Suspended sediment modelling by SVM and wavelet

Authors:

Maedeh Sadeghpour Haji
Department of environmental Engineering, Environment and Energy, Science and Research Branch
Islamic Azad University
Tehran, Iran
maedeh_sadeghpoor@yahoo.com

Prof. Seyed Ahmad Mirbagheri
K. N. Toosi University of Technology
Department of Civil and Environmental Engineering
mirbagheri@kntu.ac.ir

Assoc.prof. Amir H. Javid
Islamic Azad University
Science and Research Branch
ahjavid@gmail.com

Assoc.prof. Mostafa Khezri
Islamic Azad University
Science and Research Branch
khezri@sharif.edu

Prof. Ghasem D. Najafpour
Babol Noshirvani University of Tech.
Biotechnology Research Center
najafpour8@gmail.com

Scientific paper - Preliminary note

Maedeh Sadeghpour Haji, Seyed A. Mirbagheri, Amir H. Javid, Mostafa Khezri, Ghasem D. Najafpour

Suspended sediment modelling by SVM and wavelet

Present-day advances in artificial intelligence, as a forecaster for hydrological events, have led to numerous changes in forecasting. The wavelet support vector machine (WSWM) model is achieved by conjunction of the wavelet analysis and the support vector machine (SVM). The suspended sediment (SS) and daily stream flow (Q) data from the Iowa River in the USA were used for training and testing. The WSVM could logically be used for approximation of the suspended sediment load.

Key words:

Discrete wavelet analysis, Cumulative SS, Daily stream flow, High suspended sediment, Support vector machine

Prethodno priopćenje

Maedeh Sadeghpour Haji, Seyed A. Mirbagheri, Amir H. Javid, Mostafa Khezri, Ghasem D. Najafpour

Modeliranje lebdećeg nanosa pomoću SVM-a i valića

Napredak koji danas bilježimo u primjeni umjetne inteligencije za predviđanje hidroloških događaja doveo je do brojnih promjena u sferi predviđanja. Valićni model baziran na metodi potpornih vektora (WSVM) dobiven je spajanjem valićne analize i metode potpornih vektora (SVM). Za učenje i testiranje korišteni su podaci o lebdećem nanosu (SS) i dnevnom protoku (Q) izmjereni na rijeci Iowa u SAD-u. Provedene analize su pokazale da se valićni model WSVM može koristiti za aproksimaciju količine lebdećeg nanosa.

Ključne riječi:

diskretna valićna analiza, kumulativni lebdeći nanos, dnevni protok, visoke vrijednosti lebdećeg nanosa, metoda potpornih vektora (SVM)

Vorherige Mitteilung

Maedeh Sadeghpour Haji, Seyed A. Mirbagheri, Amir H. Javid, Mostafa Khezri, Ghasem D. Najafpour

Modellierung von Schwebstoffsedimenten mittels SVM und Wavelets

Fortschritte bezüglich der Anwendung künstlicher Intelligenz Vorhersagen hydrologischer Vorgänge haben zu zahlreichen Änderungen in diesem Bereich geführt. Das Modell der Wavelet-Support-Vektor-Maschine (WSVM) beruht auf der Verbindung von Wavelet Analysen und Support-Vektor-Maschinen (SVM). Zur Einführung und Prüfung sind Daten hinsichtlich der Schwebstoffmengen (SS) und des täglichen Stromflusses (Q) aus Messungen am Fluss Iowa in den USA angewandt worden. Das Modell WSVM kann daher für die Approximation von Schwebstoffmengen angewandt werden.

Schlüsselwörter:

diskrete Wavelet Analyse, kumulative Schwebstoffsedimente, täglicher Stromfluss, hohe Schwebstoffsedimente, Support-Vektor-Maschine

1. Introduction

In most rivers, main part of the sediment is transported in suspension. An increasing importance is given to correct sediment prediction, particularly in flood-prone areas [1]. The proper estimation of suspended sediments (SS) is essential for the operation and design of canals, dams and diversions. The forecasting of river suspended sediments is a complex work for operation of water resources in environmental engineering and hydrology [2]. A great number of unclear factors are involved in river SS modelling. Theoretical equations may not be adequate for gaining information on the whole procedure as they have restricted capability to capture non-linearities and non-stationarities in environmental and hydrologic data. Over the past decade a particular attention has been paid to the application of artificial intelligence in the field of environmental engineering and water resources management. Several studies have been made in order to develop artificial intelligence methods for modelling processes with limited knowledge [3-6]. Some trustworthy literature-based techniques for modelling water resources and river suspended sediment load such as wavelet, artificial neural network (ANN), and support vector machine (SVM), are presented in this paper.

The capability of ANNs to find nonlinear relations between inputs and outputs make them proper tools for modelling hydraulic and hydrological phenomena [7].The ANN simulation has been increasingly applied in many countries, and its use for time series modelling has recently been discovered. The wavelet-transformed data of observed time series enhance forecasting capabilities by capturing valuable information at different resolution levels [8]. The performance of multi layer feed forward (MLFF) network, and radial basis function (RBF) network, to forecast the discharge of suspended sediments has been compared [9]. The ANN employment in support of river SS prediction has been studied by many researchers [10, 11]. The ANN, neuro-fuzzy (NF), multi layer regression (MLR) and sediment rating curve (SRC) models were examined for the one day ahead simulation of SS in two hygrometry stations. It was established by comparison of modelling results that the NF model is better suited for the SS forecasting [12]. The ANN model was proposed as a means for simulating the monthly suspended sediment flux in China [9]. In the mentioned model, the suspended sediment flux was correlated with the temperature, average rainfall, rainfall intensity, and flow discharge. Results show that the ANN model is capable of simulating the monthly suspended sediment flux [13].

Other investigators have defined a model by combining the wavelet transform and the neuro-fuzzy (NF) technique to predict the daily suspended sediment [14]. The ANN approach was used to model the SS concentration on two sites on the Mississippi River. The corresponding results have revealed that the ANN technique is better when compared to conventional methods [15].

A new wavelet artificial neural network model was used for the daily SS forecast in the Yadkin River at the Yadkin College station in the USA. The wavelet transform was connected to an artificial neural network (ANN). The comparison of prediction accuracies of the wavelet artificial neural network (WANN) and other models revealed that the proposed WANN model was able to effectively predict the SS. The best model performance was achieved by Meyer wavelet with the determination coefficient value of 0.83 [16]. A combined neural-wavelet model was proposed for prediction of the Ligvanchai watershed precipitation at Tabriz, Iran. The main time series was decomposed into several multi-frequency time series by wavelet technique and imposed as input data to the ANN to predict the precipitation one month ahead [17]. A combined wavelet-ANN model was recommended for SS forecasting. The daily stream flow and SS data derived from the Iowa River station in the United States were employed to train and test the ANN, WANN, MLR and SRC models. According to results obtained, the WANN model performed better than other models [18].

The relation between the SS and stream flow discharge was identified by genetic algorithm (GA). This model had better performance than the SRC technique [19]. Another research highlights the usefulness of wavelet analysis and the power of this tool for studying time series in karst hydro-systems [20]. The wavelet analysis and NF was applied for daily SS forecasting [21]. The support vector machine (SVM) is a supervised learning technique that produces input-output mapping functions from a set of training data [22]. In training support vector machines the decision boundaries are directly determined from the training data. This learning strategy is based on statistical learning theory and minimizes the classification errors of the training data and the unknown data [23]. The SVM is also used for predicting the chlorophyll concentration in reservoirs [24]. The WSVR model is investigated as a means to predict monthly stream flows and daily precipitations. These models were developed by connecting two techniques, discrete wavelet analysis and support vector regression. The comparison results illustrate that the WSVR enhances the model performance [25, 26]. The SVM is used as a pattern-recognition predictor to simulate daily, weekly and monthly runoff and sediment yield from an Indian watershed [27]. Two input variable pre-processing methods for SVM model are explored, principal component analysis and the gamma test. The proposed methods provide a more truthful performance with regard to the monthly stream flow forecasting [28]. The least square support vector machine (LSSVM) is compared with artificial neural networks (ANNs) and sediment rating curve (SRC) in a separate prediction of upstream and downstream stations sediment data. The comparison results of the models show that the LSSVM model performs better than the ANN techniques [29]. The artificial neural network and support vector machine models were applied to predict the suspended sediment load

in Doiraj river basin situated in the west of Iran. The stream flow and rainfall data were considered as input variables and length of the training data set. The best input combination for the models was identified by the use of GT [30].

