e.g., Singh et al. (1998); Yang (1996); Cohn et al. (1992); Forman et al. (2000). Nonetheless, over the years, some sediment rating methods based on producing the SRC have been proposed by researchers to determine the average relationship between the flow discharge and SSL. Among others, Asselman (2000) fitted SRC for different locations along the Rhine River and its main tributaries. Rating curves obtained by least squares regression on logarithmic transformed data tend to underestimate the sediment transport rates. Picoet et al. (2001) applied statistical approaches and a lumped conceptual model to estimate the time evolution of Suspended Sediment Concentration (SSC) using the water discharge data. Both of the models were compared with a simple rating curve model. The obtained results confirmed the superiority of the conceptual model. Overleir (2004) showed that the classical non-linear least squares (NLS) method can model only a very limited class of variance heterogeneity. By introducing a heteroscedastic maximum likelihood model, the variance heterogeneity was treated more generally. The maximum likelihood model stabilizes the variance better than the NLS approach, and thus is a more robust and appropriate approach to fit a rating curve to a set of discharge measurements. Crowder et al. (2007) compared a nonlinear regression technique with the linear regression method at 26 United States Geological Survey gauging stations throughout the Upper Mississippi River basin for predicting annual and cumulative suspended sediment yields. Sensitivity analysis was conducted at two stations, one having a concave sediment load–discharge plot and one having a convex sediment load–discharge plot, to determine rating curve's capabilities. Results suggested that the regression technique that will yield the best sediment load estimates (with the lowest error values) will depend on at least three factors: (1) the shape and variability within the underlying sediment load–discharge relationship; (2) the amount of data used to calibrate the regression equation; and (3) the time frame over which sediment loads are being estimated (e.g. annual or longer term cumulative yields).

During recent years, different heuristic techniques [e.g. Artificial Neural Network (ANN), GEP and SVM)] have been successfully applied for modeling various hydrological processes. Alp and Cigizoglu (2007) employed ANN method using hydrometeorological data to estimate daily total SSL on rivers. The simulation results were compared with conventional Multi-Linear Regression (MLR). They found that the results of ANN were better than MLR models. Partal and Cigizoglu (2008) used a combined wavelet-ANN model to predict SSL in rivers and compared the results with those of the ANN and the SRC methods. The results showed the higher accuracy of the hybrid wavelet-ANN model. Aytek and Kisi (2008) applied Genetic Programming (GP) technique to model SSL and compared the results with SRC and MLR techniques. The results indicated that the proposed GP formulation performs better than the SRC and MLR models. Azamathulla et al. (2010) employed SVM to predict SSL. The SVM technique demonstrated a superior performance compared to other traditional sediment-load methods. Chiang and Tsai (2011) estimated SSL by using linear as well as the power regression models, ANN and SVM. The results showed the higher accuracy of the SVM-based models. Kisi and Cimen (2011) introduced a wavelet-SVM conjunction model for monthly streamflow forecasting. Azamathulla et al. (2011) introduced the GEP-based models as high-accuracy tools for modeling stage-discharge relationship. Kisi and Shiri (2011) forecasted daily precipitation values by establishing hybrid wavelet-genetic programming (WGP) and wavelet-neuro-fuzzy (WNF) approaches. Kisi (2012) studied the ability of Least Square Support Vector Machine (LSSVM) in prediction of SSC of upstream and downstream stations in rivers. Liu et al. (2013) developed a wavelet-ANN model (WANN) to predict SSC in the Kuye River. The comparison of results of WANN models with SRC and ANN models revealed that the WANN prediction accuracy is higher than the other applied techniques. Kakaei Lafdani et al. (2013) investigated the abilities of SVM and ANN models for predicting daily SSL. The obtained results illustrated that the ANN models and nu-SVR model using Gamma Test for input selection had better performance than regression combination. Roushangar et al. (2014) evaluated the performance of NF and GEP to predict total bed material load. These models were compared with some well-known traditional models. The obtained results showed that the GEP and ANFIS performed better than traditional models. Also, Roushangar and Koosheh (2015) modeled bed load transport in three gravel-bed rivers using Genetic Algorithm-Support Vector Machine (GA-SVR) method with different kernel functions. They obtained the most influential hydraulic parameters for predicting the bed load transport of each river.

The present paper aims at assessing the temporal and spatial data management scenarios for modeling SSL, using successive hydrometric stations data. Consequently, the models were assessed per station using traditional local data splitting models (two-block splitting mode), through defining a time dependent train-test selection method. Then, a cross-station approach was applied which involves using the ancillary data for the both input and target values.

1 Materials and Methods

1.1 Study area and used data

Daily stream flow and SSL data from three hydrometric stations on the Housatonic River in U.S.A (operated by US Geological Survey, USGS) were used for modeling SSL. The distance between each successive station is about 50 km. The first station (station no: 01197500) is located near Great Barrington, Berkshire County, Massachusetts. The second (station no: 01199000) is located at Falls Village, Litchfield County, Connecticut, and the third one (station no: 01200500) is located at Gaylordsville, Litchfield County, Connecticut. The corresponding basin areas of the stations are, respectively, 282, 634, and 996 sq. Mile. Figure 1 shows the geographical positions of the studied stations.

The observed data covering a period of two years (from April 01, 1979 to September 30, 1980), were utilized for training and testing the employed models. Figure 2 illustrates the time series plot of daily flow discharge (Q) and SSL for the studied stations during the study period. It is clear that the trends of flow discharge and SSL transport rate variations are similar in each station and among the all stations. Analyzing the time series plots shows that the highest SSL transport corresponds to the maximum flow discharge values in all three stations, while the lowest SSL transport rates are observed during the minimum river flow rates. This might dictate a direct relation between SSL and flow discharge, although there are some events with variable SSL rate and fixed values of flow discharge during the time period, which may be explained by the influence of other parameters on SSL transport rate. For developing and testing the models, the data were randomly selected so that the 70% of the whole data set was used for training

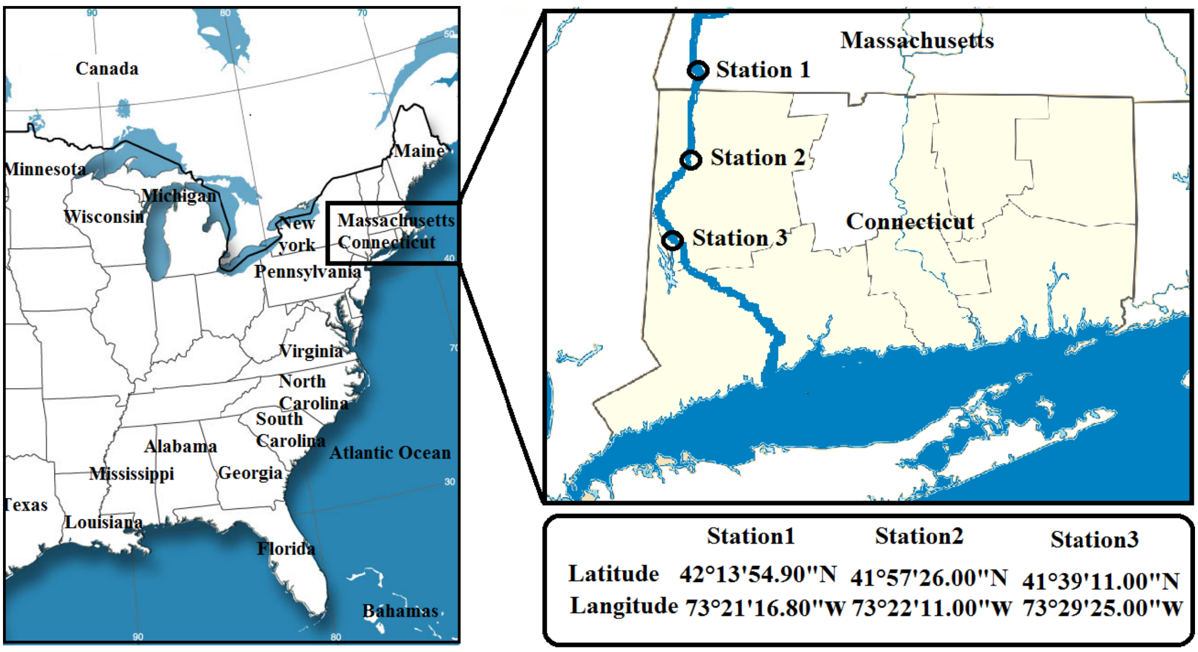

Figure 1 The locations of studied stations.

