

Development of sediment load estimation models by using artificial neural networking techniques

Muhammad Hassan · M. Ali Shamim · Ali Sikandar ·
Imran Mehmood · Imtiaz Ahmed · Syed Zishan Ashiq · Anwar Khitab

Received: 22 February 2015 / Accepted: 15 September 2015 / Published online: 13 October 2015
© Springer International Publishing Switzerland 2015

Abstract This study aims at the development of an artificial neural network-based model for the estimation of weekly sediment load at a catchment located in northern part of Pakistan. The adopted methodology has been based upon antecedent sediment conditions, discharge, and temperature information. Model input and data length selection was carried out using a novel mathematical tool, Gamma test. Model training was carried out by using three popular algorithms namely Broyden-Fletcher-Goldfarb-Shanno (BFGS), back propagation (BP), and local linear regression (LLR)

using forward selection of input variables. Evaluation of the best model was carried out on the basis of basic statistical parameters namely R -square, root mean squared error (RMSE), and mean biased error (MBE). Results indicated that BFGS-based ANN model outperformed all other models with significantly low values of RMSE and MBE. A strong correlation was also found between the observed and estimated sediment load values for the same model as the value of Nash-Sutcliffe model efficiency coefficient (R -square) was found to be quite high as well.

M. Hassan (✉) · I. Mehmood · I. Ahmed · S. Z. Ashiq ·
A. Khitab
Department of Civil Engineering, Mirpur University of Science &
Technology, Mirpur, A.K., Pakistan
e-mail: hassan25.arif@gmail.com

I. Mehmood
e-mail: imran.mehmood87@gmail.com

I. Ahmed
e-mail: Imtiaz674@gmail.com

S. Z. Ashiq
e-mail: zishanashiq@gmail.com

A. Khitab
e-mail: anwarkhitab@yahoo.com

M. Ali Shamim
Department of Civil Engineering, Bursa Orhangazi Üniversitesi,
Bursa, Turkey

A. Sikandar
Department of Civil Engineering, National University of Sciences
& Technology, Islamabad, Pakistan

Keywords Sedimentation · Dead storage · Physical parameters · Gamma test · Artificial neural networks

Introduction

The management of sediment in river basins and watercourses has been an important issue for water management engineers throughout history—from the ancient Egyptians managing sediment on floodplains to provide their crops with nutrients to today's challenges of siltation in large reservoirs. With the passage of time, the climate change that resulted from anthropogenic activities altered the nature of sediment issues within river systems thus creating complex technical and environmental challenges in relation to sediment management. This is why the accurate estimate for sediment yield has been a problem in water resource analyses and modeling. Sediment yield is a key factor in determining various important aspects such as

the performance and life of reservoirs, canals, drainage channels, harbors, and other downstream structures (Lane et al. 1997).

As far as the history of Indus Basin irrigation system is concerned, by 1947, it consisted of 26 million acres of irrigated agriculture, 34,000 miles of major canals, and 50 million people relying on a system consisting of 13 additional canals that were already in place (Alam 1998). Up to 1990, it was estimated that sediment load in the Indus River ranked fifth in the world (Ali and Boer 2003). Modern irrigation provided the framework around which both Pakistani and Indian Punjab grew to their present economic importance (Alam 1998; Kux 2006). A grid-based regional-scale soil sediment yield model (RSSYM) was set for different catchments of the Indus basin using coarse resolution grid data.

Normally, sediments are divided in two different types: suspended load and bed load. The division is necessary in modeling analyses to wholly understand the behavior of both types of sediments and their effect on different hydrologic parameters. For this purpose, the distinction between sediment loads has been decided by the American Society of Civil Engineers (ASCE) on the recommendations of the Water Environment Research Foundation (WERF) (Environmental Water Resources Institute 2007; Roesner et al. 2006). Furthermore, the sediments are sometimes separated on the basis of that particular source from which they are produced such as glacier-derived sediments and the sediments that resulted from disintegration of the rain-fed catchment. According to Jansson, the glacier-derived sediment load is often more as compared to the rain-fed catchment load. The suspended sediment concentrations observed by various researchers were used to derive source strength parameters, defined either as turbidity generation unit (Nakai 1978), suspension parameter (Pennekamp et al. 1996), or resuspension factor (Hayes and Wu 2001). Delivery of sediment rate from a watershed and its transportation in a channel are the factors that should be quantified in order to decide, which erosion control practices are best suited to address the problem (Weaver and Madej 1981). The sediment yield modeling at regional scale is helpful to solve erosion problems by adopting suitable conservation measures (Painter 1981). For this purpose, many worldwide studies are carried out for example; sediment rate and nutrient transport are assessed by S. A. Akraasi and O. D. Ansa-Asare in the Pra basin of Ghana (Akraasi and Ansa-Asare 2003). Habib-ur-Rehman (2001)

developed a RSSYM based on dynamic approach for soil erosion. The model consists of soil detachment due to raindrop impact. The effect of canopy coverage, canopy height, land use, and land coverage on soil erosion is also modeled. Further RSSYM model includes soil detachment, sediment transport, and deposition due to over land flow. The water flows through equivalent channels; the width of equivalent channels is adjusted according to the flow accumulation value.

