

Modeling river stage–discharge–sediment rating relation using support vector regression

Sharad K. Jain

ABSTRACT

A variety of data-driven approaches have been developed in the recent past to capture the properties of hydrological data for improved modeling. These include artificial neural networks (ANNs), fuzzy logic and evolutionary algorithms, amongst others. Of late, kernel-based machine learning approaches have become popular due to their inherent advantages over traditional modeling techniques. In this work, support vector machines (SVMs), a kernel-based learning approach, has been investigated for its suitability to model the relationship between the river stage, discharge, and sediment concentration. SVMs are an approximate implementation of the structural risk minimization principle that aims at minimizing a bound on the generalization error of a model. These have been found to be promising in many areas including hydrology. Application of SVMs to regression problems is known as support vector regression (SVR). This paper presents an application of SVR to model river discharge and sediment concentration rating relation. The results obtained using SVR were compared with those from ANNs and it was found that the SVR approach is better when compared with ANNs.

Key words | artificial neural network, river stage–discharge rating, sediment concentration, support vector regression

Sharad K. Jain
Indian Institute of Technology,
Roorkee-247667,
India
E-mail: s_k_jain@yahoo.com

INTRODUCTION

Behaviour of hydrologic variables can be modeled either by using laws of physics (or their simplifications) or by using the theory of statistics. In the past few decades, a wide range of hydrological models based on these concepts has been developed and applied. Among the methods based on statistical theory, conventionally, methods such as the time series analysis, simple and multiple linear regression (MLR), etc., are used to predict hydrological variables. In the 1990s, data-driven techniques became very popular among hydrologists and the published literature has witnessed a large number of successful applications of these techniques to a variety of hydrological problems.

Among the techniques that have found a large number of applications are artificial neural networks (ANNs), fuzzy logic theory, support vector machines (SVMs), and wavelet theory. ANNs are based on simplified representation of the architecture of the human brain (ASCE 2000). These have

been used in a wide range of problems in hydrology, for example, flow predictions, flow/pollution simulation, parameter identification, and modeling complex non-linear input-output time series. Recent studies on ANN application in the area of hydrology include rainfall-runoff modeling (Cigizoglu 2003; Solomatine & Dulal 2003; Wilby *et al.* 2003; Nayak *et al.* 2004; Hu *et al.* 2005; Abrahart & See 2007), river stage forecasting (Imrie *et al.* 2000; Lekkas *et al.* 2001; Campolo *et al.* 2003), reservoir inflow forecasting and operation (Jain *et al.* 1999; Coulibaly *et al.* 2001), describing soil water retention curve (Jain *et al.* 2004), estimating daily suspended sediment data (Jain 2001; Cigizoglu 2004; Partal & Cigizoglu 2009) and streamflow prediction (Zealand *et al.* 1999; Chang & Chen 2001; Cigizoglu 2003; Cigizoglu & Kisi 2005; Kisi 2005, 2007; Prada-Sarmiento & Obregón-Neira 2009).

The natural complexity of hydrologic variables poses new challenges either to develop new techniques for their

modeling or to try the techniques that have been found to be successful in the other fields. In the recent past, a statistical learning approach, known as the SVM has been employed in many fields of engineering. This technique has also been applied to some hydrological problems and this was the motivation to apply it to the present problem to examine if it leads to improved results. Before proceeding further, a brief description of SVM is in order.

Support vector machines

SVMs were developed for classification problems in the late 1990s by Vapnik and others (Vapnik 1998, 2000). The first applications of SVMs were reported in the late 1990s and since then, their applications have rapidly grown in widely varied fields such as pattern recognition, text classification and, of course, water resources.

The SVM implements the structural risk minimization (SRM) principle rather than the empirical risk minimization (ERM) principle implemented by most traditional models including the ANNs. The most important concept of SRM is minimizing an upper bound to the generalization error instead of minimizing the training error. Based on this principle, the SVM achieves an optimum network structure. In addition, SVM is equivalent to solving a linear constrained quadratic programming problem so that the solution of the SVM is always unique and globally optimal. Originally, SVMs have been successfully applied to pattern recognition problems (Borges 1998; Hsu *et al.* 2003). However, along with the introduction of Vapnik's ϵ -insensitive loss function, SVMs have been extended to solve nonlinear regression estimation by Gunn (1998) and time series forecasting (Thissen *et al.* 2003).

In this paper, the application of support vector regression (SVR) to model the river stage–discharge and sediment concentration rating relation has been presented. To judge the performance of SVR, the results have been compared with those obtained using ANN. Before describing the methodology, a brief description of feature space and kernel functions is presented in the following section.

Statistical learning theory

SVMs are supervised machine learning techniques from the family of generalized linear classifiers. The task of learning

from examples can be formulated in the following way. Given a set of functions and a set of examples, each one of them generated from an unknown probability distribution $P(x, y)$, learn a function $f(\alpha)$ which provides the smallest possible value for the average error (called the risk) when tested on independent examples randomly drawn from the same distribution P . Here α is the set of parameters. Since $P(x, y)$ is unknown, the risk $R(\alpha)$ is unknown. Therefore, an induction principle for risk minimization is required. One such principle is known as the ERM. However, if the number of training samples is small (which is quite common in hydrology), ERM principle does not guarantee a small actual risk.

Vapnik (1995, 1998) developed the statistical learning theory in which the problem of learning an input-output relationship from a data set is viewed as that of choosing, from the given set of functions, the one that best approximates the output value y for every input vector x , according to an unknown conditional distribution function. Let the expected value of the loss due to classification or estimation errors be given by a risk functional $R(\alpha)$

$$R(\alpha) = \int L[y, f(x, \alpha)] dP(x, y) \quad (1)$$

where $L[y, f(x, \alpha)]$ is a measure of deviation between the observed output y and the computed output $f(x, \alpha)$. The goal of learning is to find the function $f(x, \alpha_0)$ that minimizes the risk functional $R(\alpha)$ when the only available source of information about the system is the training set (Dibike *et al.* 2001).