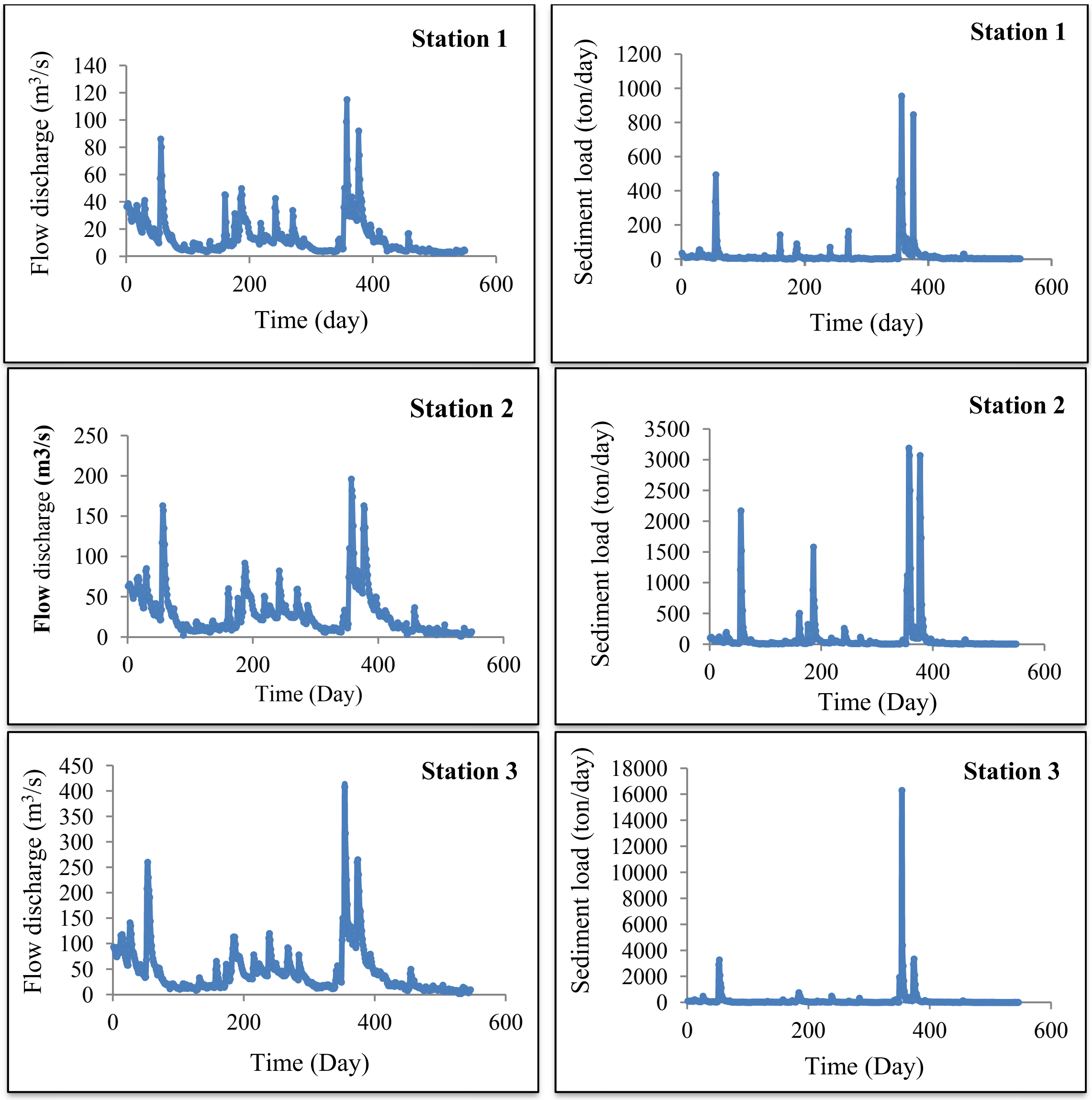

Figure 2 Time series plot of river discharge and sediment load for the Berkshire County, Massachusetts (station 1), Litchfield County, Connecticut (station 2 and station 3) during the study period.

and the remaining 30% were used for testing. Although two-block data splitting method might cause the risk of over fitting, and three-block norm or k-fold testing might allow for a suitable data scanning (Shiri et al. 2013 and 2014), the random selection of the train-test set would reduce this risk as discussed by Roushangar et al. (2014a, 2014b). Moreover, the random train-test set selection was carried out taking into consideration the statistical parameters of the observed data. Subsequently, the most skewed data were used as train data for better learning the models with extreme events. Table 1 sums up the statistical parameters of the observed flow discharge and SSL data for the train, test and whole study period. In the Table, the X_{min}, X_{max}, X_{mean}, S_X, C_V and C_{SX} present the minimum, maximum, mean, standard deviation, coefficient of variation, and skewness coefficient, respectively. It can be seen from the skewness coefficients that both the discharge and sediment data represent highly scattered distribution.

1.2 Sediment rating curve

The establishment of a SRC is an important problem in hydrologic studies. The SRC is used to

Table 1 Statistical parameters of the used data set

Statistical Parameters	Training set		Testing set		All data set	
	Q (m^3/s)	SSL (ton day^{-1})	Q (m^3/s)	SSL (ton day^{-1})	Q (m^3/s)	SSL (ton day^{-1})
Station 1						
X_{mean}	13.82	18.63	13.96	21.6	13.86	19.52
S_X	13.96	60.91	14.53	89.68	14.12	70.72
$C_v(S_x/X_{mean})$	1	3.27	1.04	4.15	1.02	3.62
C_{sx}	2.76	8.95	2.86	8.42	2.78	9.1
X_{max}	115	846	98.9	955	115	955
X_{min}	2.29	0.38	2.61	0.38	2.29	0.38
Station 2						
X_{mean}	28.47	91.92	28.85	104.57	28.58	95.72
S_X	28.16	326.68	29.63	375.16	28.58	341.68
$C_v(S_x/X_{mean})$	0.99	3.55	1.03	3.59	1	3.57
C_{sx}	2.23	6.78	2.75	5.93	2.47	6.47
X_{max}	182	3070	196	3190	196	3190
X_{min}	1.1	0.21	1.05	0.2	1.05	0.2
Station 3						
X_{mean}	44.52	158.35	44.35	125.34	44.46	148.49
S_X	47.81	918.03	43.95	423.48	46.65	802.66
$C_v(S_x/X_{mean})$	1.07	5.8	0.99	3.38	1.05	5.4
C_{sx}	3.1	14.84	2.52	6.16	2.96	15.83
X_{max}	408	16300	265	3340	408	16300
X_{min}	2.07	0.59	2.07	0.59	2.07	0.59

assess the sediment load corresponding to the measured flow discharge. The most common SRC form is a power function (Walling 1978), as:

$$S = aQ^b \tag{1}$$

where, Q and S denote flow discharge and SSL, respectively. Values of a and b for a particular flow are determined from data via a linear regression between (log$S$) and (log$Q$).

1.3 Gene expression programming

GEP was developed by Ferreira (2001), using the fundamental principles of the Genetic Algorithms (GA) and GP. GEP mimics biological evolution to create a computer program for modeling a specified phenomenon. The problems are encoded in linear chromosomes of fixed-length as a computer program (Ferreira 2001). The chromosomes are composed of multiple genes, each gene encoding a smaller sub-program. GEP technique starts the solution of a specific problem with the random generation of the chromosomes of the initial population. Each individual chromosome in the initial population is then expressed, and its fitness is evaluated using the fitness function. In this study, the Root Mean Square Error (RMSE) is applied as fitness function:

$$RMSE_i = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_{ij} - O_j)^2} \tag{2}$$

where P_{ij} is the value predicted by individual program i for fitness case j (out of sample case), and O_j is the observed value for fitness case. For a perfect fit, P_{ij} = O_j and $RMSE_i$=0. For evaluating the fitness f_i of an individual program i, the following equation should be applied:

$$f_i = 1000\frac{1}{1+RMSE_i} \tag{3}$$

The mentioned function (f_i) ranges between 0 and 1000 with 1000 corresponded to the ideal (Ferreira 2006).

The structural and functional organization of the linear chromosomes allows the unconstrained operation of important genetic operators such as mutation, transposition, and recombination.