The genetic programming (GP) method was applied for estimating the daily SS in two stations on the Cumberland River in U.S. Results show that the GP is superior to other models [31].

These surveys show that the wavelet analysis is an effective method for irregularly distributed multi-scale features of climate elements in space and time. The goal of combining the wavelet analysis with the SVM technique is to improve the SS forecasting accuracy.

2. Support vector machine

New techniques on artificial intelligence, with a great variety of applications, have been identified over the past decade. One of them is the support vector machine (SVM) which is used in the classification and regression algorithms assortment. Support vector regresses, which are extensions of support vector machines, have shown good generalization ability for various function approximation and time series prediction problems [23]. Much research has so far been made about the theory of SVM [32-35]. Therefore, only a brief explanation of ε -SVM model is given. Suppose we are given training data $\{(x_1,y_1), ..., (x_l,y_l)\} \subset X \times R$, where x and y denote the space of input patterns and target values. The aim of the SVM is to find a function $f(x)$ that has the greatest ε deviation from actual targets y_i for all training data, and it should be as flat as possible. In this paper, the ε-SVM regression is conducted for prediction of next day SS [36]. In other words, errors are not considered as long as they are less than ε, but any deviation larger than ε will not be accepted. Linear functions of $f(x)$ are described as follows:

$$f(x) = \langle w, x \rangle + b, \quad w \in X, b \in R \tag{1}$$

where $\langle w, x \rangle$ denotes the dot product in X. Flatness in the case of Eqn (1) means that a small w is sought. One method to ensure w is to minimise the norm i.e. $\|w\|^2 = \langle w, w \rangle$. This problem can be written as a convex optimization problem:

Minimize $\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}\left(\zeta_i + \zeta_i^*\right)$

Subject to $\begin{cases} y_i - \langle w, x \rangle - b \leq \varepsilon \\ \langle w, x \rangle + b - y_i \leq \varepsilon \end{cases}$ (2)

The assumption in Eqn (2) is that such a function $f(x)$ that approximates all pairs (x_i, y_i) with ε precision actually exists, and that the convex optimization problem can be solved. On occasion, this may not be the case, or we may also allow for some errors. We can introduce slack variables ξ_i, ξ_i^* to cope with otherwise insoluble constraints of the optimization problem Eqn (3). Thus, the formulation can be described as follows:

Minimize $\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}\left(\zeta_i + \zeta_i^*\right)$

Subject to $\begin{cases} y_i - \langle w, x_i \rangle - b \leq \varepsilon + \xi_i \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases}$ (3)

The constant C > 0 determines the trade-off between the flatness of $f(x)$ and the amount up to which deviations larger than ε are tolerated. This corresponds to the situation with the so called ε–insensitive loss function $|\xi|_\varepsilon$ explained in Eqn (4) as follows:

$$|\zeta|_\varepsilon = \begin{cases} 0 & \text{if } |\xi| < \varepsilon \\ |\xi| - \varepsilon & \text{otherwise} \end{cases} \tag{4}$$

The optimization problem in Eqn (3) can be solved in its dual formulation. The major plan is to make a Lagrange function from the primal objective function and the corresponding constraints, by introducing a dual set of variables. It has a saddle point with respect to the primal and dual variables at the solution. This function is shown by the following expressions:

$$L := \frac{1}{2}\|w\|^2 + c\sum_{i=1}^{1}\left(\xi_i + \xi_i^*\right) - c\sum_{i=1}^{l}\left(\eta_i\xi_i + \eta_i^*\xi_i^*\right)$$

$$-\sum_{i=1}^{l}\alpha_i\left(\varepsilon + \xi_i - y_i + \langle w, x_i \rangle + b\right) \tag{5}$$

$$-\sum_{i=1}^{l}\alpha_i^*\left(\varepsilon + \xi_i^* + y_i - \langle w, x_i \rangle - b\right)$$

The term L is the Lagrangian and η_i, η_i^*, α_i, α_i^* are factors of Lagrange multipliers. Therefore, dual variables in Eqn (6) have to allow for constraints.

$$\alpha_i, \alpha_i^*, \eta_i, \eta_i^* \geq 0 \tag{6}$$

It follows from the saddle point condition that the partial derivatives of L with respect to the primal variables (w, b, ξ_i, ξ_i^*) have to vanish for optimality.

$$\frac{\partial L}{\partial b} = \sum_{i=1}^{l}\left(\alpha_i^* - \alpha_i\right) = 0$$

$$\frac{\partial L}{\partial w} = w - \sum_{i=1}^{l}\left(\alpha_i^* - \alpha_i\right)x_i = 0 \tag{7}$$

$$\frac{\partial L}{\partial \xi_i^*} = C - \left(\alpha_i^* - \eta_i^*\right) = 0$$

Substituting Eqn (7) into Eqn (5) yields the dual optimization problem and eliminates dual variables. Eqn (8) is rewritten as follows:

$$w = \sum_{i=1}^{l}\left(\alpha_i - \alpha_i^*\right)x_i, \text{ tako da je } f(x) = \sum_{i=1}^{l}\left(\alpha_i - \alpha_i^*\right)\langle x_i, x \rangle + b \tag{8}$$

This is the so-called support vector expansion. The w could be explained as a linear algorithm that only depends on dot products between training patterns w. Linear model is not suitable for numerous hydrological events. Consequently, it becomes suitable by converting Kernel for putting data in a space with more dimensions and then by applying the standard support vector regression algorithm. These interpretations will be convenient for the formulation of a nonlinear extension. This could be done by pre-processing the training patterns x_i by a map $\varphi : x \to \Im$ into some feature space $\Im$. Thus, it is sufficient to recognize Kernel function $k(x, x_i) := \langle \varphi(x), \varphi(x_i) \rangle$ rather than φ explicitly. The Kernel function provides the opportunity for using a non-linear function in input space for varying to linear function in characteristics space. Kernel function benefits to untreated high dimensional feature space explicitly. This technique is named Kernel trick and is shown as follows [22, 30, 36, 37]:

$$f(x) = \sum_{i=1}^{l} \left(\alpha_i - \alpha_i^* \right) k(x, x_i) + b \qquad (9)$$

The standard conversion of Kernel function, which is most often used in regression and modelling, is given in Table 1 [22].

Table 1. Kernel function most often used in regression and modelling

Kernel Functions	Type of classifier
$k(x, x_i) = (x^T x_i)$	Linear, dot product, Kernel, CPD
$k(x, x_i) = \left[(x^T x_i) + 1 \right]^d$	Complete Polynomial of degree d, PD
$k(x, x_i) = e^{-\gamma \|x - x_i\|^2}$	Gaussian RBF, PD
$k(x, x_i) = \tanh \left[(x^T x_i) + b \right]^*$	Multilayer perceptron, CPD
$k(x, x_i) = = \dfrac{1}{\sqrt{\|x - x_i\|^2 + \beta}}$	Inverse multiquadric function, PD
* only for certain values of b, (C) PD= (conditionally) positive definite	

The Radial Basis Function, i.e. a function based on Radial Basis Function more similar to Gaussian (bell shape), which is shown in Table 1 and dependent to γ variable, is taken into consideration in this paper because:
1. In contrast with linear Kernel, it can hold the case when the relation between class labels and attributes is nonlinear.
2. This Kernel function has a smaller amount of hyperparameters which control the difficulty of model selection.
3. It has less numerical complexities [38, 39].

3. Wavelet transform

3.1. The Continuous wavelet transform (CWT)

The theory of wavelet analysis is founded on the Fourier analysis [40]. The wavelet decomposition is perfect for considering transient signals and obtaining a superior characterization and a more dependable discrimination technique [41]. The signal is multiplied with a function close to the window function and the transform is calculated for diverse segments of the time-domain signal. The continuous wavelet transform is described as follows:

$$CWT_x^\psi (\tau, s) = \psi_x^\psi (\tau, s) = \frac{1}{\sqrt{|s|}} \int x(t) \psi^* \left(\frac{t - \tau}{s} \right) ds \qquad (10)$$

The wavelet transform is a function of two parameters, translation (τ) and scale (s). $\Psi(\tau)$ is named the mother wavelet. The term wavelet means a small wave. The smallness implies to the condition that window function is compactly supported. The word translation relates to the place of the window, as the window is shifted from beginning to the end of the signal. Scale parameter (s) which is defined as s = 1 / frequency [42].