To make the best use of limited data, Vapnik (1995, 1998) developed a statistical technique called SRM. SRM aims to minimize the risk functional with respect to the empirical risk and the confidence interval. The confidence interval depends on the chosen class of functions whereas the empirical risk depends on the particular function chosen by the training procedure. The objective is to find that subset from the chosen set of functions which minimizes the risk bound for that subset. SVMs employ the SRM principle whereas the traditional learning methods commonly employ the ERM principle. The key difference between the two is that the SRM minimizes an upper bound on the generalization error while ERM minimizes the error on the training data. This difference between risk minimization

provides SVM with a greater ability to generalize, which is the goal in statistical learning (Dibike *et al.* 2001). Osuna *et al.* (1997), Gunn (1998) and others have shown that the SRM principle is superior to the ERM principle. Moreover, the solutions offered by traditional optimization models may fall in local optima whereas a global optimum solution is assured for SVMs.

Support vector regression

SVMs can be applied to both classification and regression problems. The term SVR describes regression using SVMs. To apply SVMs to regression problems, an alternative loss function is used that includes a distance measure. Let the problem be to approximate the data $\{(x_1, y_1), \dots, (x_n, y_n)\}$ from the real domain with a linear function, $f(x, \alpha) = (w \cdot x) + b$. Here, w is the weight vector and b is known as bias. The regression function $f(x, \alpha)$ is found such that it has at most a deviation of ε from the observed targets y_i and is as flat as possible (Dibike *et al.* 2001). The general loss function with ε -insensitive zone is described as

$$|y - f(x, \alpha)|_{\varepsilon} = \varepsilon \text{ if } |y - f(x, \alpha)| \leq \varepsilon \text{ and} \\ = |y - f(x, \alpha)| \text{ otherwise} \quad (2)$$

The strength of the SVM comes from the use of kernel functions to map input data to a higher dimensional space. A non-linear solution in the original, lower dimensional input space is equivalent to a linear solution in the higher dimensional feature space (Maity *et al.* 2010). Since many hydrological problems involve non-linear relations among the variables, SVM is an attractive tool for such problems. Least Squares SVM (LS-SVM) is a powerful method that uses non-linear kernels to solve regression problems in the framework of SVMs. Here the cost function is a regularized least squares function with equality constraints. The objective function and the equality constraints are given by:

$$\min \varphi = \frac{1}{2} w^T w + \frac{\gamma}{2} \sum_{i=1}^n e_i^2 \quad (3)$$

$$\text{subject to } y_i = w^T \varphi(x_i) + b + e_i, \quad i = 1, \dots, n \quad (4)$$

where e_i is the error and γ is a regularization parameter. The objective is to find the optimal parameters that minimize the prediction error and at the same time, give a function that is as flat as possible. Note that γ represents the trade-off between minimizing the training errors and minimizing the model's complexity. The first term on the right hand side (RHS) of Equation (3), $w^T w/2$ is called the regularization term whose minimization will make the function as flat as possible, thereby controlling the capacity of the function. The optimization problem given by Equations (3) and (4) can be solved by constructing an augmented objective function using Lagrange multipliers. By differentiating this objective function with respect to the parameters, a system of linear equations is obtained whose solution yields the regression function

$$y = f(x, \alpha) = \sum_{i=1}^n \hat{\alpha}_i K(x, x_i) + \hat{b}_i \quad (5)$$

where ' $\hat{\cdot}$ ' denotes the estimated value of the parameter and $K(x, x_i)$ is the kernel function. For example, the radial basis function (RBF) kernel is expressed by:

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{\sigma^2}\right) \quad (6)$$

where σ is the parameter of the RBF kernel.

APPLICATIONS OF SVMs TO HYDROLOGICAL PROBLEMS

In the last two decades, SVMs have been successfully used in very divergent fields of engineering and business (e.g. Smola 1996; Vapnik *et al.* 1997). In the field of hydrology and water resources, SVMs have also seen many applications. Some of these include: to forecast flood stage (Liong & Sivapragasam 2002); to extend the rating curves developed at three gauging stations in Washington, USA (Sivapragasam & Muttil 2005); to predict future water levels in Lake Erie (Khan & Coulibaly 2006); to forecast long-term discharges (Lin *et al.* 2006); to estimate the removal efficiency of settling basins in canals (Singh *et al.*

2008); to develop pedotransfer functions for water retention of soils in Poland (Lamorski *et al.* 2008); to predict river stage at a downstream station using water level time series at an upstream station (Wu *et al.* 2008); to predict suspended sediment transport in rivers (Çimen 2008), for evapotranspiration modeling (Kisi & Çimen 2009); and to predict water level fluctuations in a lake (Çimen & Kisi 2009).

Although the number of applications of SVMs in hydrology has not been very large so far, the results have been very encouraging, motivating new applications. Dibike *et al.* (2001) applied SVM in remotely sensed image classification and rainfall-runoff modeling. Liong & Sivapragasam (2002) reported that the SVM had a superior performance in forecasting flood stage compared to ANN. Bray & Han (2004) successfully used SVMs to identify a suitable structure of rainfall-runoff models and parameters. Comparison of SVM results with a transfer function model brought out the strengths of SVM. Khadam & Kaluarachchi (2004) presented a framework using the method of order of importance to incorporate soft information to describe the relative accuracy of calibration data. The applicability of the framework was demonstrated using the data of a catchment where the short streamflow record with significant gaps was reconstructed using SVM. She & Basketfield (2005) also reported superior results in forecasting spring and fall season streamflows in the Pacific Northwest region of the United States using SVM.