The following procedure was applied to model SSL through GEP according to Ferreira (2006):

i) Selecting the appropriate fitness function.

ii) Selecting the terminal set (*T*) and the operation function set (*F*) for creating the chromosomes: the terminal set includes recorded river discharge and sediment load values. Also, four basic arithmetic operators (+, -, *, /) and some basic mathematical functions ($\sqrt{\ }, \sqrt[3]{\ }, \ln(x)$, e^x , x^2, $x^3, \sin x, \cos x, \arctan x$) were utilized as function set.

iii) Selecting the chromosomal architecture, i.e. the length of head and the number of genes: Length of head, h=7, and three genes per chromosome were employed.

iv) Choosing the linking function: addition and multiplication linking function was applied in the present study. In general, the choice of linking function depends on the problem and there is not any basic rule to identify which of these functions is preferred to another. (Kisi et al. 2012)

v) Choosing the set of genetic operators: Default values of the genetic operators were applied since these values have been recommended for hydrological studies by literature (e.g., Shiri and Kisi 2011; Kisi and Shiri 2012; Kisi et al. 2012).

Further details about developing the GEP-based models for simulating SSL might be found in e.g. Kisi et al. (2012) and Shiri and Kisi (2012).

1.4 Support vector machine

The idea of support vector machines (SVMs), which are known as the classification and regression procedures, has been developed by Vapnik (1998). SVM includes the structural risk minimization (SRM) principle in its formulation, which has demonstrated its superiority to the traditional empirical risk minimization (ERM) principle, employed by conventional neural networks. SRM minimizes the upper bound on the expected risk, while ERM minimizes the errors corresponded to the training data (Vapnik 1998; Gunn 1998). The SVM method is based on the concept of optimal hyperplane that separates samples of two classes by considering the widest gap between two classes. SVM solves a quadratic problem in which the objective function is obtained by combination of loss function and regularization term (Basak et al. 2007).

The main process of SVM model building consists of selecting support vectors which support the model structure and determine their weights. The process of an SVM estimator (f) on regression can be described as:

$$f(x) = w.\phi(x) + b \tag{4}$$

where w and b denote a weight vector, a bias, respectively. $\o$ stands for a nonlinear transfer function mapping the input space into a high-dimensional feature space.

Vapnik (1998) introduced an object function of convex optimization with ε-insensitivity loss function for solving the Eq. (4), of which mathematical expressions are described as

$$\underset{w,b,\xi,\xi^*}{\text{minimize}} \qquad \frac{1}{2}\|w\|^2 + C\sum_{k=1}^{N}\left(\xi_k + \xi_k^*\right)$$

Subject to

$$\begin{cases} y_k - w^T\varphi(\mathbf{x}_k) - b \le \varepsilon + \xi_k \\ w^T\varphi(\mathbf{x}_k) + b - y_k \le \varepsilon + \xi_k^* \qquad k = 1,2,\dots,N \\ \xi_k, \xi_k^* \ge 0 \end{cases} \tag{5}$$

where ξ and ξ^* are slack variables penalizing estimation error by the ε-insensitivity loss function, and C is a positive tradeoff parameter for the degree of the empirical error. Finally, Karush-Kuhn-Tucker conditions are applied to the regression, and Eq. (5) thus yields the dual Lagrangian

Maximize

$$R(\alpha,\alpha^*) = -\frac{1}{2}\sum_{i,j=1}^{N}\begin{array}{l}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)K(x_i,x_j) - \\ \varepsilon\sum_{i=1}^{N}(\alpha_i + \alpha_i^*) + d_i\sum_{i=1}^{N}(\alpha_i - \alpha_i^*)\end{array}$$

Subject to

$$\begin{cases} \sum_{i=1}^{N}(\alpha_i - \alpha_i^*) = 0 \\ 0 \le \alpha_i \le C \\ 0 \le \alpha_i^* \le C, i = 1,2,\dots,N \end{cases} \tag{6}$$

where α_i, α_i^* are Lagrangian multipliers and $K(x_i, x_j)$ denotes the kernel function which transferred the input space into feature space. The value of the Kernel equals the inner product of two vectors, x_i and x_j in the feature space $\o(x_i)$ and $\o(x_j)$, that is $K(x_i, x_j)= \o(x_i).\ \o(x_j)$. A suitable choice of kernel allows the data to become separable in the feature space despite being non-separable in the original input space.

In the present study, Radial Basis Function (RBF) was used as well-known kernel function for SVM application. RBF kernel is the general purpose kernel used when there is no prior knowledge about the data which has the form:

$$K(x_i,x_j) = exp(-\frac{1}{2}(\frac{\|x_i - x_j\|}{\sigma})^2) \tag{7}$$

After calculating α_i and α_i^* , an optimal desired weights vector of regression is:

$$w^* = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)K(x_i, x_j) \quad (8)$$

Therefore, the regression function can be defined as:

$$f(x) = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)K(x_i, x_j) + b \quad (9)$$

The selection of three parameters, σ, ε and C, of a SVM model is important to the accuracy of predicting. Thus the optimal values of mentioned parameters should be set to obtain the most accurate result.

1.5 Performance criteria

The capabilities of models are evaluated by using the coefficient of determination (R^2), Nash-Sutcliffe coefficient (*NS*), and root mean square error (*RMSE*), as follows:

$$R^2 = \left(\frac{\sum_{i=1}^{N}(S_i^o - \bar{S}^o)(S_i^p - \bar{S}^p)}{\sqrt{\sum_{i=1}^{N}(S_i^o - \bar{S}^o)^2 \sum_{i=1}^{N}(S_i^p - \bar{S}^p)^2}} \right)^2 \quad (10)$$

$$NS = 1 - \frac{\sum_{i=1}^{N}(S_i^o - S_i^p)^2}{\sum_{i=1}^{N}(S_i^o - \bar{S}^o)^2} \quad (11)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(S_i^o - S_i^p)^2} \quad (12)$$

Here, S_i^o and S_i^p denote the observed suspended sediment values at ith time step, and the corresponding simulated suspended sediment values, respectively. *N* stands for the number of time steps. $\bar{S}^o$ and $\bar{S}^p$ represent the mean observed and simulated values, respectively.

R^2 provides information for linear dependence between observations and corresponding simulated values and should not be applied alone as a performance criterion. R^2 suffers from limitations that make them poor measures of model performance. Although these statistics continue to be used to determine how well a model simulates the observed data, they nevertheless provide a biased view of the efficacy of a model (Legates and McCabe 1999). The coefficient of efficiency *NS* has been widely used to evaluate the performance of hydrologic models. Nash and Sutcliffe (1970) defined the coefficient of efficiency (NS) which ranges from minus infinity to 1.0, with higher values indicating better agreement. *NS* is the coefficient is used to indicate the relative assessment of the model performance in dimensionless measures .the optimal value of *NS* is equal to 1. The *NS* represents an improvement over the R^2 for model evaluation purposes in that it is sensitive to differences in the observed and model simulated means and variance. The *RMSE* represents the sample standard error of the differences between predicted values and observed values. The *RMSE* serves to aggregate the magnitudes of the errors in predictions for various times into a single measure of predictive power. It ranges between 0 and 1 which have the optimal value equal to zero.

1.6 Input selection

River flow and SSL data from three studied stations were applied in different architectures to build the input configurations to feed the heuristic models. The configurations were defined in two main groups:

Group1. Local SSL estimation: Different input configurations were constructed for all three stations, including various combinations of the river flow (*Q*) and SSL values of successive time steps. In this context, the partial auto correlation functions (PACF) of the river flow and SSL data were evaluated for choosing the best time lag which comprises the best memory of time series. Figure 3 displays the PACF graphs of *Q* and *SSL* data for the studied stations. From the Figure it is seen that the first four time lags are influential in modeling time series. Therefore, the *Q* and *SSL* values of the current day as well as the corresponding values of three immediate previous days were introduced as input variables. Although a little discrepancy might be assumed for station2 data, where three time lags seem to be influential, the similar configurations were defined for all stations for consistency. Then, two scenarios were defined to introduce these values to the employed models (see Table 2):

Scenario I- Daily values of current and previously recorded *Q* values as well as the previously recorded values of *SSL* data were introduced.

Scenario II- Average values of the *Q* and *SSL* data corresponding to previous time steps were introduced.