3.2. Discrete Wavelet Transform (DWT)

The concept of DWT is similar to that of the continuous wavelet transform (CWT). The time-scale representation of a digital signal is achieved by digital filtering methods. The CWT is a correlation between a wavelet at diverse scales and the signal with the scale (or the frequency) being used as a measure of similarity. In the discrete case, filters of different cutoff frequencies are used to analyze the signal at diverse scales. The signal is approved throughout a series of high pass filters to analyze the high frequencies, and it is passed through a series of low pass filters to analyze the low frequencies. Sub-sampling by a factor n decreases the number of samples in the signal n times. Up-sampling a signal corresponds to enhancing the sampling rate of a signal by adding new samples to the signal; while the signal is a discrete time function, the terms function and sequence will be applied interchangeably. This sequence is indicated by $x[n]$, where n is an integer. The process begins with passing this signal through a half band lowpass filter with impulse response $h[n]$. Filtering a signal communicates to the numerical operation of convolution of the signal with the impulse response of the filter. The equation is defined as follows:

$$x[n] * h[n] = \sum_{k=-\infty}^{\infty} x[k] h[n - k] \qquad (11)$$

A half band lowpass filter eliminates the entire frequencies that are over half of the highest frequency in the signal. The lowpass filtering halves the resolution, but leaves the scale unchanged. The signal is then subsampled by 2 since one half of samples are unnecessary.
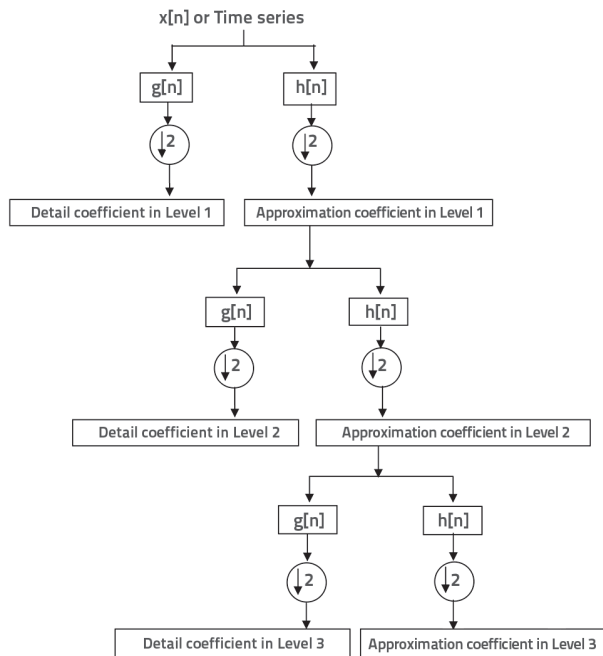
Figure 1. Original signal decomposition with lowpass and highpass filters

These procedures double the scale and could be expressed as follows:

$$y[n] = \sum_{k=-\infty}^{\infty} h[k]x[2n-k]$$ (12)

The DWT converts the signal into an approximation and detailed information. The original signal $x[n]$ is primary passed through a half band highpass filter $g[n]$ and a lowpass filter $h[n]$. These equations could be expressed as follows:

$$y_{high}[k] = \sum_n x[n]g[2k-n]$$ (13)

$$y_{low}[k] = \sum_n x[n]h[2k-n]$$ (14)

where $y_{high}[k]$ and $y_{low}[k]$ are the output of the highpass and lowpass filters, after subsampling by 2. The former system, which is identified as the subband coding, could be repeated for supplementary decomposition. The method is illustrated in Figure 1 where $x[n]$ is the original signal, and $h[n]$ and $g[n]$ are lowpass and highpass filters, respectively. The original signal decomposition in three levels is illustrated as follows [43]:

4. Gauging station and statistical analysis of data

The data collected for a period of 9 years (01-October-1978 to 30-September-1987) from the Iowa River at Wapello, NC gauging station (USGS station No: 05465500, Basin Area (sq. mi.): 12,499, longitude: 091°10'57" and latitude: 41°10'48"), operated by the U.S. Geological Survey (USGS), were applied for training and testing the models. Actually, uninterrupted time series data such as Q and SS were needed in modelling. Data from October 1, 1978 to September 30, 1985 (7 years), and the data from October 1, 1985 to September 30, 1987 (2 years), were used as training and validation data sets. The time series of data related to daily Q and SS are shown in Figure 2.

Statistical parameters for stream flow and suspended sediment such as maximum, minimum, mean, average and standard deviation (SD), are shown in Table 2. It is clear from the Table that the stream flow and suspended sediment data present spread distribution. Maximum experimental Q and SS were also used in the test period. These statistical parameters signify the vastly comprehensive performance of stream flow and suspended sediment phenomenon.

The selection of correct combination of pre-process input factors is one of complex procedures in modelling. The autocorrelation and cross-correlation between the Q and SS
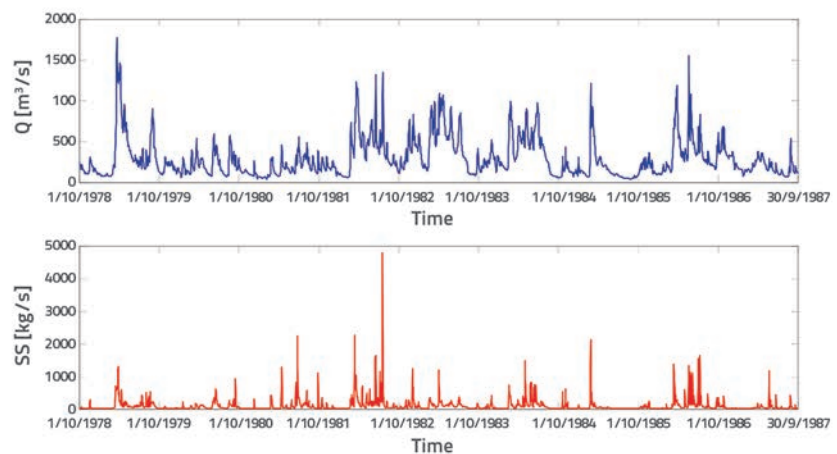

Figure 2. Stream flow and suspended sediment time series for a period of 9 years

Table 2. Statistic parameters for stream flow and suspended sediment

Data set	Period	Maximum	Minimum	Mean average	Standard deviation
Suspended sediments SS [kg/s]	Učenje	4780	0,06	96,65	206,94
	Testiranje	1631,94	0,85	86,91	181,65
	Sve	4780	0,06	94,49	201,6
Stream flow Q [m³/s]	Učenje	1770	39,6	292,78	247,11
	Testiranje	1550	64,3	274,32	194,90
	Sve	1770	39,6	288,69	236,62

data were investigated to predict the daily SS. This procedure was used in several research papers [14, 16, 18, 21, 26, 29, 30]. Cross-correlation coefficients, relating between observed SS_t and stream flow time series (Q_t, Q_{t-1}, Q_{t-2},...Q_{t-10}), and autocorrelation coefficients such as lag 1 day autocorrelation coefficient (R_1), lag 2 days autocorrelation coefficient (R_2), lag 3 days autocorrelation coefficient (R_3) .... and lag 10 days autocorrelation coefficient (R_{10}), are illustrated for Q and SS in Figure 3. It can be seen that autocorrelation coefficients of stream flow are higher than autocorrelation coefficients of suspended sediment in both training and testing data sets. The autocorrelation between SS_t and $SS_{t-4,t-5,t-6...}$ is quite low, and the same happens for cross-correlation coefficients between SS_t and $Q_{t-4,t-5,t-6...}$. Consequently, the models, whose input were the SS and Q of three previous days, were examined. We took different lag time series of stream flow and SS that included maximum three time steps into the past as input data.

The semilog scale scatter plots of this stations' data are given in Figure 4. It can be noticed that there is a nonlinear and complex relationship between the stream flow and suspended sediment. This Figure demonstrates the existence of an outlier in the data range. A SS concentration value of 4780 kg/s was observed, while other concentration values were below 2245.37 kg/s. These values were generated through extra complex calculation for the models.