Empirical results obtained from these models showed promising performance in solving site-specific, real-time water resources management problems. Qin *et al.* (2005) used LS-SVMs, a non-linear kernel-based machine, to demonstrate the generalization capabilities of SVMs and their potential for applications in the field of hydrology. Tripathi *et al.* (2006) employed the SVM approach for downscaling precipitation to meteorological subdivisions in India and suggested that these could be a potential alternative to the ANN models. In the same vein, Anandhi *et al.* (2008) presented a methodology to downscale monthly precipitation to river basin scale in an Indian context using SVM and the results indicated the superiority of SVM in downscaling precipitation to river basin scale.

Behzad *et al.* (2009) employed SVM to predict one-day lead stream flow of the Bakhtiyari River in Iran using the

data of local climate and rainfall. The results were compared with those of ANN and ANN integrated with genetic algorithms (ANN-GA). Considerable improvements in root mean squared error (RMSE) and squared correlation coefficient (R^2) were noted when SVM was used in place of ANN models. Li *et al.* (2010) presented a modified SVM based prediction framework to improve the predictability of the inflow to a reservoir using climate data from the prior period and claimed that the proposed modified SVM-based model outperforms the prediction ability of MLR and simple SVM. Maity *et al.* (2010) presented an application of SVR to predict the monthly streamflow of the Mahanadi River, India and compared the results with those derived from ARIMA (autoregressive integrated moving average) model. They found SVR to outperform the ARIMA model.

In this paper, the performance of SVM was compared with ANN which itself has been found to be a powerful mapping method. To that end, it would be helpful to give a comparison of SVM and ANN. Table 1 gives a comparison of the key features of the two techniques.

Along with its advantages, every technique has certain pitfalls. The pitfalls of ANNs have been studied and documented in detail (e.g. see Zhang 2007). SVM also has some pitfalls. Among these, over-fitting is a common pitfall. As pointed out by Han *et al.* (2007), SVMs also have

Table 1 | A comparison of the key features of SVM and ANN

Artificial neural networks (ANN)	Support vector machines (SVM)
Hidden layers map to lower dimensional spaces	Kernel maps to a very-high dimensional space
Search space has multiple local minima	Search space has a unique minimum
Training of an ANN is expensive and ANN initialization is particularly problematic	Training of SVMs is extremely efficient
Classification is extremely efficient	Classification is extremely efficient
An ANN requires a number of hidden layers and neurons	Kernel type, kernel function parameter, and cost are typically selected by trial and error, or the prior knowledge
ANNs have very good accuracy in typical domains	SVMs have very good accuracy in typical domains and are extremely robust

over-fitting and under-fitting problems and that over-fitting is more damaging than the under-fitting. Selection of kernel functions also needs to be studied in detail and in some situations, a linear kernel can give better results than a non-linear one. In this regard, the following observation of Han *et al.* (2007) assumes significance: ‘that linear and nonlinear kernel functions (i.e. RBF) can yield superior performance against each other under different circumstances in the same catchment. It is not a simple task to simply declare one kernel is better than another one in complicated hydrological simulations’.

APPLICATIONS OF SVR FOR RATING CURVE DEVELOPMENT

Encouraged by the findings of the reported applications of SVMs, this study was taken up to model river stage–discharge–sediment concentration relation using SVR.

Study area and data used

In this study, data for two gauging stations on the Mississippi River were used. The gauging stations selected are in Illinois State of the USA and are operated by US Geological Survey (USGS). These are located near Chester (USGS station no. 07020500, drainage area = 1,835,276 sq. km) and Thebes (USGS station no. 07022000, drainage area =

1,847,190 sq. km). The requisite data were downloaded from the USGS web site (water.usgs.gov). For the Chester station, the data from 1 January 2004 to 31 December 2005 (2 years) were chosen for calibration and data from January 1 to 31 December 2006 were chosen for validation. For the Thebes station, the data for 1 January 1990 to 30 September 1990 were used for calibration and the data for 15 January 1991 to 10 August 1991 were used for validation. Note that for this station, data of a full year or more were used for training and testing. Table 2 gives key statistical properties of the data used in the study. For both the sites, the coefficient of variation and skewness coefficient for sediment concentration was higher than the same for discharge. Further, correlation between discharge and sediment for both the sites was very high.

Variation of discharge with river stage for gauging stations shows that although there is scatter in the data it is not much. However, when the river stage and sediment concentration data are plotted, large scatter in the data is noted at most of the sites, particularly at high stage values. It is well known that sediment transport is a complex process that depends upon a large number of interacting factors and it is a challenge to model such behaviour.

Data processing and parameter estimation

The data pertaining to river stage, discharge, and sediment concentration were scaled to fall well in the range 0 to 1

Table 2 | Statistical properties of the data used in the study

Statistics	Thebes		Chester	
	Observed discharge (m ³ s ⁻¹)	Observed sed. concentration (mg l ⁻¹)	Observed discharge (m ³ s ⁻¹)	Observed sed. concentration (mg l ⁻¹)
Mean	7,526.59	462.12	4,604.00	261.85
Minimum	1,386.0	40.0	1,779.61	58.00
Maximum	12,900.0	1,680.0	13,576.15	1,540.00
Standard deviation	2,713.94	343.32	2,386.63	229.31
Coeff. of variation	0.361	0.743	0.52	0.876
Skewness coefficient	0.181	1.189	1.25	2.653
Lag-1 correlation	0.986	0.935	0.99	0.945
Correlation between discharge–sediment data	0.838		0.975	

by the following formula:

$$Q_s = \frac{Q - Q_{\min}}{Q_{\max} - Q_{\min}} \quad (7)$$

where Q_s is standardized flow, Q_{\min} is 0.9 times the minimum flow, and Q_{\max} is 1.1 times the maximum of the flow values. The river stage and sediment data were also standardized in the similar way. Previous studies suggest that the performance SVR using scaled input data is better when compared with that using unscaled data (Bray & Han 2004). After modeling, the outputs were re-transformed to the original domain and performance statistics were determined.