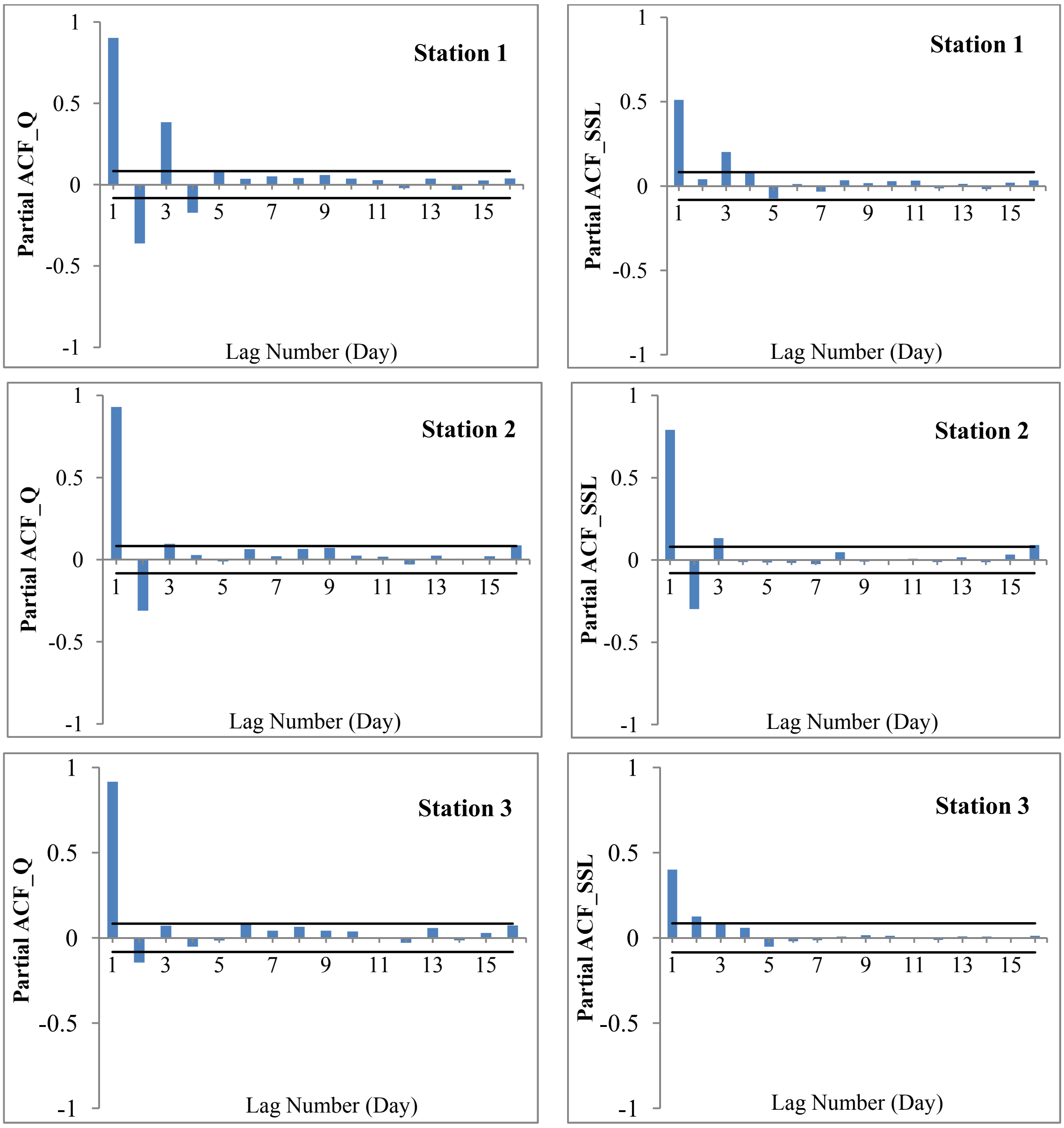

Figure 3 Partial auto correlation functions (PACF) for the river discharge (*Q*)and suspended sediment load(SSL) data of station 1, station 2 and station 3.

Group2. Cross-station SSL estimation: The SSL values of the downstream stations were assumed as a function of the SSL and *Q* values of the upstream station. Consequently, the input patterns from upstream station were used as inputs for estimating SSL at downstream station (ancillary data application), through the following data input selection typologies:

Typology a- the SSL values of the downstream station was assumed as a function of SSL and *Q* values of the upstream station(s), simultaneously (the effect of time lags were omitted).

Typology b- the best input configurations of the local application scenario (Producing the most accurate prediction results)

Table 3 gives the input configurations used to feed the models in cross-station application.

1.7 Models implementation

At the first step, both the local and cross-station application groups will be simulated using GEP to identify the best input configurations. Although PACF technique was used to detect the most influential input parameters on the SSL, the technique describes the linear dependency, so any nonlinear manner would be ignored leading to picking unsuitable features, and increasing possible redundancy. So, the GEP capability will be used to the same inputs, in order to select the most appropriate features which have the maximum relevance with the target values. Once the GEP-based SSL estimation models were established and analyzed, the important inputs will be identified from the global input matrix. Then, the best input configurations will be used to feed the SVM. Finally, the results will be compared and discussed. Figure 4 represents the flowchart of the present study.

Table 2 Input configurations used to feed the models-local application

Scenario	Input configurations
I	1. Q_t
	2. Q_t, Q_{t-1}
	3. Q_t, Q_{t-1}, Q_{t-2}
	4. Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}
	5. S_{t-1}
	6. S_{t-1}, S_{t-2}
	7. S_{t-1}, S_{t-2}, S_{t-3}
	8. Q_t, S_{t-1}
	9. Q_t, Q_{t-1}, S_{t-1}
	10. Q_t, Q_{t-1}, Q_{t-2}, S_{t-1}
	11. Q_{t-1}, Q_{t-2}, S_{t-1}
II	1. $\bar{Q}_{t-1,t-2}{}^{*}$
	2. $\bar{S}_{t-1,t-2}{}^{**}$
	3. $\bar{Q}_{t-1,t-2}$, $\bar{S}_{t-1,t-2}$

Note: Q_t: river flow at *i th* time step; S_i: SSL at *i th* time step, $^{*}\bar{Q}_{t-1,t-2}=\frac{Q_{t-1}+Q_{t-2}}{2}$, $^{**}\bar{S}_{t-1,t-2}=\frac{S_{t-1}+S_{t-2}}{2}$

Table 3 Input configurations of the cross-station application

No	Target station	Input variables
Typology-a		
1	station 2	$Q_1(t)$
2	station 2	$S_1(t)$
3	station 2	$Q_1(t)$, $S_1(t)$
4	station 3	$Q_1(t)$
5	station 3	$S_1(t)$
6	station 3	$Q_1(t)$, $S_1(t)$
7	station 3	$Q_2(t)$
8	station 3	$S_2(t)$
9	station 3	$Q_2(t)$, $S_2(t)$
10	station 3	$Q_1(t)$, $Q_2(t)$
11	station 3	$S_1(t)$, $S_2(t)$
12	station 3	$Q_1(t)$, $Q_2(t)$, $S_1(t)$, $S_2(t)$
Typology-b		
The best input configurations of local application scenario		

Note: $Q_1(t)$: Q value of the Station1 at the current time step; $S_1(t)$: SSL value of the Station1 at the current time step; $Q_2(t)$: Q value of the Station2 at the current time step; $S_2(t)$: SSL value of the Station2 at the current time step.

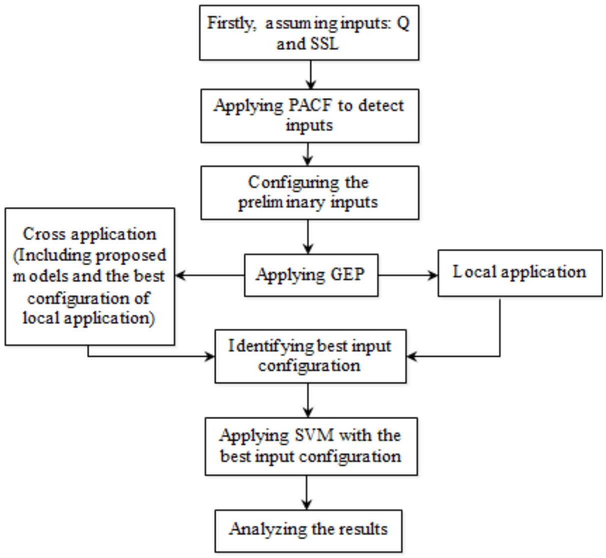

Figure 4 The flowchart of study.