A number of conventional evaluations such as the coefficient of determination (R^2), sum of square error (SSE), and root mean square error (RMSE), were considered [44, 45]. In this study, the performance of the models was evaluated by R^2, RMSE, mean of Error (Error mean), and Error Standard Deviation (Error STD). These equations are described as follows:

$$Error_i = SS_{i(measured)} - SS_{i(predicted)} \quad (15)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(Error_i\right)^2}{\sum_{i=1}^{n}\left(SS_{i(measured)} - SS_{i(mean)}\right)} \quad (16)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(Error_i\right)^2}{n}}$$

$$ErrorSTD = \sqrt{\frac{\sum_{i=1}^{n}\left(Error_i - Error_{i(mean)}\right)^2}{n}} \quad (17)$$

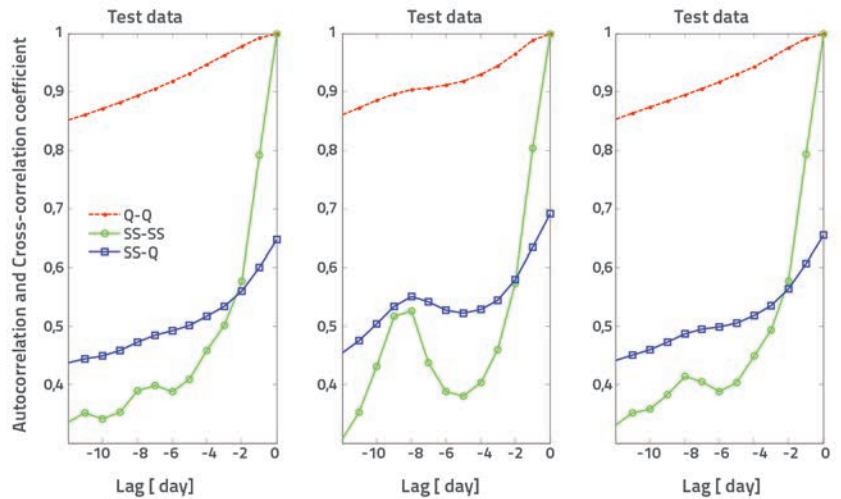where n is the number of data.

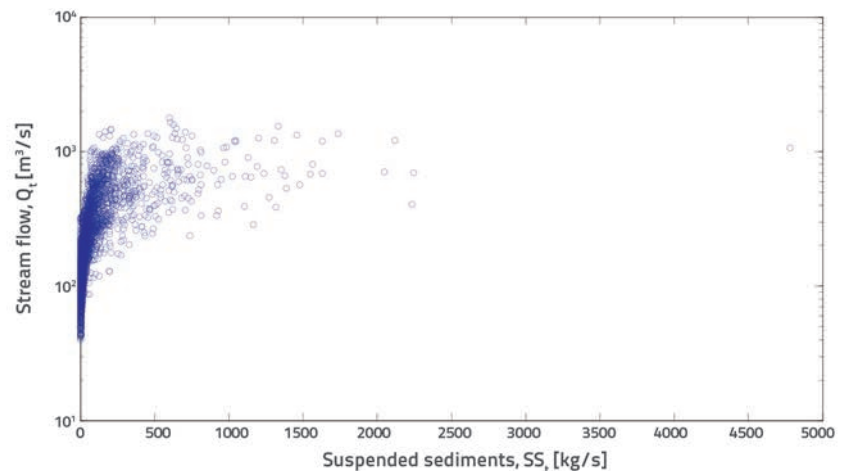Figure 3. The autocorrelation for Q and SS, cross-correlation of SS with Q

Figure 4. The scatter plots of SS_t, Q_t

5. Application and results

Suspended sediment and stream flow time series were decomposed (by wavelet analysis) into some multi-frequency time series for details (different resolution levels) and approximation. They were linked to SVM technique as inputs for predicting one-day-ahead SS. Number of nodes in input layer is determined with (i+1) × 2 because the WSVM combination technique applies two variables (suspended sediment and stream flow). Each time series is decomposed into i, i = (1, 2, 3) details time series, and one approximation time series. The wavelet technique can divide the SS time series properties into various scales of wavelet transform at the same time. Different mother wavelet was decomposed into Q and SS, in diverse levels, from 1 to 3. It was established that with these data, the decomposition in one level for SS and Q signals that yield two subsignals by sym3, the wavelet had the best performance. The details of subsignals are presented in Figure 5.

In this figure, SSApp and QApp are wavelet coefficients of SS and Q in approximation mode, and *SSDet1, QDet1* are wavelet
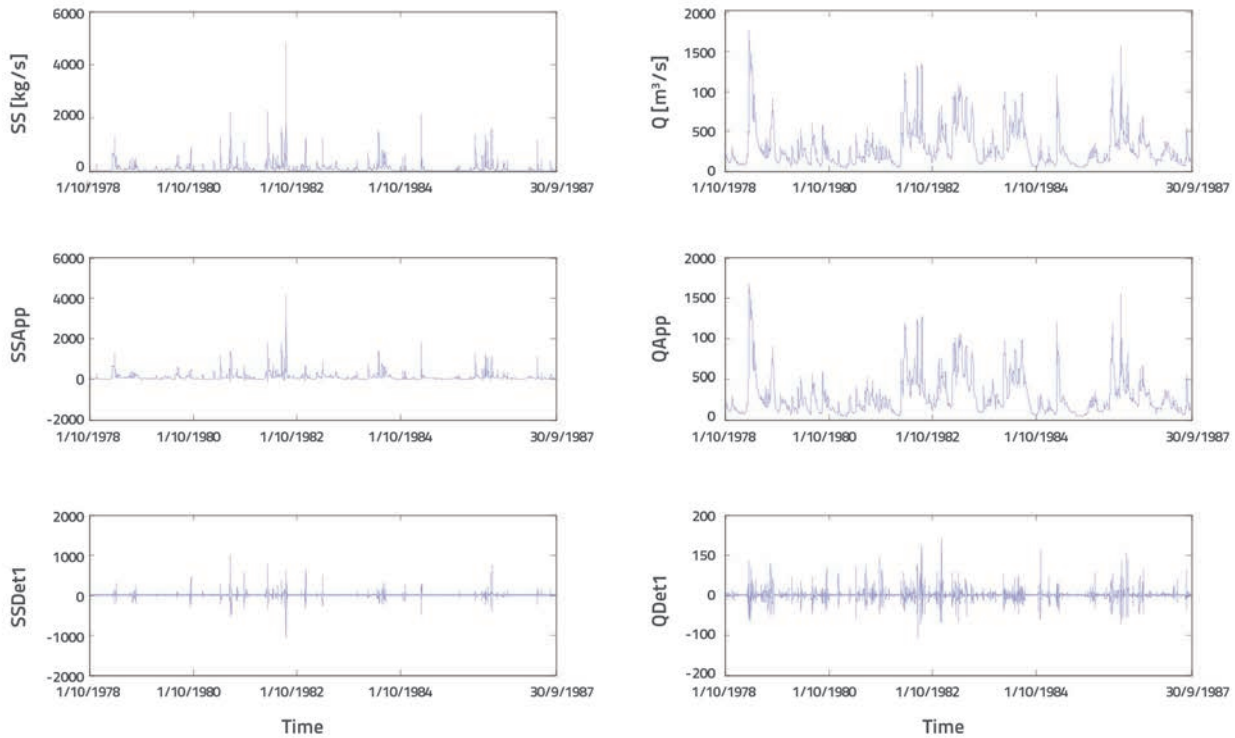
Figure 5. Detail sub-signals and approximation of sym3 wavelet (level 1)

coefficients of SS and Q in detail mode at level 1. SS and Q time series were decomposed at diverse levels by wavelet analysis and imposed as inputs to SVM method for forecasting one-day-ahead SS. By this decomposition, the periodic property of SS was considered. Each time subsignal plays a diverse function in original time series, and the performance of each signal is distinct. The study checks different combinations of Q and SS with maximum three time steps into the past, as inputs to the models. The following input combinations were investigated:

1. Q_{t-1}, Q_{t-2}, Q_{t-3}, SS_{t-1}, SS_{t-2}, SS_{t-3}
2. Q_{t-1}, Q_{t-2}, SS_{t-1}, SS_{t-2}
3. Q_{t-1}, SS_{t-1}
4. SS_{t-1}, SS_{t-2}
5. Q_{t-1}, SS_{t-1}, SS_{t-2}

Optimum parameters of the SVM models were defined by minimizing the objective function (RMSE) error between calculated and observed suspended sediment values in test period for each input combination. In the application of SVM, C and ε were parameters that needed to be specified. In fact, if good care is not taken in parameter selection, the resulting regression model may yield large prediction errors on unobserved future data [46]. The parameter selection tool assumes that the RBF (Gaussian) Kernel was used although extensions to other Kernels and SVR could be easily made. Values for , ε and γ were needed for building a ε - SVM model from training data, and the RBF Kernel function was used [47]. The γ expression, mentioned in Table 1, was significant in the

RBF model and could lead to under fitting and over fitting in prediction [48].