Over the last few decades, the hydrologic parameter prediction has been influenced to a greater degree through using artificial intelligence techniques (Cobaner et al. 2009). Some examples of this include streamflow prediction (Hassan et al. 2014), flood forecasting (Remesan et al. 2010), runoff prediction (Remesan et al. 2009), prediction of reservoir levels (Shamim et al. 2015), identification of physical process for rainfall runoff model (Jain et al. 2004) solar radiation estimation (Shamim et al. 2014; Shamim et al. 2010; Remesan et al. 2008), and for ground water contamination (Shamim et al. 2004). Kisi et al. (2008) has also used black box models in water resources management. Many other researchers including Partal and Cigizoglu (2008); Tayfur (2002); Kisi (2004a, 2005); Kisi et al. (2008); Tayfur and Guldal (2006); Nourani (2009); Zhu et al. (2007); Rajaei et al. (2009); and Cigizoglu (2004) have also used artificial intelligence techniques like artificial neural networking and support vector machines for predicting suspended sediment load in different river basins. In addition, sediment loads have also been predicted by Alp and Cigizoglu (2005), Nagy et al. (2002), Lin and Namin (2005), and Agarwal et al. (2006) by using ANN techniques. Alizdeh et al. (2015) have developed the Wavelet ANFIS model to estimate sedimentation in a dam reservoir. Olyaie et al. (2015) compared the efficiency of three different sediment prediction methods, namely artificial neural networks (ANNs), adaptive neuro-fuzzy inference system (ANFIS), coupled wavelet and neural network (WANN), and conventional sediment rating curve (SRC). They concluded that WANN outperformed the other methods for sediment prediction. The researchers including Mustafa et al. 2011 have also compared the efficiency of two different neural networking techniques for prediction of suspended sediment discharge. The obtained results show that the RBF network model has better performance than the MLFF network model in suspended sediment discharge prediction. The load transport has

also been predicted using machine learning approach by Bhattacharya et al. (2007) and Azamathulla et al. (2010). Jie and Yu (2011) have used ANN techniques for estimating the sediment load in Kaoping River basin in Taiwan. Heng and Suetsugi (2013) estimated the sediment load in un-gauged catchments of the Tonle Sap River basin, Cambodia, by using ANN techniques. Similarly, Kisi (2004b) has estimated the suspended sediment load at two stations in the Tongue River in Montana, USA, by using ANN techniques. Also, Nourani et al. (2014) studied the capability of artificial neural network for detecting hysteresis phenomenon involved in hydrological processes. Conclusively, ANN techniques have been widely used and accepted for efficient prediction of different hydrological parameters and also, it is a very inexpensive tool for prediction as it can be performed in even slow computers (Shamseldin 2010). The current study focuses on accurate estimation of sediment loads by using ANN techniques at “Basham,” an upstream station of the most important water reservoir of Pakistan, Tarbela. It has been observed that the capacity of the reservoir has reduced 29 % by sediment depositions in recent years resulting in storage capacity of 6.56 MAF as compared to the design capacity 9.59 MAF. The aroused situation requires proper planning and management of sediments to save the precious source of water for Pakistan as the Tarbela reservoir has been used for multi benefits such as for irrigation, hydropower generation, and recreational purposes. The ANN-based models can be an economical approach for this purpose where it is practically impossible to have detailed information about the topography, precipitation, and other physical parameters due to difficult terrain of Upper Indus Basin. The study is also important in the sense that it is time saving and it is the first of its kind ever used in Pakistan.

Methodology

Study area and data set

The Indus River has its source near Lake Mansrowar in the Himalayan catchment and emerges from the land of glaciers on the northern slopes of the Kailash ranges 5182 m above sea level. It largely brings in snowmelt runoff in addition to monsoon rains. Four main upstream tributaries join the Indus: Shyok River at an

elevation of 2438 m near Skardu, the Gilgit, and Hunza rivers near Bunji and Siran River just north of Tarbela. All these tributaries, except Siran, are located in a semiarid environment and hardly any vegetation exists in the watershed area. Geologic erosion is a common phenomenon in the Indus basin.

The Indus river is in fact one of the largest sediment-producing rivers of the world and is fed mainly by snow and glaciers of the Himalayas. River Indus determines the lives and livelihoods of around one million people in the territories through which it flows. It is estimated that the Indus and its tributaries carry about 0.35 MAF (0.435 BCM) of sediment annually. Of this, nearly 60 % remains in the system and finally it is deposited in the reservoirs, canals, and irrigation fields. This ultimately requires costly silt clearance measurements in the system. Data regarding sediments, discharge, precipitation, etc. was observed on daily, weekly, and monthly basis by the Water and Power Development Authority (WAPDA) at different stations on River Indus. The upper Indus basin along different stream gauges with their respective elevations is shown in Fig. 1.