In this study, the Matlab code for SVR based on LS-SVM and developed by Suykens *et al.* (2002) was used. To use the SVR model selected for this study, values of two parameters are required. These parameters are γ and σ . Currently no guidelines are available to determine the best values of these parameters for the given problem. Therefore, near optimal values of these two were determined by trial-and-error. For the data used in this study, by increasing the values of γ , the calibration results were found to improve whereas the validation results were deteriorating. Further, the opposite behaviour was found when σ was increased. It was also found that the optimal values of γ and σ depend upon the statistical properties of the data. Here, a systematic grid search was performed by varying γ and σ and the combination which gave good results for learning and testing was accepted. During the use of SVR, it was also realized that to obtain improved results, different values of γ and σ are needed to model discharge and sediment concentration because of difference in their statistical properties. For example, for the data of Thebes station, values of γ and σ at 20 and 2, respectively, were found to give good results for discharge, whereas for sediment concentration good results were obtained for 5 and 0.5, respectively. It is felt that the determination of the values of γ and σ requires a separate detailed investigation.

In the case of ANN, the training of three-layer feed forward networks was carried out using the Levenberg–Marquardt algorithm. It was observed that the best results were obtained with five neurons in the hidden layer. The transfer function for the hidden and the output layers were

log-sigmoid and pure-linear, respectively. The ‘trainbr’ function of Matlab which was used here to implement Bayesian regularization for weights and biases of the network. Mean square error was minimized and a program was written in Matlab to train the network and for validation.

Various combinations of input variables (i.e. river stage, discharge, and sediment concentrations) were used and computations were performed for training/calibration and testing/validation data sets. The idea behind selection of these inputs was that H_t and/or H_{t-1} will provide the information about the current stage and its change, previous discharge data will help in attaining better modeling due to high correlation. Further, S_{t-1} or S_{t-2} will provide information about previous sediment concentration and is an important input to determine the same at the current time step. Finally, an integrated framework should enable capture of interaction of discharge and sediment concentration and thus improved results.

The performance criteria reported are the sum of squares of errors (SSE), the correlation coefficient between the observed and computed values (CORR) and the Nash–Sutcliffe efficiency. These values are reported in Table 3 (for Chester site) and Table 4 (for Thebes site). In each cell, the first row refers to the ANN model and the second to SVR. A higher value of CORR indicates a better performance of the model whereas for SSE, a smaller value indicates a better model. As far as the Nash–Sutcliffe efficiency measure is concerned, the closer the value is to unity, the better the model. Among these three performance measures, CORR is least sensitive and in most of the cases in this study, a value exceeding 0.98 could be obtained.

RESULTS AND DISCUSSION

Table 3 gives the results for the training/calibration data for the Chester station for the ANN, SVR models as well as MLR. Six models were constructed by using the various combinations of river stage, discharge, and sediment concentration as input data. Starting with H_t as single input, progressively more inputs were added. The performance of any model should improve as more and more variables are included as inputs and this indeed was found to be the case here. It can be readily seen from the table that for the

Table 3 | Sum of squares of errors and coefficient of correlation for ANN and SVR models – training and testing data of Chester site

Model inputs	Model	Training/calibration data		Test/validation data	
		Discharge	Sediment	Discharge	Sediment
		Correlation/SSE/Nash eff.	Correlation/SSE/Nash eff.	Correlation/SSE/Nash eff.	Correlation/SSE/Nash eff.
H_t	ANN	0.999/0.880*10 ⁷ /0.998	0.807/1.701*10 ⁷ /0.651	0.999/0.082*10 ⁷ /0.999	0.737/0.278*10 ⁷ /0.191
	SVR	0.999/0.895*10 ⁷ /0.998	0.797/1.785*10 ⁷ /0.635	0.999/0.102*10 ⁷ /0.999	0.747/0.249*10 ⁷ /0.949
	MLR	-/12.022*10 ⁷ /-		-/4.825*10 ⁷ /-	
H_t , and H_{t-1}	ANN	0.999/0.673*10 ⁷ /0.999	0.813/1.655*10 ⁷ /0.661	0.999/0.098*10 ⁷ /0.999	0.748/0.263*10 ⁷ /0.235
	SVR	0.999/0.621*10 ⁷ /0.999	0.811/1.674*10 ⁷ /0.658	0.999/0.086*10 ⁷ /0.999	0.747/0.252*10 ⁷ /0.948
	MLR	-/11.576*10 ⁷ /-		-/4.747*10 ⁷ /-	
H_t , Q_{t-1} , and S_{t-1}	ANN	0.999 /0.659*10 ⁷ /0.999	0.972/0.268*10 ⁷ /0.945	0.999/0.100*10 ⁷ /0.999	0.940/0.042*10 ⁷ /0.879
	SVR	0.999 /0.615*10 ⁷ /0.999	0.974/0.252*10 ⁷ /0.949	0.999/0.122*10 ⁷ /0.999	0.952/0.034*10 ⁷ /0.993
	MLR	-/8.154*10 ⁷ /-		-/2.072*10 ⁷ /-	
H_t , H_{t-1} , and S_{t-1}	ANN	0.999/ 0.619*10 ⁷ /0.999	0.972/0.273*10 ⁷ /0.944	0.999/0.097*10 ⁷ /0.999	0.939/0.046*10 ⁷ /0.868
	SVR	0.999/ 0.559*10 ⁷ /0.999	0.973/0.257*10 ⁷ /0.948	0.999/0.106*10 ⁷ /0.999	0.950/0.036*10 ⁷ /0.993
	MLR	-/10.212*10 ⁷ /-		-/3.935*10 ⁷ /-	
H_t , H_{t-1} , Q_{t-1} and S_{t-1}	ANN	0.999/0.601*10 ⁷ /0.999	0.973/0.262*10 ⁷ /0.946	0.999/0.109*10 ⁷ /0.999	0.946/0.038*10 ⁷ /0.889
	SVR	0.999/0.398*10 ⁷ /0.999	0.977/0.226*10 ⁷ /0.954	0.999/0.079*10 ⁷ /0.999	0.953/0.036*10 ⁷ /0.993
	MLR	-/1.084*10 ⁷ /-		-/0.357*10 ⁷ /-	
H_t , H_{t-1} , H_{t-2} , Q_{t-1} and S_{t-1}	ANN	0.999/0.470*10 ⁷ /0.999	0.974/0.255*10 ⁷ /0.948	0.999/0.107*10 ⁷ /0.999	0.938/0.045*10 ⁷ /0.871
	SVR	0.999/0.364*10 ⁷ /0.999	0.981/0.181*10 ⁷ /0.963	0.999/0.076*10 ⁷ /0.999	0.947/0.042*10 ⁷ /0.992
	MLR	-/1.078*10 ⁷ /-		-/0.357*10 ⁷ /-	