2 Result and Discussion

2.1 Local application

The error statistics of the GEP and SRC models during the test period are given in Table 4. Among the scenario-I models, the Q-based and SSL-based models generally produces low accuracy results in all three stations. In Q-based context, the first model (model-1) generally shows high error magnitudes which might be due to the hysteresis effects. However, introducing the antecedent river flow values improves the models' accuracy. The most accurate results of Q-based models are obtained when the river flow values of the current as well as 1- and 2- previous days values are introduced as input variables in the first and second stations. But, the model including current and 1-previous day river flow values provides better results than the other Q-based models for the third station. The results show the adverse effect of introducing the Q_{t-3} values, which might be explained by cross-correlation effects of the river

Table 4 Testing statistics of the GEP and SRC models- local application (The best input configurations were marked in bold)

Method	Scenario	Input configurations	*Station 1*			*Station 2*			*Station 3*		
			R^2	*NS*	*RMSE*	R^2	*NS*	*RMSE*	R^2	*NS*	*RMSE*
GEP	I	1. Q_t	0.726	0.634	4.23	0.879	0.757	26.098	0.834	0.745	40.855
		2. Q_t, Q_{t-1}	0.760	0.75	4.139	0.845	0.772	20.124	0.956	0.783	36.954
		3. Q_t, Q_{t-1}, Q_{t-2}	0.819	0.817	3.124	0.834	0.847	17.052	0.642	0.625	40.789
		4. Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}	0.776	0.527	10.283	0.884	0.825	17.373	0.648	0.353	62.144
		5. S_{t-1}	0.818	0.814	4.435	0.986	0.864	16.94	0.904	0.739	35.312
		6. S_{t-1}, S_{t-2}	0.756	0.735	5.12	0.721	0.651	28.297	0.638	0.4	46.069
		7. S_{t-1}, S_{t-2}, S_{t-3}	0.812	0.729	5.137	0.933	0.842	19.32	0.737	0.719	32.853
		8. Q_t, S_{t-1}	0.755	0.64	5.185	0.868	0.704	19.744	0.839	0.789	32.66
		9. Q_t, Q_{t-1}, S_{t-1}	**0.912**	**0.909**	**1.449**	**0.944**	**0.929**	**14.403**	**0.990**	**0.98**	**17.361**
		10. Q_t, Q_{t-1}, Q_{t-2}, S_{t-1}	0.839	0.723	4.01	0.713	0.664	24.221	0.944	0.937	30.094
		11. Q_{t-1}, Q_{t-2}, S_{t-1}	0.916	0.786	3.958	0.918	0.863	18.651	0.9203	0.673	39.881
	II	1. $\bar{Q}_{t-1,t-2}$	0.633	0.611	5.761	0.572	0.47	46.089	0.491	0.469	47.895
		2. $\bar{S}_{t-1,t-2}$	0.560	0.45	13.25	0.660	0.591	33.411	0.819	0.501	47.707
		3. $\bar{Q}_{t-1,t-2}$,$\bar{S}_{t-1,t-2}$	**0.879**	**0.837**	**1.86**	**0.891**	**0.883**	**16.259**	**0.941**	**0.94**	**29.419**
SRC		Q_t	0.607	0.456	5.17	0.838	0.397	290.43	0.731	0.551	54.24

Note: R^2 represents coefficient of determination; *NS* means Nash- Sutcliffe coefficient; *RMSE* means Root Mean Square Error.

flow on SSL transport rate. Although cross correlation functions (not presented here) confirm the interdependence of the *Q* and *SSL* values up to 7 time lags, this correlation has linear inherent, so the nonlinear relation between *Q* and *SSL* values might not be explained by through this linear process. Nevertheless, the relative fluctuations of SSL and river flow data has a power function nature, so the corresponding time variations may display different manner for extreme events. This might cause time shift in occurrence of high- and low- sediment flows in relation to river flow variations. Attending to the SSL-based models, the most accurate results correspond to the first model, where only the *SSL* values of one immediate previous day were used to feed the model. Unlike to the PACF output, no direct link is seen between the current SSL values and the intervals beyond the *t-1* time step. This may remember the stochastic nature of the SSL transport phenomenon as well as failure of auto- regressive modeling tool in this regard. Comparing the performance of the GEP models based on the simultaneous using the *Q* and *SSL* data, the model comprising the Q_t, Q_{t-1}, and S_{t-1} values as input variables (GEP-I-9, in Table4) shows the most accurate results in all studied stations. Nonetheless, its performance is highest among the scenario I models, with the lowest *RMSE* value of 1.449, 14.403, and 17.361 ton/day, for stations 1, 2, and 3, respectively. Attending to the models built up with scenario-II configurations, the GEP-II-3 gives the most accurate results among other three models. Except GEP-I-9 model, the GEP-II-3 outperforms all the models of the both scenarios. Thus, the GEP-I-9 and GEP-II-3 models could be introduced as the best models of local applications according to performance criteria (*RMSE* and *NS*) .Beside this might show the amalgamated influence of the *Q* and *SSL* data on modeling process, the mathematic average of the parameters seems to provide a preliminary sufficient insight about the phenomenon to be learnt with heuristic model. As could be foreshadowed, introducing the sole averaged river flow or SSL values can't produce a high-accuracy results when learning with data driven technique, which is mainly due to the difficulties exist for mapping the nonlinear behavior of the SSL data.

Also, as can be seen from Table 4, the performance of the SRC method is noticeably lower in the modeling of the complex nonlinear relation between the river flow and the suspended sediment with respect to the GEP. The fitted SRCs and their proposed formulations are shown in Figure 5 for three stations.

2.2 Cross station application

Table 5 sums up the statistical indices of the GEP-based SSL models for cross station application. The models of *typology-a* show the input configurations which were proposed in *input selection* section. Also, the models of *typology-b*

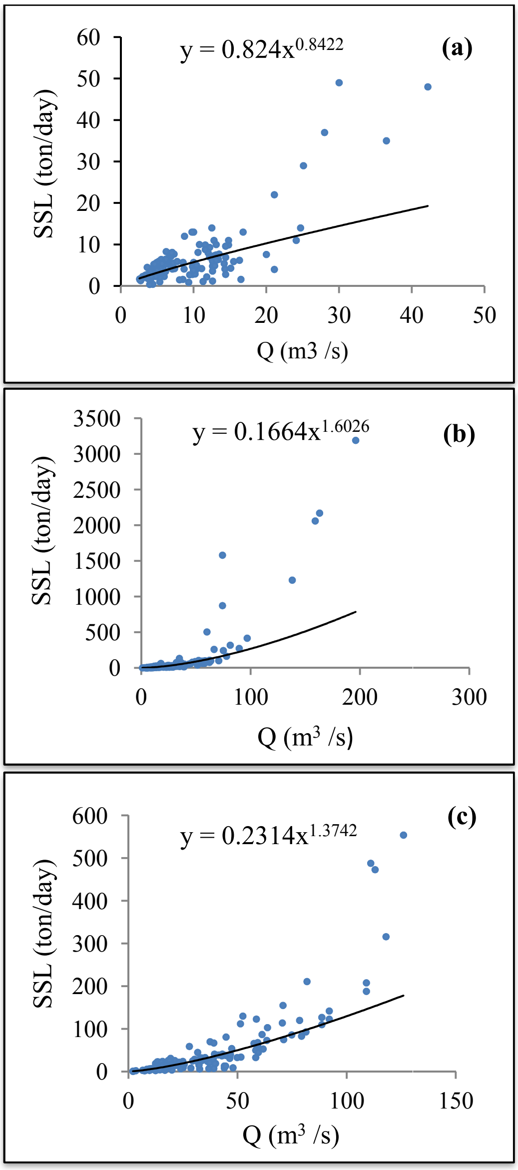

Figure 5 Applied SRC for the station 1 (a), station 2 (b) and station 3 (c).