Datasets

The available datasets comprised of weekly records of discharge, temperature, and suspended sediment load at three different meteorological stations Bunji, Shatial, and Besham as located in Fig. 1 with red circles. The detail of these stations and selected data set is shown in Table 1. Datasets for the study area were obtained from Pakistan Meteorological Department and Surface Water Hydrology Department, WAPDA.

Table 1 illustrated that nine parameters have been considered for estimating suspended sediment load at Besham. In this case, eight parameters were considered as inputs for model training and the one suspended sediment load at Besham has been considered as the desired output.

Training and testing data has been decided by performing Gamma test in conjunction with the *M*-test. From the total available data length (5 years), 180 weeks (approximately 3.5 years) has been selected for training phase and the remaining data, 80 weeks (approximately 1.5 years), was used for testing and validation purposes. Two normalization processes were adopted: Box-Cox transformation and normalization by scaling the data

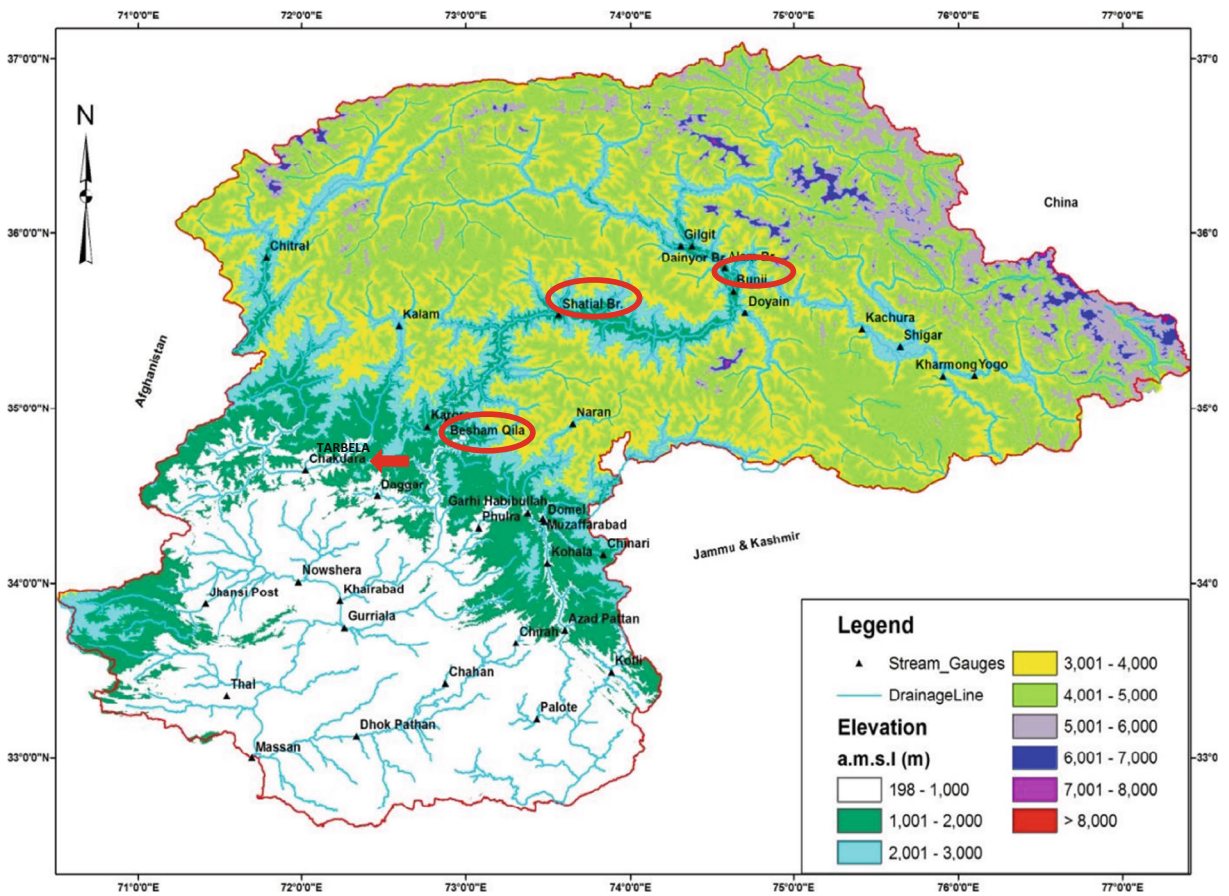


Fig. 1 Upper Indus basin and location of stream gauges

between 0 and 1. Former transformation has been applied to the data using suitable value of power factor (λ) which was selected, on the basis of least value of standard deviation, by applying distributive statistical analysis on the data. The value derived out in this case was 0.1. Jason has proved the significance of Box-Cox transformation upon other types of data transformations (Osborne 2010). The second transformation was performed using following

standard formula:

$$(\text{Normalized})P_i = \frac{P_i - P_{\min}}{P_{\max} - P_{\min}} \tag{1}$$

where

- P_i i th value of variable P
- P_{\min} the minimum value for variable P
- P_{\max} the maximum value for variable P

The purpose of applying two different types of transformations was to check the suitability of technique on this data by comparing the final results.