In this research, the γ parameter had a default value equal to 1/num-feature in LIBSVM software, and the C and ε parameters were set to several values and various SVM models were developed. Optimum parameters for each combination in WSVM model by Symlet (Sym3) wavelet are given in Table 3, and the optimal one that presented the minimum RMSE error, mean error, and STD error, and the maximum R^2 in the test period, was chosen. Before applying the WSVM to the data, the training input and output values were normalized using the logarithm function, respectively.

According to this Table 3, the WSVM model provided the best performance criteria for Combination 2 Q_{t-1}, Q_{t-2}, SS_{t-1}, SS_{t-2}. It is clear from the Table that the minimum RMSE=70.8 kg/s, while the highest R^2 value of 0.847 was obtained when the C = 6 and ε = 0,07 were chosen. To characterize periodic properties, the observed Q and SS time series were decomposed into several multi-frequency time series by wavelet transform. Each subsignal participated with different and distinct behaviour in the original time series. These time series were imposed as input data to the SVM technique to forecast the next day SS. The results obtained by the WSVM model demonstrate that the performance increases if decomposition levels for river stream flow and suspended sediment time series are considered equal, while the model performance decreases by increasing the decomposition level to levels greater than 1. This research was made to investigate effects of the employed mother wavelet type and decomposition level on the projected

Table 3. WSVM performance evaluation in SS prediction in the testing period

Combination	1	2	3	4	5
C	1	6	3	2	1
ε	0,001	0,07	0,2	0,001	0,001
R^2	0,765	0,847	0,735	0,821	0,83
RMSE [kg/s]	87,86	70,83	93,34	76,66	74,66
Mean error [kg/s]	5,51	2,15	6,09	2,42	2,53
STD error [kg/s]	87,74	70,86	93,21	76,68	74,67

Table 4. SVM performance evaluation in SS prediction in the testing period

Combination	1	2	3	4	5
C	1	2	1	2	1
ε	0,001	0,001	0,001	0,3	0,01
R^2	0,678	0,636	0,588	0,674	0,61
RMSE [kg/s]	102,96	109,38	116,5	103,52	109,34
Mean error [kg/s]	11,67	9,26	14,83	8,9	10,02
STD error [kg/s]	102,37	109,06	115,63	103,21	108,95

WSVM model efficiency. Therefore, the Q and SS time series were decomposed into one, two, and three levels by six different kinds of wavelet transforms such as Daubechies-2 (db2) wavelet, the Haar wavelet, the Biorthogonal (Bior1.1), Coif1, Symlets (Sym2, Sym3), wavelets. The sym3 wavelet, which had superior results, was chosen. Optimum parameters of each combination in SVM model are given in Table 4. For this station, the SVM model presented the best performance criteria for Combination 1 (Q_{t-1}, Q_{t-2}, Q_{t-3}, SS_{t-1}, SS_{t-2}, SS_{t-3}) and the relative RMSE and R^2 for the model were 102.96 kg/s and 0.678 in the test period.

Compared to the best SVM model, the optimal WSVM model increased R^2 (almost 20 %), while the RMSE decreased by 31.2 %. To support the previous results by visual assessment, the semilog-scaled time series of the measured and predicted SS, and residuals of errors employing SVM and WSVM models for the test period, are depicted in Figures 6 and 7. The log-scaled scatter plots of observed and estimated SS load values and normal distribution of errors are also shown.

Figure 6 shows a good agreement between the observed and predicted values. High and low SS values were better forecast by the WSVM model. The predicted values were close to the measured values and results were closer to the 45° straight line in the scatter plots than the other model. This is due to
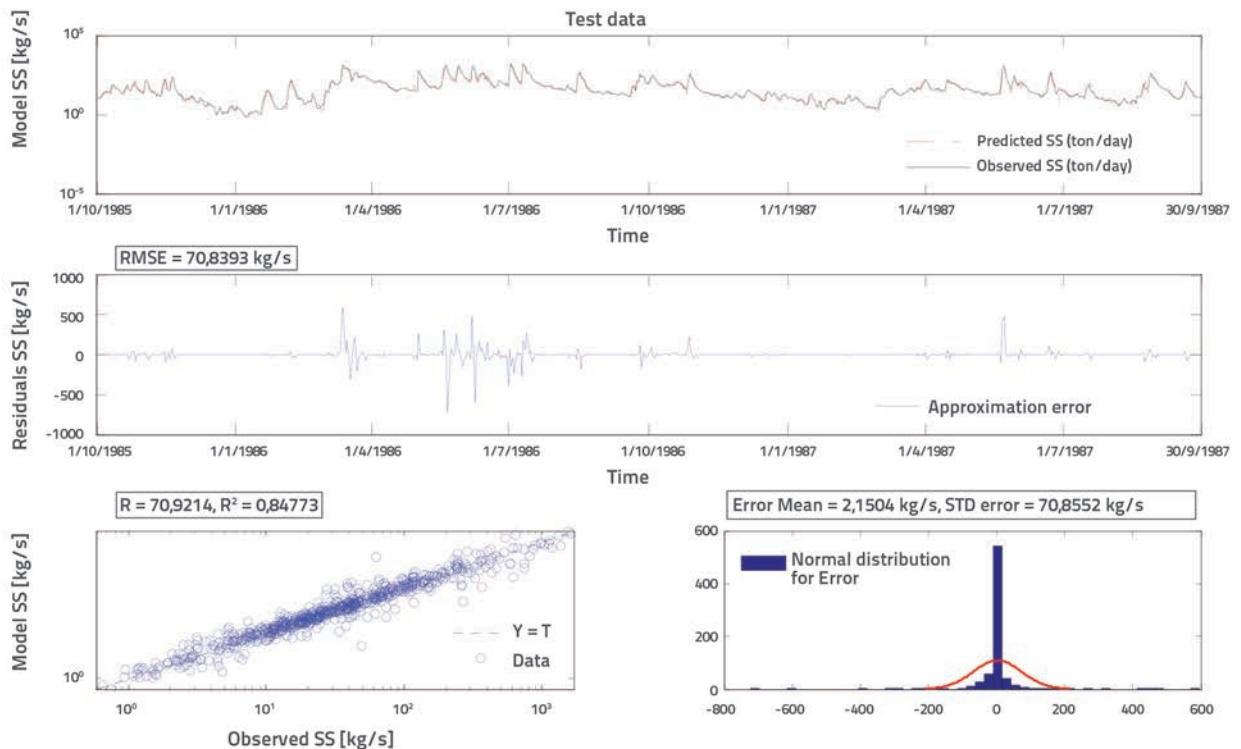