Table 5 Testing statistics of the GEP models- cross station application. (The best input configurations were marked in bold)

No.	Targets	Inputs	R^2	*NS*	*RMSE* (ton/day)
Typology-a:					
1	$\boldsymbol{S_2(t)}$	$\boldsymbol{Q_1(t)}$	**0.954**	**0.949**	**18.85**
2	$S_2(t)$	$S_1(t)$	0.570	0.56	32.339
3	$S_2(t)$	$Q_1(t)$, $S_1(t)$	0.878	0.782	31.88
4	$S_3(t)$	$Q_1(t)$	0.968	0.571	118.514
5	$S_3(t)$	$S_1(t)$	0.761	0.726	78.784
6	$S_3(t)$	$Q_1(t)$, $S_1(t)$	0.807	0.755	73.987
7	$\boldsymbol{S_3(t)}$	$\boldsymbol{Q_2(t)}$	**0.880**	**0.874**	**41.428**
8	$S_3(t)$	$S_2(t)$	0.748	0.746	75.385
9	$S_3(t)$	$Q_2(t)$, $S_2(t)$	0.914	0.854	42.562
10	$S_3(t)$	$Q_1(t)$, $Q_2(t)$	0.644	0.296	130.543
11	$S_3(t)$	$S_1(t)$, $S_2(t)$	0.854	0.736	85.544
12	$S_3(t)$	$Q_1(t)$, $Q_2(t)$, $S_1(t)$, $S_2(t)$	0.696	0.523	120.345
Typology-b:					
1	$S_2(t)$	$Q_1(t)$, $Q_1(t-1)$,$S_1(t-1)$	0.811	0.802	31.23
2	$\boldsymbol{S_2(t)}$	$\boldsymbol{\bar{Q}_{1(t-1,t-2)}}$,$\boldsymbol{\bar{S}_{1(t-1,t-2)}}$	**0.904**	**0.851**	**24.143**
3	$S_3(t)$	$Q_1(t)$, $Q_1(t-1)$, $S_1(t-1)$	0.810	0.639	88.561
4	$S_3(t)$	$\bar{Q}_{1(t-1,t-2)}$,$\bar{S}_{1(t-1,t-2)}$	0.876	0.821	50.872
5	$S_3(t)$	$Q_2(t)$, $Q_2(t-1)$, $S_2(t-1)$	0.744	0.662	86.956
6	$\boldsymbol{S_3(t)}$	$\boldsymbol{\bar{Q}_{2(t-1,t-2)}}$,$\boldsymbol{\bar{S}_{2(t-1,t-2)}}$	**0.924**	**0.89**	**39.586**

are the best models of local application which were taken from Table 4 (GEP-I-9 and GEP-II-3 input configurations). From the Table 5 it is seen that the most accurate result for *typology-a* models correspond to SSL modeling of the station 2 when the current *Q* value of the upstream station is introduced as input variable (*RMSE* = 18.85 ton/day, *NS*=0.949). Also, the same result could be observed for station 3 where the model with the current *Q* value of the station 2 gives the most precise SSL value estimation (*RMSE* = 41.428 ton/day, *NS*=0.874).The comparison of the results of models 4-9 shows that the station 2 has more effective role than station 1 in estimation of SSL values for station 3. Although the *RMSE* values seem to be high for these models, comparing the *RMSE* with average values of SSL for test patterns (Table 1) reveals that the dimensionless *RMSE* values (the ratio between *RMSE* and average SSL value) are, respectively, 0.180 and 0.330 for S_2(t) and S_3(t) estimation, describing the high-accuracy performance of GEP-based models in these cases. Nevertheless, the systematic *RMSE* values (not presented here) are lower than unsystematic values, exhibiting the higher accuracy of the applied models. In terms of typology-b, combinations of $\bar{Q}_{(t-1,t-2)}$and $\bar{S}_{(t-1,t-2)}$ from upstream stations seem to be the best model for both station 2 and 3. Attending to the *typology-b*, where the best input configurations of the local application are used with ancillary data supplementation, the overall accuracy of the models are higher for this typology, than those of the first typology. This might be attributed to inclusion of the previously recorded Q and SSL values to models' architecture.

Comparing among the stations, when the SSL values of station2 [i.e. S2(t)] are modeled using the Q1(t) values, the error value is much lower than the model corresponded to the S3(t) simulation using Q2(t) values (18.850 vs. 41.428 ton/day, Owing that the stations are located with equal distances, such discrepancy might be explained by the natural catchment characteristics affecting the SSL value between stations 2-3.

Totally, although the prediction results of cross application are generally inferior than those of local application, the cross application seems to be applicable for the stations with lack of data.

2.3 Best models

One of the strong points of using GEP is in producing explicit formulations (i.e. model expression) of the relationship that rules the physical phenomenon. Such expressions may be subject to some physical interpretations. Table 6 summarizes the GEP expressions of the best models for local and cross-station approaches. It should be noted that the presented formulas only are valid in their under studied cases. According to Table 6, the GEP formulation is undefined in the case of scenario I (station 3) for Q_{t-1} =0 or Q_t=0values.

In order to assess the GEP performance in relation to other heuristic models, the SVM technique was applied with the best input configurations of local and cross-station

Table 6 GEP expressions for the best input configurations of local and cross-station applications

Applied Scenario	Inputs	target	GEP expressions
Local application			
Station1:			
Scenario-I	Q_t, Q_{t-1}, S_{t-1}	$S_1(t)$	$S_t=\left[\frac{\log(S_{t-1})}{\cos(Q_{t-1})+S_{t-1}+5.27Q_{t-1}}\right]*\left[Q_t^2+Q_t(12+\sqrt{S_{t-1}})\right]*\left[\tan^{-1}\left(\frac{S_{t-1}^{4/3}+Q_t-Q_{t-1}}{Q_{t-1}}\right)\right]$
Scenario-II	$\bar{Q}_{(t-1,t-2)}$,$\bar{S}_{(t-1,t-2)}$	$S_1(t)$	$S_t = \left[\frac{3.229\bar{S}_{(t-1,t-2)}}{3.229+\bar{S}_{(t-1,t-2)}}+(\bar{Q}_{(t-1,t-2)})+2.522-\sin(\bar{Q}_{(t-1,t-2)})\right] * \left[\frac{\sqrt{\frac{5.282\bar{S}_{(t-1,t-2)}}{\bar{Q}_{(t-1,t-2)}}}}{6.282\bar{S}_{(t-1,t-2)}+\bar{Q}_{(t-1,t-2)}}\right] * \left[2(\bar{S}_{(t-1,t-2)})-\sin\left(\left(log\left(\bar{Q}_{(t-1,t-2)}\right)\right)^2\right)\right]$
Station2:			
Scenario-I	Q_t, Q_{t-1}, S_{t-1}	$S_2(t)$	$S_t=\left[1.743Q_t^{1/3}*(Q_t-Q_{t-1}-10.83)\right]+\left[1.128S_{t-1}^{1/3}(2Q_t-Q_{t-1}-1.441)\right]+\left[(29.774-Q_t)*\text{atan}\,(S_{t-1}+7.346-Q_t)\right]$
Scenario-II	$\bar{Q}_{(t-1,t-2)}$,$\bar{S}_{(t-1,t-2)}$	$S_2(t)$	$S_t=\left[\frac{0.086\bar{S}_{(t-1,t-2)}}{0.103-\bar{Q}_{(t-1,t-2)}}\right]+\left[(1.174\bar{S}_{(t-1,t-2)}-\bar{Q}_{(t-1,t-2)})*cos(\bar{Q}_{(t-1,t-2)}-7.149)^3\right]+\left[\bar{S}_{(t-1,t-2)}+(\bar{S}_{(t-1,t-2)}*cos\left(\frac{\frac{Q_{t-1}+Q_{t-2}}{2}}{8.987}\right)^3)\right]$
Station3:			
Scenario-I	Q_t, Q_{t-1}, S_{t-1}	$S_3(t)$	$S_t=\left[\frac{Q_t^3}{(\sin(S_{t-1})+0.931Q_{t-1})^2}\right]+\left[\left(\frac{9.974Q_t*\text{atan}\,(Q_{t-1}-Q_t)}{Q_{t-1}}\right)-Q_{t-1}\right]+\left[\frac{(S_{t-1}-6.607)*(Q_t^2-4.296S_{t-1})}{Q_{t-1}Q_t}\right]$
Scenario-II	$\bar{Q}_{(t-1,t-2)}$,$\bar{S}_{(t-1,t-2)}$	$S_3(t)$	$S_t=\left[\frac{\sqrt{\bar{S}_{(t-1,t-2)}}-9.989+\bar{S}_{(t-1,t-2)}}{\sqrt{\bar{Q}_{(t-1,t-2)}}-\tan^{-1}(\bar{S}_{(t-1,t-2)})}\right] + \left[1.517(\bar{S}_{(t-1,t-2)}*\bar{Q}_{(t-1,t-2)})^{1/3}+0.504\bar{S}_{(t-1,t-2)}\right]+\left[(1.776-0.196\bar{S}_{(t-1,t-2)})*\sin(\log(\bar{S}_{(t-1,t-2)})^3)\right]$
Cross station application			
Station 2:			
Typology-a	$Q_1(t)$	$S_2(t)$	$S_t=\left[\frac{Q_1^3(t)*\sin(-5.081Q_1(t))}{921.491}\right]+\left[0.855\cos\left(\sqrt{Q_1(t)}\right)*(-3.666-Q_1(t))\right]+\left[\frac{2Q_1^3(t)}{247.348}-Q_1(t)\right]$
Typology-b	$\bar{Q}_{1(t-1,t-2)}$,$\bar{S}_{1(t-1,t-2)}$	$S_2(t)$	$S_t = \left[-0.924+\sqrt{log\left(\sin(\bar{S}_{1(t-1,t-2)})+7.52\right)}\right] * \left[(\bar{Q}_{1(t-1,t-2)})^2-21.442+2.505\bar{S}_{1(t-1,t-2)}+104.465\right] * \left[cos(sin(1.199\bar{S}_{1(t-1,t-2)}))\right]$
Station 3:			
Typology-a	$Q_2(t)$	$S_3(t)$	$S_t=\left[\frac{Q_{2(t)}^2-Q_{2(t)}}{13.796}+\cos\left(\frac{Q_{2(t)}}{6.55}\right)\right]+\left[(11.489-1.598Q_{2(t)})*\sin\left(\log\left(Q_{2(t)}^2\right)\right)\right]+\left[(Q_{2(t)}-15.018)*\sin\left(5.278Q_{2(t)}^2\right)\right]$
Typology-b	$\bar{Q}_{2(t-1,t-2)}$,$\bar{S}_{2(t-1,t-2)}$	$S_3(t)$	$S_t=\left[\bar{S}_{2(t-1,t-2)}-10.469+\left(\frac{(\bar{S}_{2(t-1,t-2)})^{0.33}}{1.25}\right)*\bar{Q}_{2(t-1,t-2)}\right]+0.376*\left[(\bar{Q}_{2(t-1,t-2)})^2-\bar{S}_{2(t-1,t-2)}\right]+\left[\frac{\bar{Q}_{2(t-1,t-2)}}{12.359}-2.739\right]^2$