Table 1 Available datasets and length of observations

Sr No.	Stations	Datasets available	Period of record
1	Shatial	Discharge (Q_{SHA}), temperature (T_{SHA}), sediment load (S_{SHA})	2004–2008
2	Bunji	Discharge (Q_{BUN}), temperature (T_{BUN}), sediment load (S_{BUN})	2004–2008
3	Besham	Discharge (Q_{BESH}), temperature (T_{BESH}), sediment load (S_{BESH})	2004–2008

Determination of best input combination using Gamma test

A novel mathematical tool, Gamma test, was used for the evaluation of best input combination which is proficient enough to develop a smooth model for a specified set of real variables. Basically, this test makes us enable to calculate the variance of noise on our desired output,

called Gamma statistics or best mean square error. This can be performed by decomposing the relationship between two variables x and y into smooth and noisy parts, assuming that y is a function of x . This relationship between assumed variables x and y can be captured by developing an algorithm on the basis of an initial data set $\{(x_i, y_i), 1 \leq i \leq M\}$.

If f is a smooth function between x and y and r is the part of noise which cannot be considered for by any even model, then their relationship can be shown as (Agalbjorn et al. 1997)

$$Y = f(x) + r \tag{2}$$

If the mean of this noise “ r ” is zero, then a constant bias can be engaged into the unknown function f . The interesting thing is that despite the fact that f is unknown, this tool enables us to calculate the value of noise on an output on this condition: As the number of data samples increase, the gamma value becomes equal to a an asymptotic value which represents the variance of a noise on an output (Stefánsson et al. 1997).

Determination of data length through M -test

Once the noise-free combination was determined through Gamma test then it is necessary to determine that length of data which is the best suitable for the development of models which are neither under-fitted nor over-fitted. Basically, M -test is a criterion for determining whether an infinite series of functions converges uniformly. It applies to series whose terms are functions with real or complex values and is analogous to the comparison test for determining the convergence of series of real or complex numbers.

If $\{f_n\}$ is a sequence of real- or complex-valued functions defined on a set A and that there exist positive constants M_n such that

$$|f_n(x)| \leq M_n \tag{3}$$

for all $n \geq 1$ and all, $x \in A$ where the series converges

$$\sum_{n=1}^{\infty} M_n,$$

then it can be concluded that the series

$$\sum_{n=1}^{\infty} f_n(x)$$

converges uniformly on A .

Model development

Local linear regression

This is a reliable technique widely used in many low dimensional forecasting that can predict to a high degree of accuracy. The main requirement of the model is the number of nearest neighbors, p max, chosen by influence statistics which are used to solve a linear matrix Eq. (4) for a given neighborhood of p max points.

$$Xm = Y \tag{4}$$

where X is a d matrix of p max input points in d dimension, $x_i(1 \leq i \leq p$ max) are the nearest neighbor points, Y is a column vector of length p max, of the corresponding outputs, and m is a column vector of parameters that should be determined to provide the optimal mapping from X to Y so that:

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1d} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2d} \\ x_{31} & x_{32} & x_{33} & \dots & x_{3d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{xpmax1} & x_{xpmax2} & x_{xpmax3} & \dots & x_{xpmaxd} \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ m_3 \\ \vdots \\ m_d \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{max} \end{bmatrix} \tag{5}$$

The number of linearly independent rows is the rank r of matrix X that will affect the existence or uniqueness of the solution for m . If matrix X is considered to be a nonsingular square matrix, then the solution to Eq (6) will be given as (Evans 2002)

$$m = X^{-1}y \tag{6}$$

On the other hand, if X is not square or singular, then Eq. (6) is modified and an attempt is made to find a vector m , which minimizes $|Xm - y|^2$.

Artificial neural networks

Artificial neural networks (ANNs) were originally developed by McCulloch and Pitts (1943) which were further modified by Rosenblatt (1962) who proposed the idea of perceptrons that focused on weights and training algorithms for computational tasks. ANNs are inspired by the biological (brain) neuron system, in which each neuron receives, processes, and sends information so as to construct functional relationships between past and future events or values (Shamim et al. 2010). As far as an artificial neuron model is concerned, it consists of a set of connecting links that receives the n input signals X_i , multiplies these with corresponding synoptic weights, W_i , $i = 1, 2, 3, \dots, n$. Afterwards, signals are added up in the summing junction through a linear combiner. The activation function $f(\cdot)$ limits the amplitude of the output and gives its nonlinearity. Mathematically, it can be given by (Agalbjorn et al. 1997)

$$\sigma = \sum_{i=1}^n X_i W_i \quad (7)$$

and

$$Y = f(\sigma + b) \quad (8)$$

where b represents *bias*.