Figure 6. The semilog-scaled time series and log-scaled scatter plots of observed and estimated SS, residuals and normal distribution of errors by best WSVM model in test period
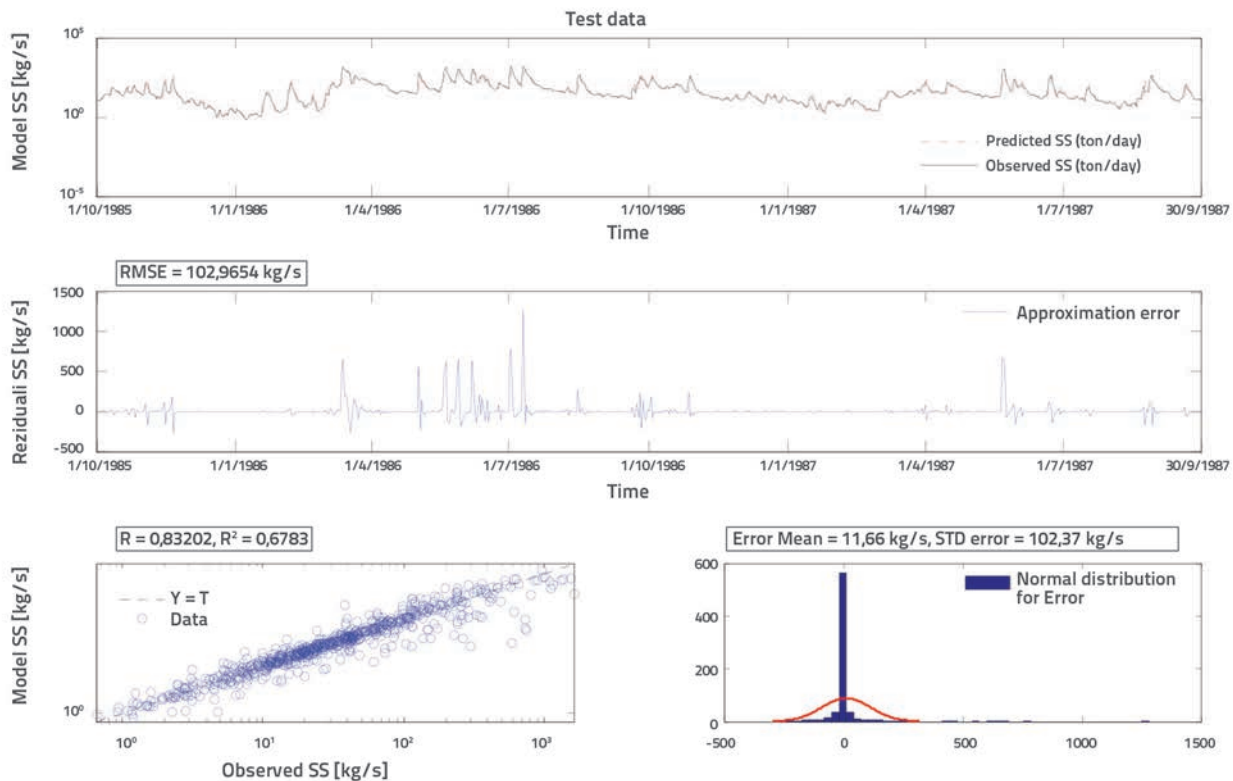
Figure 7. The semilog-scaled time series and log-scaled scatter plots of observed and estimated SS, residuals and normal distribution of errors by best SVM model in test period

the fact that specific features of the subtime series could be noticed more clearly than the original signal. These figures illustrate the accuracy of each model in the calculation of error in standard deviation. In WSVM model the error in standard deviation was about 70.85 kg/s and the bell-shaped curve was steep. In the other technique, the standard deviation was flatter and the examples are spread apart. It amounted to about 102.3 kg/s.

Improvements in prediction can clearly be observed if these methods are compared with a more primitive one such as the sediment rating curve (SRC). Generally, the SRC has the form of SS = aQ^b, where a and b are constants. Thus, the SRC has an important bearing on the correct assessment of SS. The power equation covers both the effect of increased stream power at higher Q, and the extent to which new sources of SS become available in weather conditions that cause high Q. Despite its general use, several problems are recognized with regard to the accuracy of the fitted curve and the physical meaning of its regression coefficients [49]. The SRC was fitted to the training data in this station, and the following equation was obtained:

$$SS = 3 \times 10^{-11}Q^{1,9}$$

The predictive ability of the SRC model was also tested with the same data sets that were employed to test the SVM and WSVM models, thus making the result comparable. To

support previous results by visual assessment, the semilog-scaled time series of the measured and predicted SS, and residuals of errors employing SRC model for the test period, are depicted in Figure 8. RMSE and R^2 for the model were 119.89 kg/s and 0.28 in test period, respectively. The SVM model was demonstrated to be in better agreement with the original SS data than the SRC model. However, the errors show that some contributions of the physics were not considered. Therefore, the WSVM model was more efficient and had better predictions than the SRC and SVM models.

The SS calculation is essential in any reservoir problems. Furthermore, correct assessment of SS is needed for the design and operation of canals, dams and diversions. The cumulative SS was estimated by these models in test period. The WSVM and SVM methods underestimated cumulative SS by about 2.48 % and 13.4 %, respectively. The prediction of high SS values greater than 810.18 kg/s was performed in this paper. Twelve high SS cases were selected for the analysis in the test period and time of each occurrence is shown in Figures 8 and 9. The WSVM model results were nearer to the 1:1 line model. RMSE and R^2 values for this model were 251.42 kg/s and -0.09. The RMSE decreased by 57 % in comparison with the best SVM model. The SVM technique underestimated the SS in most cases. The semilog-scaled and Log-scaled scatter plots of observed and predicted SS for values more than 810.18 kg/s and residuals and normal distribution of errors are shown in Figures 9 and 10.
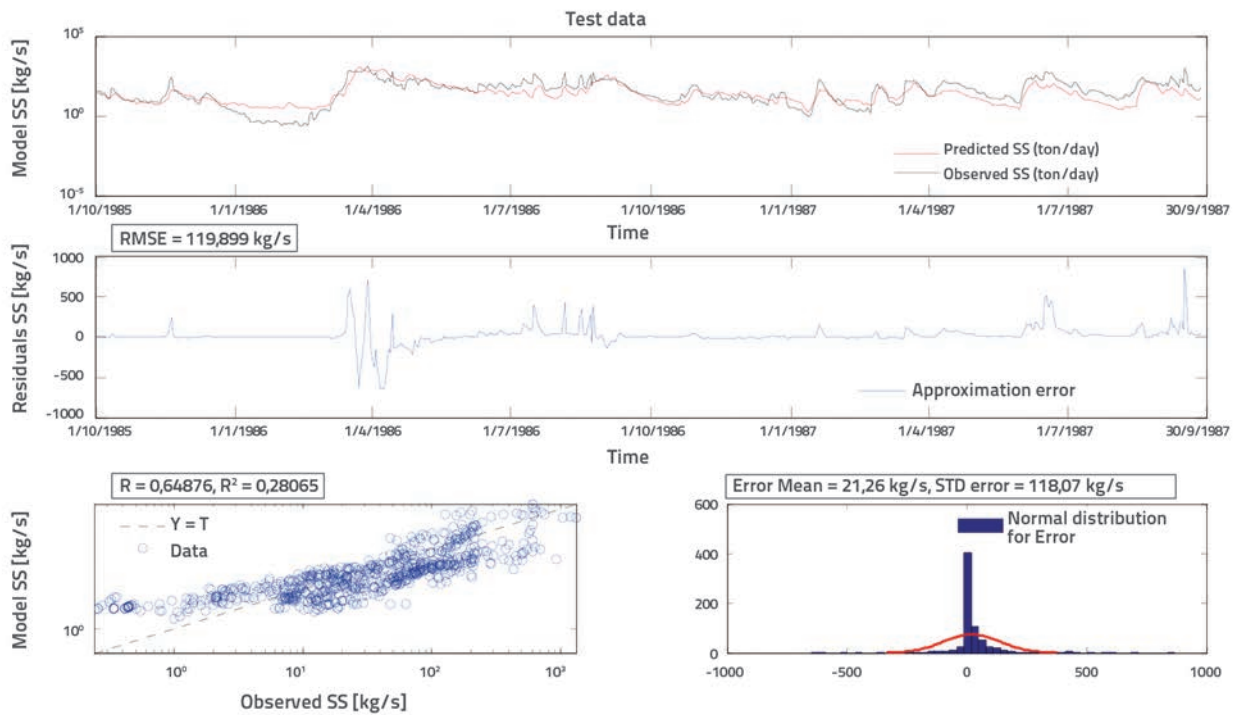
Figure 8. The semilog-scaled time series and log-scaled scatter plots of observed and estimated SS, residuals and normal distribution of errors by best SRC model in test period
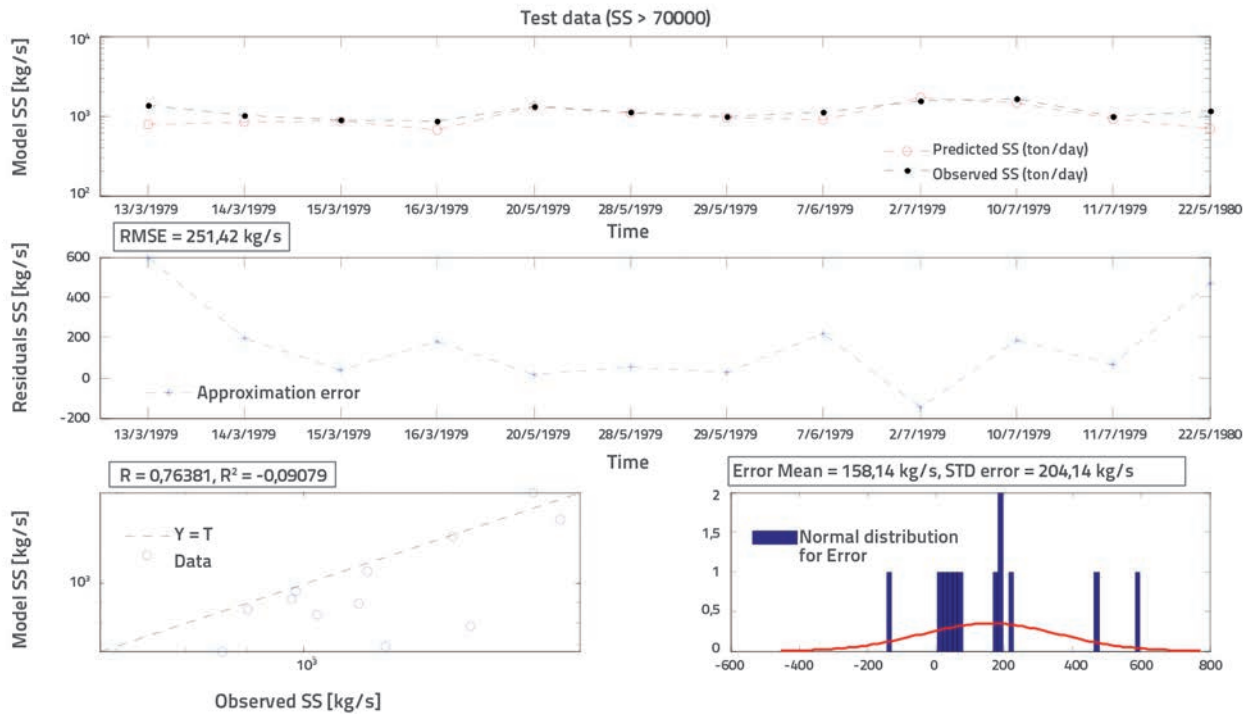