Table 7 Comparison of the GEP and SVM models for the best input configurations of local and cross-station applications (The best models were marked in bold for the stations)

Models		Target	Input Configurations	SVM					GEP	
				NS	RMSE (ton/day)	C	ε	σ	NS	RMSE (ton/day)
Local application	***I***	$S_1(t)$	Q_t, Q_{t-1}, S_{t-1}	**0.923**	**1.365**	5	0.01	5	**0.909**	**1.449**
	II	$S_1(t)$	$\bar{Q}_{(t-1,t-2)}$,$\bar{S}_{(t-1,t-2)}$	0.864	1.78	9	0.01	2	0.837	1.86
	III	$S_2(t)$	Q_t, Q_{t-1}, S_{t-1}	0.813	18.312	10	0.01	0.33	0.704	19.744
	IV	$S_2(t)$	$\bar{Q}_{(t-1,t-2)}$,$\bar{S}_{(t-1,t-2)}$	0.914	15.14	10	0.01	0.5	0.883	16.259
	V	$S_3(t)$	Q_t, Q_{t-1}, S_{t-1}	**0.984**	**13.562**	10	0.01	0.5	**0.980**	**17.361**
	VI	$S_3(t)$	$\bar{Q}_{(t-1,t-2)}$,$\bar{S}_{(t-1,t-2)}$	0.956	27.35	10	0.01	1	0.940	29.419
Cross application	***VII***	$S_2(t)$	$Q_1(t)$	**0.951**	**14.369**	20	0.01	1	**0.949**	**18.85**
	VIII	$S_2(t)$	$\bar{Q}_{1(t-1,t-2)}$,$\bar{S}_{1(t-1,t-2)}$	0.895	17.196	5	0.01	0.5	0.851	24.143
	IX	$S_3(t)$	$Q_2(t)$	0.88	39.988	10	0.01	1	0.874	41.428
	X	$S_3(t)$	$\bar{Q}_{2(t-1,t-2)}$,$\bar{S}_{2(t-1,t-2)}$	0.931	31.741	20	0.01	0.9	0.89	39.586

applications and the testing results are reported in Table 7. Furthermore, Table 7 shows the optimal values of SVM parameters (C, ε and σ) which are effected by inherent characteristics of available data set. Comparison between ε and C values indicates that the loss function (C) has a wider range than ε for all cases. The table clearly shows the SVM superiority over the GEP technique in modeling SSL in the both local and cross-station scales. The results obtained demonstrated the GEP and SVM capabilities in modeling SSL values in the both local and cross-station approaches.

Figure 6 shows the log-scaled scattered plots of observed and predicted SSL values of GEP and SVM models which were marked in bold as the best models for the stations in Table 7. As can be seen from Figure 6 both GEP and SVM show better

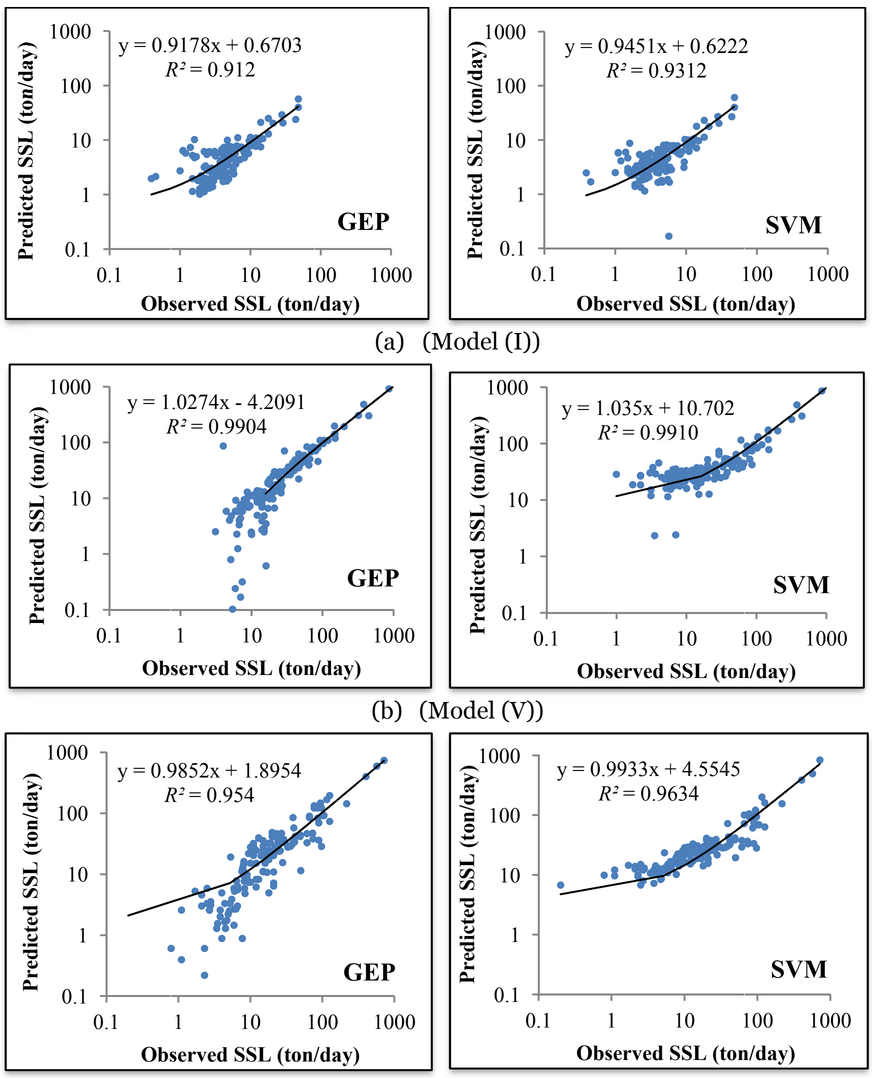

Figure 6 The log-scaled scatter plots of the best models using Gene Expression Programming (GEP) and Support Vector Machine (SVM) approaches for three stations (test data set).

performance in low SSL values than in high rates of SSL transport. It seems SSL transport to be more complicated in high rate than in low transport rate. Therefore, it is recommended that SSL modeling is divided into different ranges of SSL values according to appropriate hydraulic criteria, and more effective parameters should be considered in studies to achieve acceptable prediction accuracy.