A neuron alone cannot do anything and is of no use until and unless it is combined with other neurons to form a network. Over the years, a countless number of ways have been proposed by which a number of neurons can be connected together to form a network, and algorithms deployed to train them. A few examples include single and multilayer feed forward networks, recurrent networks, radial basis function networks, Lekkas et al. (2001); Maier and Dandy (1998); Haykin (1999); Fine (1999), Fletcher (1987). Three or more layers are most commonly used that are normally comprised of an input layer, one or more intermediate layers (also called hidden layers), and an output layer (Remesan et al. 2008). An important aspect of model development using ANNs is training (learning) phase. Two types of learning methodologies, namely supervised and unsupervised learning, exist. Supervised learning is most

commonly used within which inputs and desired outputs are fed into the network so that weights are adjusted so as to minimize output error.

Another significant aspect in ANN-based model development is the choice of training algorithm. Despite the fact that back propagation training algorithms are quite popular, its performance is too slow as it requires small learning rates for stable learning. On the other hand, conjugate gradient (CG), quasi Newton, and Levenberg-Marquardt (LM) make use of standard optimization techniques and are considered to be quite fast as far as processing is concerned.

In this study, feed forward artificial neural network models with two hidden (intermediate) layers, trained by both back propagation (BP) neural networks and quasi-Newton-based Broyden-Fletcher-Goldfarb-Shanno (BFGS) (Fletcher 1987), have been developed for suspended sediment estimation at Besham station located in the upper Indus basin of Pakistan. The selection of hidden layers and number of nodes/neurons in each layer has been carried out using a simple trial-and-error method as too few neurons may cause under-fitting and too many may result in over-fitting problem in model development processing. Another reason for selecting two hidden layers for ANN model is their inability to solve nonlinear problems as reported by Minsky and Papert (1969) and Jones (2004). A comparison of ANN-based model output has also been made with local linear regression-based models and the best models have been evaluated which can be used for an accurate and reliable estimate of suspended sediment quantity entering in the reservoir.

Results & discussions

Gamma test and M -test

Gamma test has been performed on the weekly observations to reduce the noise for smooth model development. In general, $2^n - 1$ combinations were possible but only 200 realistic combinations/masks has been formed, and Gamma value was calculated for each. The variation of gamma values with respect to the combinations is explanatory in Fig. 2.

On the basis of minimum Gamma value criterion, 10 out of 200 combinations were selected for model development. The selected input combinations/masks with their Gamma values are listed in Table 2. It

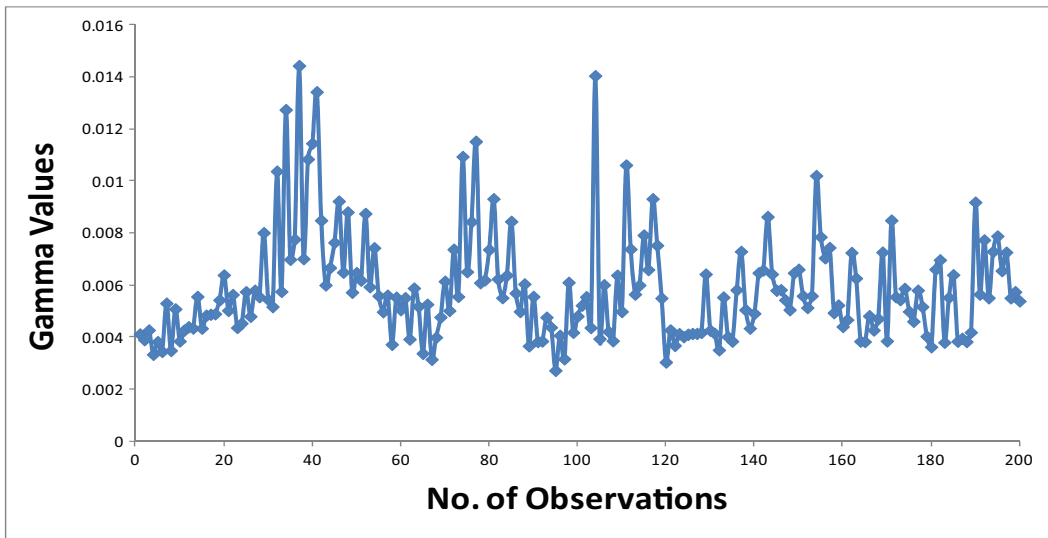


Fig. 2 Gamma values vs no. of observations

clearly shows that the Gamma values are quite less and the smooth models build on the selected data will have a standard deviation of less prediction error.