Figure 9. The semilog-scaled time series and Log-scaled scatter plots of observed and estimated SS for values of more than 810.18 kg/s, residuals and normal distribution of errors by best WSVM model in test period
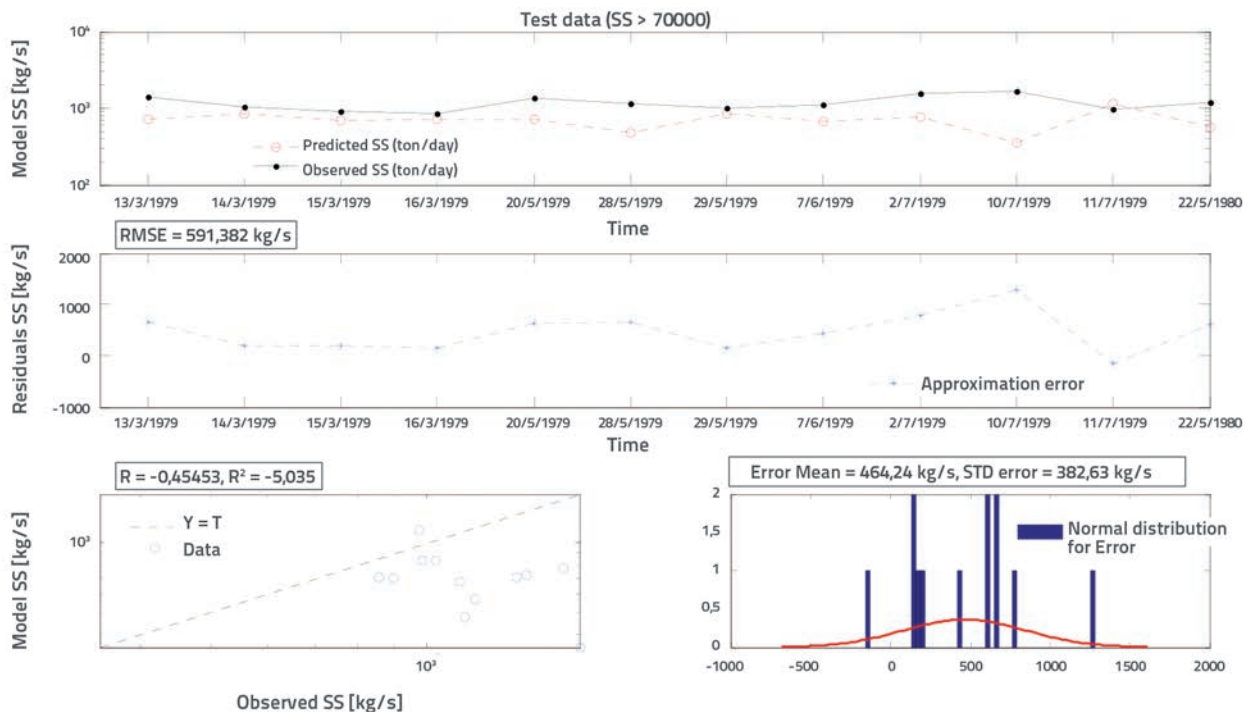
Figure 10. The semilog-scaled time series and Log-scaled scatter plots of observed and estimated SS for values of more than 810.18 kg/s, residuals and normal distribution of errors by best SVM model in test period

6. Conclusion

In this research, WSVM and SVM models were applied to predict the daily suspended sediment load. The stream flow and suspended time series were decomposed into several multi-frequency time series by wavelet transform. Each subsignal participated with different and distinct behavior in the original time series. Observed Q and SS data were considered as input variables, and the best input combination for the models was recognized by using the autocorrelation and cross-correlation. Performances of these models were compared by R^2, RMSE, mean error and standard deviation error. In this investigation, RBF Kernel functions were considered as more logical and efficient consequences compared to other Kernel functions. Conclusions revealed that the WSVM model with Combination 2 (Q_{t-1}, Q_{t-2}, SS_{t-1}, SS_{t-2}) had superior performance than the other model with the minimum RMSE=70.83 kg/s, and the highest R^2 value of 0.847 was obtained when the CC=6 and $\varepsilon = 0{,}07$ was chosen. In general, both SVM and WSVM models exhibited better predictions than SRC model, while the WSVM forecast was more trustful. By comparing these intelligent methods with the sediment rating curve, improvements in prediction can clearly be seen. The WSVM and SVM methods underestimated cumulative SS by about 2.48 % and 13.4 %, respectively. The prediction of high SS values was performed in this paper, and prediction of high suspended sediment load values (greater than 810.18 kg/s) was compared with observation values. Results have revealed that the WSVM model could reasonably be used to estimate cumulative and high suspended sediment load.

REFERENCES

[1] Yang, C.T.: Sediment transport.: theory and practice: McGraw-hill New York, 1996.

[2] Kisi, O.: Suspended sediment estimation using neuro-fuzzy and neural network approaches/Estimation des matières en suspension par des approches neurofloues et à base de réseau de neurones, Hydrological Sciences Journal, 50 (2005) 4, pp. 683-696, DOI: 10.1623/hysj.

[3] Maheswaran, R., Khosa, R.: Wavelet–Volterra coupled model for monthly stream flow forecasting, Journal of Hydrology, 450, pp. 320-335, 2012.

[4] Lin, G.-F., Chou, Y.-C., Wu, M.-C.: Typhoon flood forecasting using integrated two-stage Support Vector Machine approach, Journal of Hydrology, pp. 334-342, 2013.

[5] Nayak, P., Venkatesh, B., Krishna, B., Jain, S.K.: Rainfall runoff modelling using conceptual, data driven and wavelet based computing approach, Journal of Hydrology, pp. 57-67, 2013.

[6] He, J., Valeo, C., Chu, A., Neumann, N.F.: Prediction of event-based stormwater runoff quantity and quality by ANNs developed using PMI-based input selection, Journal of Hydrology, 400, pp.10-23, 2011.

[7] Govindaraju, R.S.: Artificial neural networks in hydrology. II: hydrologic applications, Journal of Hydrologic Engineering, 5, pp. 124-137, 2000.