3 Conclusion

The present study evaluated the ability of GEP and SVM approaches for modeling flow discharge-SSL relationship in three successive hydrometric stations of Housatonic River in USA. First, the local and cross-station data management strategies were used with different input configurations through GEP application. The selection of the input parameters was based on the analysis of the statistical correlation between the studied parameters, i.e. flow discharge and SSL. The results of the local application revealed that the model comprising the current flow discharge (Q_t), one previous day flow discharge (Q_{t-1}) and one previous day SSL (S_{t-1}) inputs, presents the highest accuracy in all three stations. In the cross station application, introducing the current flow discharge of the upstream station ($Q_1\,(t)$) as the unique input parameter provides the best SSL predictions for the downstream (second) station. Nonetheless, the simultaneous introduction of the ($\bar{Q}_{2(t-1,t-2)}$) and ($\bar{S}_{2(t-1,t-2)}$) as the model input parameters gives the most accurate results among others, for predicting SSL value of the third station. Obtained results revealed that the cross station application can be utilized as a reliable scenario to predict SSL values for the stations with the lack of the observational data. Then, the best input configurations which have been identified by using GEP for the both local and cross-station applications were used by SVM. The comparison of results clearly showed the superiority of SVM approach over the GEP for SSL modeling. Finally, the conventional SRC method was compared with GEP and SVM. The results illustrated that the traditional SRC model cannot surpass the applied GEP and SVM approaches.

The present study employed data from three successive stations. Further studies might be necessary for validating the obtained conclusions for different stations. The GEP formulas given here might be used to any specified site with similar conditions, e.g. SSL variations. Nonetheless, further researches might be necessary for assessing the GEP capabilities for externally calibrated models, i.e. when the train-test set belong to different stations.

Acknowledgement

Authors would like to express their gratitude to University of Tabriz for their supports during the present study.

References

Alp M, Cigizoglu HK (2007) Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. Environmental Modeling & Software 22(1): 2-13. DOI: 10.1016/j.envsoft.2005.09.009

Asselman NEM (2000) Fitting and interpretation of sediment rating curves. Journal of Hydrology 234(3): 228-248. DOI: 10.1016/S0022-1694(00)00253-5

Aytek A, Kişi Ö (2008) A genetic programming approach to suspended sediment modeling. Journal of Hydrology 351(3): 288-298. DOI: 10.1016/j.jhydrol.2007.12.005

Azamathulla HM, Ghani AA ,Chang CK, et al. (2010) Machine learning approach to predict sediment load–a case study. CLEAN–Soil, Air, Water 38(10): 969-976. DOI: 10.1002/clen.201000068

Azamathulla HM, Ghani AA, Leow CS, et al. (2011) Gene-expression programming for the development of a stage-discharge curve of the Pahang River. Water Resources Management 25(11): 2901-2916. DOI: 10.1007/s11269-011-9845-7

Basak D, Pal S, Patranabis DC (2007) Support Vector Regression. Neural Information Processing-Letters and Reveiews 11(10): 203-224.

Chiang JL, Tsai YS (2011) Suspended sediment load estimate using support vector machines in Kaoping river basin. In Consumer Electronics, Communications and Networks (CECNet), 2011 International Conference on (pp. 1750-1753). IEEE. DOI: 10.1109/CECNET.2011.5769267

Cohn TA, Caulder DL, Gilroy EJ, et al. (1992) The validity of a simple statistical model for estimating fluvial constituent loads: An empirical study involving nutrient loads entering Chesapeake Bay. Water Resources Research 28(9): 2353-2363. DOI: 10.1029/92WR01008

Crowder DW, Demissie M, Markus M (2007) The accuracy of sediment loads when log-transformation produces nonlinear sediment load–discharge relationships. Journal of Hydrology 336(3): 250-268. DOI: 10.1016/j.jhydrol.2006.12.024

Ferreira C (2001) Gene expression programming: a new adaptive algorithm for solving problems. Complex Systems 13 (2): 87-129. DOI: 10.1016/j.jhydrol.2006.12.024

Ferreira C (2006) Gene expression programming: mathematical Modeling by an artificial intelligence. Springer, Berlin, Heidelberg, New York, p. 478.

Forman SL, Pierson J, Lepper K (2000) Luminescence geochronology. Quaternary geochronology: Methods and applications: 157-176.

Gunn SR (1998) Support vector machines for classification and regression. ISIS technical report, 14.

Lafdani EK, Nia AM, Ahmadi A (2013) Daily suspended sediment load prediction using artificial neural networks and support vector machines. Journal of Hydrology 478: 50-62. DOI: 10.1016/j.jhydrol.2012.11.048

Kisi O, Cimen M (2011) A wavelet-support vector machine conjunction model for monthly stream flow forecasting. Journal of Hydrology 399(1): 132-140. DOI: 10.1016/j. jhydrol.2010.12.041

Kisi O, Shiri J (2011) Precipitation forecasting using wavelet-genetic programming and wavelet-neuro-fuzzy conjunction models. Water Resources Management 25(13): 3135-3152. DOI: 10.1007/s11269-011-9849-3

Kisi O (2012) Modeling discharge-suspended sediment relationship using least square support vector machine. Journal of Hydrology 456: 110-120. DOI: 10.1016/ j.jhydrol. 2012.06.019

Kisi O, Dailr AH, Cimen M, et al. (2012) Suspended sediment modeling using genetic programming and soft computing techniques. Journal of Hydrology 450: 48-58. DOI: 10.1016/ j.jhydrol.2012.05.031

Kisi O, Shiri J (2012) River suspended sediment estimation by climatic variables implication: comparative study among soft computing techniques. Computers & Geosciences 43: 73-82. DOI: 10.1016/j.cageo.2012.02.007

Legates DR, McCabe GJ (1999) Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. Water Resources Research 35(1): 233-241. DOI: 10.1029/1998WR900018

Liu QJ, Shi ZH, Fang NF, et al. (2013) Modeling the daily suspended sediment concentration in a hyperconcentrated river on the Loess Plateau, China, using the Wavelet–ANN approach. Geomorphology 186: 181-190. DOI: 10.1016/ j.geomorph.2013.01.012

Nash J, Sutcliffe JV (1970) River flow forecasting through conceptual models part I—A discussion of principles. Journal of Hydrology 10(3): 282-290. DOI: 10.1016/0022-1694(70) 90255-6

Overleir A.P (2004) Accounting for heteroscedasticity in rating curve estimates. Journal of Hydrology 292(1): 173-181. DOI: 10.1016/j.jhydrol.2003.12.024

Partal T, Cigizoglu HK (2008) Estimation and forecasting of daily suspended sediment data using wavelet–neural networks. Journal of Hydrology 358(3): 317-331. DOI: 10.1016/j.jhydrol.2008.06.013

Picouet C, Hingray B, Olivry JC (2001) Empirical and conceptual modeling of the suspended sediment dynamics in a large tropical African river: the Upper Niger river basin. Journal of Hydrology 250(1): 19-39. DOI: 10.1016/ S0022-1694(01)00407-3

Roushangar K, Mehrabani FV, Shiri J (2014a) Modeling river total bed material load discharge using artificial intelligence approaches (based on conceptual inputs). Journal of Hydrology 514: 114-122. DOI: 10.1016/j.jhydrol.2014.03.065

Roushangar K, Mouaze D, Shiri J (2014b) Evaluation of genetic programming-based models for simulating friction factor in alluvial channels. Journal of Hydrology 517: 1154-1161. DOI: 10.1016/j.jhydrol.2014.06.047

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. DOI: 10.1016/j.jhydrol.2015.06. 006

Shiri J, KişI Ö (2011) Comparison of genetic programming with neuro-fuzzy systems for predicting short-term water table depth fluctuations. Computers & Geosciences 37(10): 1692-1701. DOI: 10.1016/j.cageo.2010.11.010

Shiri J, Kisi O (2012) Estimation of daily suspended sediment load by using wavelet conjunction models. Journal of Hydrologic Engineering 17(9): 986-1000. DOI: 10.1061/ (ASCE)HE.1943-5584.0000535

Shiri J, Sadraddini AA, Nazemi AH, et al. (2013) Evaluation of different data management scenarios for estimating daily reference evapotranspiration. Hydrology Research 44(6): 1058-1070. DOI: 10.2166/nh.2013.154

Shiri J, Marti P, Singh VP (2014) Evaluation of gene expression programming approaches for estimating daily evaporation through spatial and temporal data scanning. Hydrological Processes 28(3): 1215-1225. DOI: 10.1002/hyp.9669

Singh VP, Krstanovic PF, Lane L J (1988) Stochastic models of sediment yield. Modeling Geomorphological Systems. John Wiley and Sons New York. 1988. p 259-285, 7 fig, 3 tab, 64 ref.

Vapnik VN (1998) Statistical Learning Theory. Wiley, New York. p 768.

Walling DE (1977) Assessing the accuracy of suspended sediment rating curves for a small basin. Water Resources Research 13(3): 531-538. DOI: 10.1029/WR013i003p00531

Yang CT (1996) Sediment Transport: Theory and Practice. McGraw-Hill, New York ; Singapore. p 396.