When the true noise variance is positive, the Gamma statistic should also be positive and there will come a point where using more data to build our model will not actually improve the quality of the predictions when compared with the actual values of the output. As in this case, an asymptotic gamma value against every combination is positive so the training data should be limited by using the *M*-test. *M*-test was performed, by

taking initially 10 vectors, for increasing values of data points *M* to determine the best suitable length of data. The nearest neighbors were selected 10 as by default in the WinGamma software. The selection of nearest neighbors has a relation with the length of data. For shorter data length, as in our case, 10 to 20 is a good option but if the data is lengthy then this value should be increased to get more degree of accuracy (Jones). The selected data length/no. of vectors through this test give us the best solution in order to develop best fitted models for the prediction of suspended sediment load at Besham.

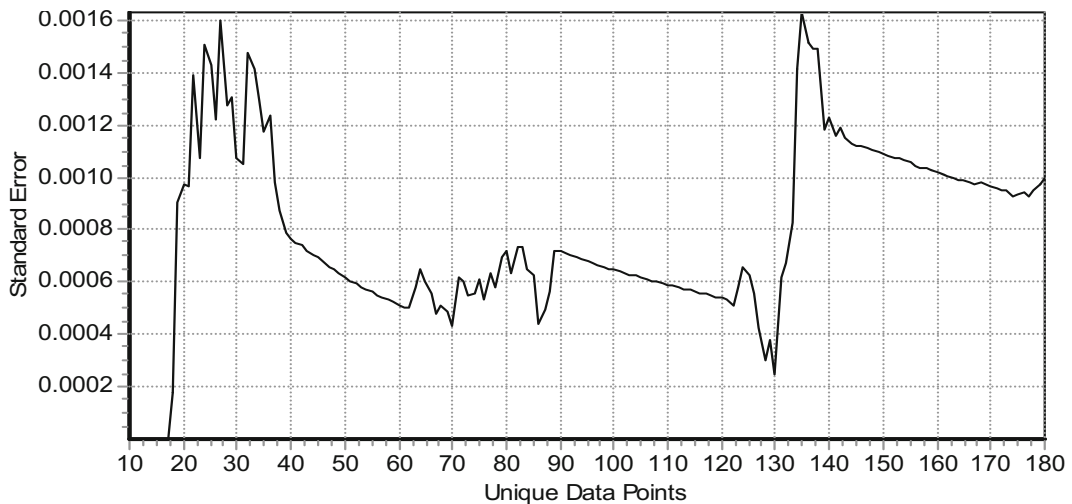
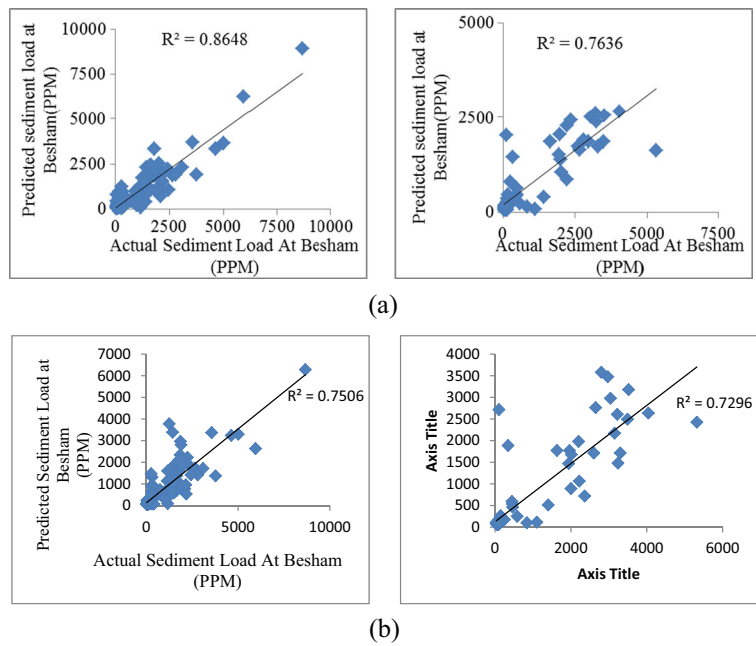


Fig. 3 Unique data points vs. standard error for experiment no. 04

Fig. 4 **a** Scatter plots for training and testing of the best model based upon BFGS (log normalization). **b** Scatter plots for training and testing of the best models based upon BFGS (Box-Cox)



This can be performed with the help of explanatory data analysis results, generated in WinGamma environment, of unique data points versus standard error. The data length was limited up to the point

where standard error line becomes stable or nearly stable as for experiment no.4, combination 11101111, the data length was limited up to 140 vectors (Fig. 3).