training data pertaining to discharge, SSE for SVR is much smaller than for ANN while CORR and Nash efficiency are nearly the same or somewhat better in case of SVR. For the training data of sediment concentration, SVR models show better performance most of the times on all the statistical measures.

For the testing/validation data, SVR models give better performance most of the time for both discharge and sediment concentration. In most of the cases, SVR models gave much smaller SSE and higher Nash efficiency and these imply better modeling. MLR technique was also applied to the same data and for all the combinations of inputs. The results for the discharge data for Chester site are shown in Table 3. It can be seen that SSE for various MLR models was much higher than ANN and SVR for the calibration and validation data.

To provide a visual interpretation and appreciation of the results, Figure 1 contains the graphs of variations of observed discharge and that estimated by SVR and ANN

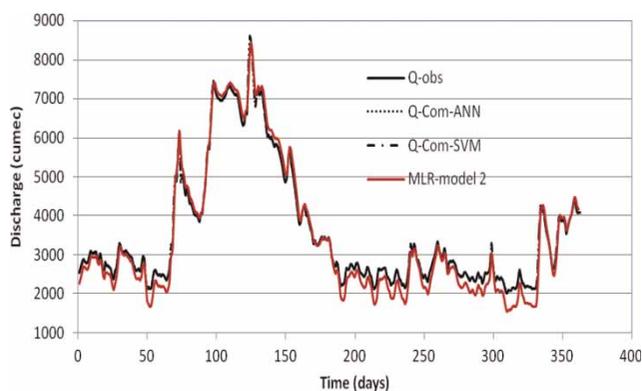
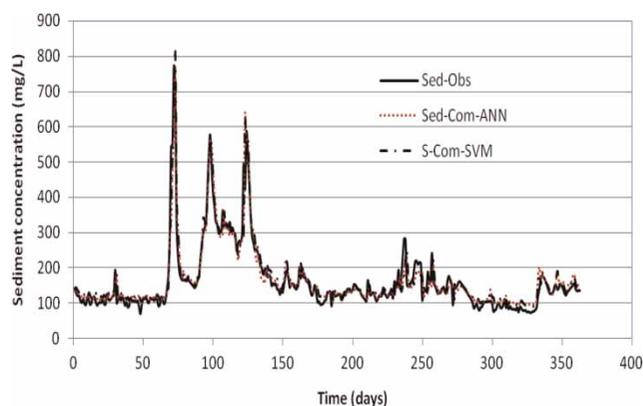
for the Chester station. Both the approaches appear to be performing well but the curve for SVR is closer to the observed values than those for ANN. Figure 1 also contains the graph of the discharge computed by MLR using the same combination of inputs. It can be noted that the curve for MLR is farther away from the other curves and the match of low discharges is poor. Obviously, the linear model is unable to mimic high non-linearity in the data. In view of this, application of MLR to the other data was not pursued further.

Figure 2 shows graphical comparison of observed and modeled sediment concentration for the Chester site for the validation period data and this graph also confirms the statistical results. The curve of computed sediment concentration for SVR is closer to the observed curve compared to the curve for ANN.

In Table 4, the performances of ANN and SVR models for the data of the Thebes site are given. For the training data of discharge, CORR is higher for SVR; SSE is also

Table 4 | Sum of squares of errors and coefficient of correlation for ANN and SVR models – training and testing data of Thebes site