[8] Kim, T.-W., Valdés, J.B.: Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks, Journal of Hydrologic Engineering, 8, pp. 319-328, 2003.

[9] Mustafa, M., Isa, M., Rezaur, R.: A Comparison of Artificial Neural Networks for Prediction of Suspended Sediment Discharge in River - A Case Study in Malaysia, World Academy of Science, Engineering and Technology, 81, pp. 372-376, 2011.

[10] Nagy, H., Watanabe, K., Hirano, M.: Prediction of sediment load concentration in rivers using artificial neural network model, Journal of Hydraulic Engineering, 128, pp. 588-595,2002.

[11] Raghuwanshi, N., Singh, R., Reddy, L.: Runoff and sediment yield modelling using artificial neural networks: Upper Siwane River, India, Journal of Hydrologic Engineering, 11, pp. 71-79, 2006.

[12] Rajaee, T., Mirbagheri, S.A., Zounemat-Kermani, M., Nourani, V.: Daily suspended sediment concentration simulation using ANN and neuro-fuzzy models, Science of the total environment, 407, pp. 4916-4927, 2009.

[13] Zhu, Y.-M., Lu, X., Zhou, Y.: Suspended sediment flux modelling with artificial neural network: An example of the Longchuanjiang River in the Upper Yangtze Catchment, China, Geomorphology, 84, pp.111-125, 2007.

[14] Rajaee, T.: Wavelet and Neuro-fuzzy Conjunction Approach for Suspended Sediment Prediction, CLEAN–Soil, Air, Water, 38, pp. 275-286, 2010.

[15] Jain, S.K.: Development of integrated sediment rating curves using ANNs, Journal of Hydraulic Engineering, 127, pp. 30-37, 2001.

[16] Rajaee, T.: Wavelet and ANN combination model for prediction of daily suspended sediment load in rivers, Science of the total environment, 409, pp. 2917-2928,2011.

[17] Nourani, V., Alami, M.T., Aminfar, M.H.: A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation, Engineering Applications of Artificial Intelligence, 22, pp. 466-472, 2009.

[18] Rajaee, T., Nourani, V., Zounemat-Kermani, M., Kisi, O.: River suspended sediment load prediction: Application of ANN and wavelet conjunction model, Journal of Hydrologic Engineering, 16, pp. 613-62, 2010.

[19] Pour, O.M.R., Shui, L.T., Dehghani, A.A.: Genetic Algorithm Model for the Relation Between Flow Discharge and Suspended Sediment Load (Gorgan River in Iran), Electronic Journal of Geotechnical Engineering, 16, pp. 539-553, 2011.

[20] Massei, N., Dupont, J.-P., Mahler, B., Laignel, B., Fournier, M., Valdes, D., Ogier, S.: Investigating transport properties and turbidity dynamics of a karst aquifer using correlation, spectral, and wavelet analyses, Journal of Hydrology, 329, pp. 244-257, 2006.

[21] Rajaee, T., Mirbagheri, S., Nourani, V., Alikhani, A.: Prediction of daily suspended sediment load using wavelet and neuro-fuzzy combined model, International Journal of Environmental Science Technology, 7, pp. 93-110, 2010.

[22] Wang, L.: Support Vector Machines: theory and applications. 177. Springer, 2005.

[23] Abe, S.: Support vector machines for pattern classification: Springer, 2010.

[24] Rogers, L.L., Dowla, F.U.: Optimization of groundwater remediation using artificial neural networks with parallel solute transport modelling, Water Resources Research, 30, pp. 457-481, 1994.

[25] Kisi, O., Shiri, J.: Precipitation forecasting using wavelet-genetic programming and wavelet-neuro-fuzzy conjunction models, Water resources management, 25, pp. 3135-3152, 2011.

[26] Kisi, O., Cimen, M.: A wavelet-support vector machine conjunction model for monthly streamflow forecasting, Journal of Hydrology, 399, pp. 132-140, 2011.

[27] Misra, D., Oommen, T., Agarwal, A., Mishra, S.K., Thompson, A.M.: Application and analysis of support vector machine based simulation for runoff and sediment yield, Biosystems engineering, 103, pp. 527-535, 2009.

[28] Noori, R., Karbassi, A., Moghaddamnia, A., Han, D., Zokaei-Ashtiani, M., Farokhnia, A., Gousheh, M.G.: Assessment of input variables determination on the SVM model performance using PCA, Gamma test, and forward selection techniques for monthly stream flow prediction, Journal of Hydrology, 401, pp. 177-189, 2011.

[29] Kisi, O.: Modelling discharge-suspended sediment relationship using least square support vector machine, Journal of Hydrology, 456, pp. 110-120, 2012.

[30] Kakaei Lafdani, E., Moghaddam Nia, A., Ahmadi, A.: Daily Suspended Sediment Load Prediction Using Artificial Neural Networks and Support Vector Machines Machine, Journal of Hydrology, pp. 50-62, 2012.

[31] Kisi, O., Dailr, A.H., Cimen, M., Shiri, J.: Suspended sediment modelling using genetic programming and soft computing techniques, Journal of Hydrology, 450, pp. 48-58, 2012.

[32] Theodoridis, S., Pikrakis, A., Koutroumbas, K., Cavouras, D.: Introduction to Pattern Recognition: A Matlab Approach: A Matlab Approach: Access Online via Elsevier, 2010.

[33] Cherkassky, V., Ma, Y.: Practical selection of SVM parameters and noise estimation for SVM regression, Neural networks, 17, pp. 113-126, 2004.

[34] Cristianini, N., Shawe-Taylor, J.: An introduction to support vector machines and other kernel-based learning methods: Cambridge university press, 2000.

[35] Kecman, V.: Learning and soft computing: support vector machines, neural networks, and fuzzy logic models: MIT press, 2001.

[36] Vapnik, V., Golowich, S.E., Smola, A.: Support vector method for function approximation, regression estimation, and signal processing, Advances in neural information processing systems, pp. 281-287,1997.

[37] Smola, A.J., Schölkopf, B.: A tutorial on support vector regression, Statistics and computing, 14, pp. 199-222, 2004.

[38] Liu, Z., Wang, X., Cui, L., Lian, X., Xu, J.: Research on water bloom prediction based on least squares support vector machine. in Computer Science and Information Engineering, WRI World Congress on. 2009. IEEE, 2009.

[39] Xu, Y., Chen, X., Li, Q.: INS/WSN-Integrated Navigation Utilizing LS-SVM and H∞ filtering, Mathematical Problems in Engineering, 2012, pp. 1-19,2012. doi:10.1155/2012/707326.

[40] Kucuk, M., Ağirali-super, N.: Wavelet regression technique for streamflow prediction, Journal of Applied Statistics, 33, pp. 943-960, 2006.

[41] Youssef, O.A.: A wavelet-based technique for discrimination between faults and magnetizing inrush currents in transformers, Power Delivery, IEEE Transactions on, 18, pp. 170-176, 2003.

[42] Polikar, R.: The wavelet tutorial part iii, IOWA State University, USA, 1996

[43] Polikar, R.: The Wavelet Tutorial Part IV, Multiresolution Analysis: The Discrete Wavelet Transform, ww. cs. ucf. edu/courses/cap5015/WTpart4. pdf, 2008.

[44] Legates, D.R., McCabe, G.J.: Evaluating the use of "goodness of fit" measures in hydrologic and hydroclimatic model validation, Water Resources Research, 35, pp. 233-241, 1999.

[45] Nash, J., Sutcliffe, J.: River flow forecasting through conceptual models part I, A discussion of principles, Journal of Hydrology, 10, pp. 282-290, 1970.

[46] Xie, Z., Lou, I., Ung, W.K., Mok, K.M.: Freshwater algal bloom prediction by support vector machine in macau storage reservoirs, Mathematical Problems in Engineering, 2012.

[47] Chang, C.-C., Lin, C.-J.: LIBSVM: a library for support vector machines, ACM Transactions on Intelligent Systems and Technology (TIST), 2:3, article no. 27, 2011. doi: 10.1145/1961189.1961199.

[48] Han, D., Chan, L., Zhu, N.: Flood forecasting using support vector machines, Journal of hydroinformatics, 9, pp. 267-276, 2007.

[49] Asselman, N.: Fitting and interpretation of sediment rating curves, Journal of Hydrology, 234, pp. 228-248, 2000.