Fig. 5 **a** Scatter plots for training and testing of the best model based upon BP (log normalization). **b** Scatter plots for training and testing of the best model based upon BP (Box-Cox transformation)

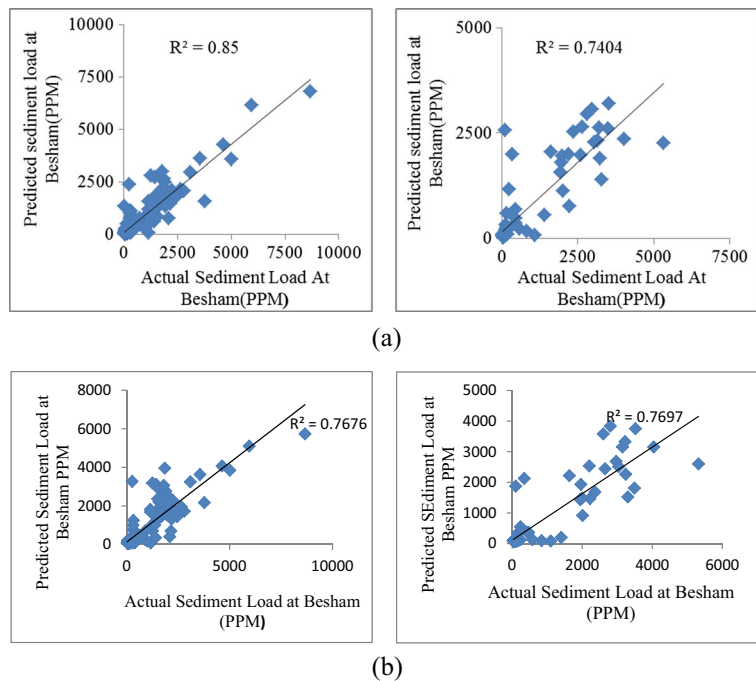
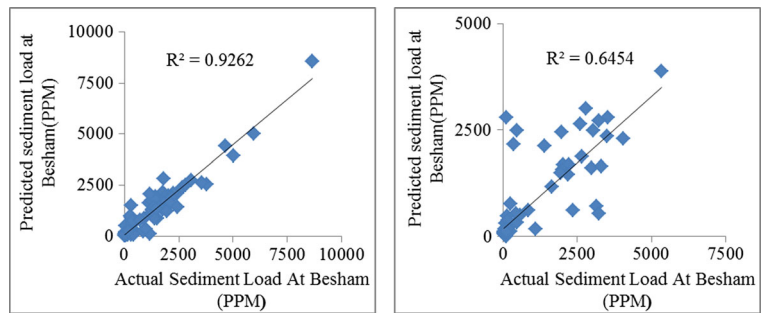


Fig. 6 Scatter plots for training and testing of the best model based upon LLR (log normalization)



Similarly, the data length has been fixed, for model development, for every selected combination on the minimum Gamma value criterion. The decided data lengths are listed in the Table 2.

Nonlinear modeling

The selected combinations on the criterion of Gamma and *M*-test were used to train nonlinear models which were included both artificial neural networks (ANNs) and local linear regression (LLR) type models.

Two types of ANN models were trained that were based upon two-layered Broyden-Fletcher-Goldfarb-Shanno (BFGS) and back propagation (BP) training algorithm. Results from these models were also compared with the local linear regression models, within the WinGamma environment.

For LLR-type models, the value of nearest neighbors (NN), which gave the minimum gamma value, was selected by trial-and-error method. In this case, the best value of NN was found to be 12. The number of hidden layers and nodes for ANN type models was also selected by hit-and-trial method to obtain the minimum values of estimated mean square error for our desired output and found to be two layers with the number of nodes 3 and 5, respectively, in the first and second layer. The structure

of developed nonlinear models is briefly explained in Table 3.

Out of three novel techniques, the most robust technique has been evaluated on the basis of maximum model efficiency and minimum systematic and random errors. Scatter plots for all developed models were drawn (Figs. 4a, b; 5a, b; and 6), and coefficient of linear model co-relation (*R*-square) was calculated for each. Furthermore, the detailed comparison of all the developed nonlinear models has been carried out on the basis of statistical parameters like root mean square error (RMSE), mean bias error (MBE), and model efficiency (*R*-square) (Tables 4 and 5).

Comparison of results

The Box-Cox transformation has been applied and the final results of developed models have been compared with the models developed on the basis of log normalization as shown in the Figs. 4 and 5. It is well clear from the figures that log normalization shows comparatively better results than that of Box-Cox transformation. According to Gamma test (Table 2), the input combination 6 (11101011) should outperform for estimation of sediment load at Besham because it has the relatively low Gamma statistic as compared to the rest. This has been verified in the case of BFGS with a noticeable high

Fig. 7 Deviation of predicted sediment load from observed load calculated through BFGS-based model