Model inputs	Model	Training/calibration data		Test/validation data	
		Discharge	Sediment	Discharge	Sediment
		Correlation/SSE/Nash eff.	Correlation/SSE/Nash eff.	Correlation/SSE/Nash eff.	Correlation/SSE/Nash eff.
H_t	ANN	0.9826/14.3184*10 ⁷ /0.7721	0.9655/1.5588*10 ⁷ /0.5961	0.9652/10.2362*10 ⁷ /0.7828	0.9298/1.1095*10 ⁷ /0.5433
	SVR	0.9827/14.2811*10 ⁷ /0.7747	0.9657/1.5454*10 ⁷ /0.6001	0.9705/9.2189*10 ⁷ /0.7969	0.9395/1.0651*10 ⁷ /0.7244
H_t , and H_{t-1}	ANN	0.9835/13.5952*10 ⁷ /0.8016	0.9673/1.3799*10 ⁷ /0.6425	0.9646/10.6095*10 ⁷ /0.7735	0.9273/0.9819*10 ⁷ /0.5959
	SVR	0.9850/12.4129*10 ⁷ /0.8520	0.9702/1.0616*10 ⁷ /0.7253	0.9693/9.4132*10 ⁷ /0.7774	0.9383/0.9955*10 ⁷ /0.7424
H_t , Q_{t-1} , and S_{t-1}	ANN	0.9954/3.8325*10 ⁷ /0.9484	0.9908/0.3881*10 ⁷ /0.8994	0.9867/5.0008*10 ⁷ /0.8917	0.9657/0.6111*10 ⁷ /0.7484
	SVR	0.9958/3.4569*10 ⁷ /0.9731	0.9917/0.2156*10 ⁷ /0.9442	0.9843/5.1965*10 ⁷ /0.9018	0.9659/0.4955*10 ⁷ /0.8718
H_t , H_{t-1} , and S_{t-1}	ANN	0.9878/10.0629*10 ⁷ /0.9351	0.9758/0.4848*10 ⁷ /0.8744	0.9798/5.9521*10 ⁷ /0.9158	0.9592/0.5431*10 ⁷ /0.7764
	SVR	0.9943/4.7192*10 ⁷ /0.9851	0.9887/0.1187*10 ⁷ /0.9693	0.9780/7.1209*10 ⁷ /0.9214	0.9533/0.5013*10 ⁷ /0.8703
H_t , H_{t-1} , Q_{t-1} and S_{t-1}	ANN	0.9978/1.8063*10 ⁷ /0.9460	0.9957/0.4056*10 ⁷ /0.8949	0.9954/1.4039*10 ⁷ /0.9113	0.9904/0.4164*10 ⁷ /0.8286
	SVR	0.9982/1.5207*10 ⁷ /0.9832	0.9963/0.1332*10 ⁷ /0.9655	0.9960/1.3198*10 ⁷ /0.9016	0.9913/0.5150*10 ⁷ /0.8667
H_t , H_{t-1} , H_{t-2} , Q_{t-1} and S_{t-1}	ANN	0.9978/1.7899*10 ⁷ /0.9454	0.9957/0.4102*10 ⁷ /0.8937	0.9948/1.5561*10 ⁷ /0.8895	0.9893/0.5410*10 ⁷ /0.7773
	SVR	0.9984/1.3748*10 ⁷ /0.9840	0.9967/0.1356*10 ⁷ /0.9649	0.9948/1.6246*10 ⁷ /0.9284	0.9893/0.3857*10 ⁷ /0.9002

**Figure 1** | Observed and computed discharge by ANN and SVR for validation period for Chester site for inputs H_t , Q_{t-1} and S_{t-1} .**Figure 2** | Observed and computed sediment concentration and SVR with inputs H_t , Q_{t-1} and S_{t-1} for validation period for the Chester site.

smaller for SVR which implies a better match and Nash efficiency is higher which also means a better model. In case of validation data, CORR is slightly better, SSE is significantly better, and Nash efficiency is also better for SVR. For the training data of sediment concentration, SVR was found to perform in a much better way on the counts. For the validation data, the results are mixed. In CORR was better for ANN in three cases but SSE was much better for SVR in two cases and nearly the same or inferior in other cases. Overall, the performance of SVR was found to be better than ANN.

Figure 3 contains the graphs of variations of observed and computed discharge values for the Thebes station. The

graph confirms the statistical evaluation and shows that the match of observed values with the results of SVR is indeed better. Figure 4 shows the graphs of observed sediment concentration and that computed by ANN and SVR. In the case of ANN, there were several instances where the computed curve had higher deviations from the observed curve. In several cases, the deviation between the observed and computed values for ANN is larger as compared to SVR. Both these figures confirm that the deviation of the graph computed by ANN is larger than that by SVR.

Results in the first two rows of Tables 3 and 4 pertain to models with limited inputs (current and previous stage

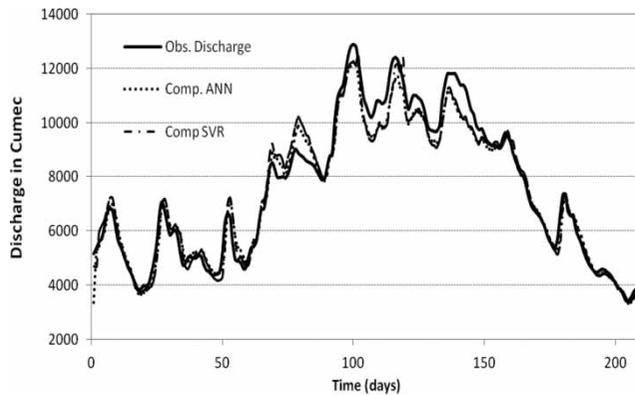


Figure 3 | Observed and computed discharge by ANN and SVR for validation period for Thebes site for inputs H_t , Q_{t-1} and S_{t-1} .

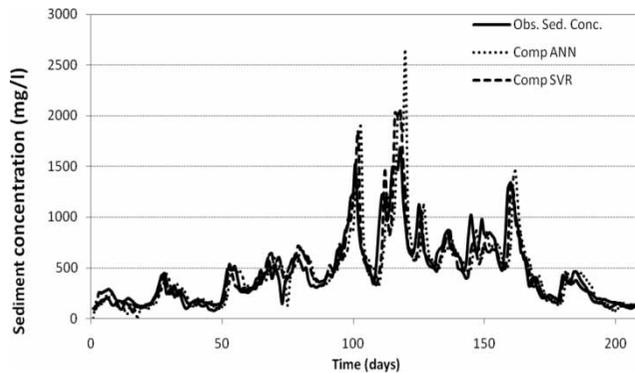


Figure 4 | Observed and computed sediment concentration by ANN and SVR for validation period for Thebes site for inputs H_t , Q_{t-1} and S_{t-1} .

values). For these inputs, the performances of the models for the validation data are not that good. In both the cases, the model performance significantly improved as more inputs were added but the inclusion of H_{t-2} in the sixth (last) input combination only marginally improved the results. It may be mentioned that this study had used stage, flow, and sediment concentration data observed at daily intervals and H_{t-2} is not likely to contain significant additional information about these variables. Thus, in a typical application, the inputs to be included in the model should be chosen based upon hydrological factors.

CONCLUSIONS

In this study, modeling of stage–discharge–sediment concentration rating relation using SVR and ANN models was

studied and the results were compared. Examination of the results using statistical indices and visual display shows that the performance of SVR is much better than the ANN. SVR works better because it utilizes the non-linear feature space. Keeping in mind the accuracy of discharge and sediment concentration measurements in typical field conditions, the accuracy of this model should be adequate in most practical applications. Hence, based on the principle of parsimony, a model with smaller number of inputs may be preferred for practical applications. Further, a SVR model with limited inputs can give reasonably good accuracy for discharge and sediment concentration but the model's ability to correctly estimate these variables critically depends upon the statistical properties of the data.

In addition to better performance, SVR methodology has many advantages compared to ANNs. Generalized software can be developed for SVR. Use of SVR technique is comparatively easy whereas the training of an ANN requires a certain level of skills and experience.

REFERENCES

- Abrahart, R. J. & See, L. M. 2007 [Neural network emulation of a rainfall-runoff model](#). *Hydrol. Earth Syst. Sci. Discuss.* **4**, 287–326.
- Anandhi, A., Srinivas, V. V., Nanjundiah, R. S. & Nagesh Kumar, D. 2008 [Downscaling precipitation to river basin in India for IPCC SRES scenarios using support vector machine](#). *Intl. J. Climatol.* **28**, 401–420.
- ASCE 2000 [Task Committee on application of artificial neural networks in hydrology](#). *artificial neural networks in hydrology, I: preliminary concepts*. *J. Hydrol. Eng. ASCE* **5**, 115–123.
- Behzad, M., Asghari, K., Eazi, M. & Palhang, M. 2009 [Generalization performance of support vector machines and neural networks in runoff modeling](#). *Expert Syst. Appl.* **36**, 7624–7629.
- Bray, M. & Han, D. 2004 [Identification of support vector machines for runoff modeling](#). *J. Hydroinformatics* **6**, 265–280.
- Burges, C. J. C. 1998 [A tutorial on support vector machines for pattern recognition](#). *Data Mining and Knowledge Discovery* **2**, 121–167.
- Campolo, M., Soldati, A. & Andreussi, P. 2003 [Artificial neural network approach to flood forecasting in the River Arno](#). *Hydrol. Sci. J.* **48**, 381–398.
- Chang, F. J. & Chen, Y. C. 2001 [A counterpropagation fuzzy-neural network modeling approach to real time streamflow prediction](#). *J. Hydrol.* **245**, 153–164.

- Cigizoglu, H. K. 2003 Estimation, forecasting and extrapolation of river flows by artificial neural networks. *Hydrol. Sci. J.* **48**, 349–361.
- Cigizoglu, H. K. 2004 Estimation and forecasting of daily suspended sediment data by multi layer perceptrons. *Adv. Water Resour.* **27**, 185–195.
- Cigizoglu, H. K. & Kisi, O. 2005 Flow prediction by three back propagation techniques using K-fold partitioning of neural network training data. *Nordic Hydrol.* **36**, 49–64.
- Çimen, M. 2008 Estimation of daily suspended sediments using support vector machines. *Hydrol. Sci. J.* **53**, 656–666.
- Çimen, M. & Kisi, O. 2009 Comparison of two different data-driven techniques in modeling lake level fluctuations in Turkey. *J. Hydrol.* **378**, 253–262.
- Coulibaly, P., Anctil, F. & Bobee, B. 2001 Multivariate reservoir inflow forecasting using temporal neural networks. *J. Hydrol. Eng. ASCE* **6**, 367–376.
- Dibike, Y. B., Velickov, S., Solomatine, D. P. & Abbott, M. B. 2001 Model induction with support vector machines: introduction and applications. *J. Comput. Civil Eng. ASCE* **15**, 208–216.
- Gunn, S. R. 1998 *Support vector machines for classification and regression*. Technical Report, University of Southampton, UK.
- Han, D., Chan, L. & Zhu, N. 2007 Flood forecasting using support vector machines. *J. Hydroinformatics* **9**, 267–276.
- Hsu, C. W., Chang, C. C. & Lin, C. J. 2003 A practical guide to support vector classification. Technical Report, Department of Computer Science and Information Engineering, National Taiwan University, Taiwan.
- Hu, T. S., Lam, K. C. & Ng, S. T. 2005 A modified neural network for improving river flow prediction. *Hydrol. Sci. J.* **50**, 299–318.
- Imrie, C. E., Durucan, S. & Korre, A. 2000 River flow prediction using artificial neural networks: generalization beyond the calibration range. *J. Hydrol.* **233**, 138–153.
- Jain, S. K. 2001 Development of integrated sediment rating curves using ANNs. *J. Hydraul. Eng. ASCE* **127**, 30–37.
- Jain, S. K., Das, D. & Srivastava, D. K. 1999 Application of ANN for reservoir inflow prediction and operation. *J. Water Resour. Plan. Manage. ASCE* **125**, 263–271.
- Jain, S. K., Singh, V. P. & van Genuchten, M. Th. 2004 Analysis of soil water retention data using Artificial Neural Networks. *J. Hydrol. Eng. ASCE* **9**, 415–420.
- Khadam, I. M. & Kaluarachchi, J. J. 2004 Use of soft information to describe the relative uncertainty of calibration data in hydrologic models. *Water Resour. Res.* **40**, W11505.
- Khan, M. S. & Coulibaly, P. 2006 Application of support vector machine in lake water level prediction. *J. Hydrol. Eng.* **11**, 199–205.
- Kisi, O. 2005 Suspended sediment estimation using neuro-fuzzy and neural network approaches. *Hydrol. Sci. J.* **50**, 683–696.
- Kisi, O. 2007 Streamflow forecasting using different artificial neural network algorithms. *J. Hydrol. Eng. ASCE* **12**, 532–539.
- Kisi, O. & Çimen, M. 2009 Evapotranspiration modelling using support vector machines. *Hydrol. Sci. J.* **54**, 918–928.
- Lamorski, K., Pachepsky, Y., Slawinski, C. & Walczak, R. T. 2008 Using support vector machines to develop pedotransfer functions for water retention of soils in Poland. *Soil Sci. Soc. Am. J.* **72**, 1243–1247.
- Lekkas, D. F., Imrie, C. E. & Lees, M. J. 2001 Improved non-linear transfer function and neural network methods of flow routing for real-time forecasting. *J. Hydroinformatics* **3**, 153–164.
- Li, P.-H., Kwon, H.-H., Sun, L., Lall, U. & Koa, J.-J. 2010 A modified support vector machine based prediction model on streamflow at the Shihmen Reservoir, Taiwan. *Intl. J. Climatol.* **30**, 1256–1268.
- Lin, J.-Y., Cheng, C.-T. & Chau, K.-W. 2006 Using support vector machines for long-term discharge prediction. *Hydrol. Sci. J.* **51**, 599–612.
- Liong, S. & Sivapragasam, C. 2002 Flood stage forecasting with support vector machines. *J. Am. Water Resour. Assoc.* **38**, 173–186.
- Maity, R., Bhagwat, P. P. & Bhatnagar, A. 2010 Potential of support vector regression for prediction of monthly streamflow using endogenous property. *Hydrol. Process.* **24**, 917–923.
- Nayak, P. C., Sudheer, K. P., Rangan, D. M. & Ramasastri, K. S. 2004 Neuro-fuzzy computing technique for modelling hydrological time series. *J. Hydrol.* **291**, 52–66.
- Osuna, E., Freund, R. & Girosi, F. 1997 An improved training algorithm for support vector machines. In: *Proceedings of IEEE Workshop on Neural Networks for Signal Processing 7*, Institute of Electrical and Electronics Engineers, New York, pp. 276–285.
- Prada-Sarmiento, F. & Obregón-Neira, N. 2009 Forecasting of monthly streamflows based on artificial neural networks. *J. Hydrol. Eng. ASCE* **14**, 1390–1395.
- Partal, T. & Cigizoglu, H. K. 2009 Estimation and forecasting of daily suspended sediment data using wavelet–neural networks. *J. Hydrol.* **358**, 317–331.
- Qin, Z., Qiang, Y. U., Jun, L., Zhi-yi, W. & Bing-min, H. 2005 Application of least squares vector machines in modelling water vapor and carbon dioxide fluxes over a cropland. *J. Zhejiang Univ. SCI* **6B** (6), 491–495.
- She, N. & Basketfield, D. 2005 Long range forecast of streamflow using support vector machine. In: *Proceedings of the World Water and Environment Resources Congress ASCE* (R. Walton, ed.). May 15–19, 2005, Anchorage, Alaska, USA.
- Singh, K. K., Pal, M., Ojha, C. S. P. & Singh, V. P. 2008 Estimation of removal efficiency for settling basins using neural networks and support vector machines. *J. Hydrol. Eng. ASCE* **13**, 146–155.
- Sivapragasam, C. & Muttill, N. 2005 Discharge rating curve extension: a new approach. *Water Resour. Manage.* **19**, 505–520.
- Smola, A. J. 1996 Regression estimation with support vector learning machines. MSc Thesis, Technische Universität München, Germany.

- Solomatine, D. P. & Dulal, K. N. 2003 [Model trees as an alternative to neural networks in rainfall–runoff modelling](#). *Hydrol. Sci. J.* **48**, 399–411.
- Suykens, J. A. K., Van Gestel, T., Vandewalle, J. & De Moor, B. 2002 [A support vector machine formulation to PCA analysis and its Kernel version](#). Internal Report 02-68, ESAT-SISTA, K.U. Leuven, Belgium.
- Thissen, U., van Brakel, R., de Weijer, A. P., Melssen, W. J. & Buydens, L. M. C. 2003 [Using support vector machines for time series prediction](#). *Chemometrics and Intelligent Laboratory Systems* **69**, 35–49.
- Tripathi, S., Srinivas, V. V. & Nanjundian, R. S. 2006 [Downscaling of precipitation for climate change scenarios: a support vector machine approach](#). *J. Hydrol.* **330**, 621–640.
- Vapnik, V. 1995 *The Nature of Statistical Learning Theory*. Springer, New York, 187 pp.
- Vapnik, V. N. 1998 *Statistical Learning Theory*. John Wiley & Sons, New York, 740 pp.
- Vapnik, V. 2000 *The Nature of Statistical Learning Theory*. Springer Verlag, New York, USA, 315 pp.
- Vapnik, V., Golwisch, S. & Smola, A. J. 1997 Support vector method for function approximation, regression estimation, and signal processing. In: *Advances in Neural Information Processing Systems 9* (M. Mozer, M. Jordan & T. Petsche, eds). MIT Press, Cambridge, MA, USA, pp. 281–287.
- Wilby, R. L., Abrahart, R. J. & Dawson, C. W. 2003 [Detection of conceptual model rainfall–runoff processes inside an artificial neural network](#). *Hydrol. Sci. J.* **48**, 163–181.
- Wu, C. L., Chau, K. W. & Li, Y. S. 2008 [River stage prediction based on a distributed support vector regression](#). *J. Hydrol.* **358**, 96–111.
- Zealand, C. M., Burn, D. H. & Simonovic, S. P. 1999 [Short term stream flow forecasting using artificial neural networks](#). *J. Hydrol.* **214**, 32–48.
- Zhang, G. P. 2007 [Avoiding pitfalls in neural network research](#). *IEEE Transactions on Systems, Man, and Cybernetics–Part C: Applications and Reviews* **37**, 3–16.

First received 22 October 2010; accepted in revised form 1 June 2011. Available online 27 January 2012