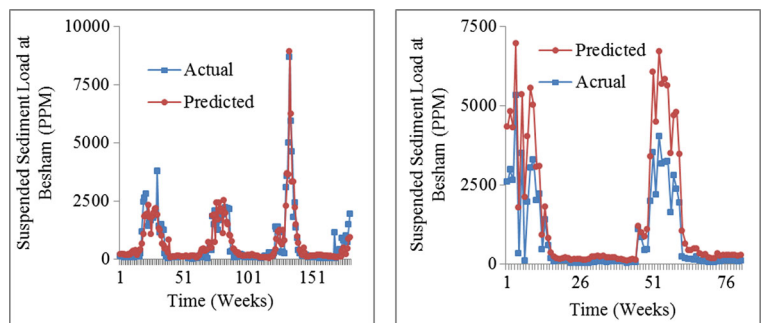


Table 2 Selected combinations on the basis of minimum Gamma value

Sr. no.	Experiment no.	Mask	Gamma values Normalization b/w 0 to 1	Gamma values Box-Cox transformation	Final vector length
1	4	11101111	0.0033304	0.00911481	140
2	6	11111011	0.0034418	0.0111108	145
3	8	11111110	0.0034786	0.0090359	138
4	65	11111010	0.0033681	0.010868	143
5	67	11111111	0.0031394	0.00952736	138
6	95	11101011	0.0027128	0.0109276	135
7	97	01101011	0.0031548	0.0106006	135
8	120	11101010	0.0030348	0.0106459	140
9	132	11011011	0.0035003	0.00937429	140
10	180	10010011	0.0036178	0.0112429	135

value of R -square, 0.864 and 0.763, low bias values -3.46 and -61 in training and testing phases, respectively (Tables 4 and 5). The negative bias values for this model show that the model is underestimating the sediment load at Besham. All models based upon LLR exposed strong linear correlation coefficient (R -square) in the training phase but failed to produce good results in the testing phase with low R -square and high root mean square errors (RMSE) as shown in the following comparison (Tables 4 and 5). It has also been observed that LLR models developed by the Box-Cox transformation showed very poor results. In the case of BP, again, test no. 6 shows good values of R -square = 0.921 in the training phase but in the testing phase it is quite less as compared to other tests in the same phase. In its place, test no. 5 has a quite reasonable value of R -square in

both training and testing phases which are 0.74 and 0.85, respectively.

The comparison clearly depicts that the BFGS-based model no. 6 outperformed all the other nonlinear models with a significant high value of R -square and low RMSE and MBE in both training and testing phases.

Figure 7 shows the deviation of predicted sediment load, for both training and testing phases, at Besham with respect to the observed sediment load based upon the evaluated best BFGS model no.6.

Summary and conclusion

The study was mainly concentrated on the estimation of suspended sediment load at Besham located just

Table 3 Structure of LLR, BFGS, and BP models

Test	LLR	BFGS		BP	
	Nearest neighbor (NN)	Nodes (layer 1)	Nodes (layer 2)	Nodes (layer 1)	Nodes (layer 2)
1	12	3	5	3	5
2	12	3	5	3	5
3	12	3	5	3	5
4	12	3	5	3	5
5	12	3	5	3	5
6	12	3	5	3	5
7	12	3	5	3	5
8	12	3	5	3	5
9	12	3	5	3	5
10	12	3	5	3	5

Table 4 Training phase: model comparison results (log normalization)

Model	LLR			BP			BFGS		
	<i>R</i> -square	BIAS	RMSE	<i>R</i> -square	BIAS	RMSE	<i>R</i> -square	BIAS	RMSE
1	0.858	29.38	457.83	0.84	<i>-1.43</i>	464.01	0.84	92.57	472
2	0.915	<i>-11.56</i>	339.84	0.837	-21.68	470.66	0.843	10.19	460.63
3	0.838	82.07	489.59	0.839	-10.43	467.92	0.837	-7.25	468.96
4	<i>0.938</i>	-12.75	<i>287.37</i>	0.839	<i>-33.52</i>	<i>467.57</i>	0.834	-21.38	472.68
5	0.907	-39.4	362.16	0.85	-34.88	450.66	0.848	-28.54	453.12
6	0.932	-13.94	302.5	<i>0.921</i>	<i>-205.13</i>	391.62	<i>0.864</i>	<i>-3.24</i>	<i>426.39</i>
7	0.926	-41.75	321.4	0.851	7.39	448.31	0.858	75.69	445.24
8	0.91	-18.18	348.15	0.847	-23.29	454.54	0.85	25.94	449.2
9	0.887	27.15	392.56	0.833	-13.4	474.29	0.837	-50.57	470.58
10	0.871	26.69	420	0.909	143.72	<i>386.69</i>	0.829	25	482.06

Italicized values present the highest value of each statistical parameter column-wise

upstream of Tarbela in order to check the variation of entering suspended sediment load to the Tarbela reservoir with respect to other meteorological variables. For this purpose, two different neural network techniques (two-layer BP and BFGS) along with LLR technique were used in order to develop a variety of models. Basically, the estimations are based upon antecedent weekly observations of meteorological data including discharge, temperature, and suspended sediment load itself. It has been exposed that the methods adopted provide satisfactory estimations as reflected in the evaluation criteria (graphs and tables). Gamma test was performed on 200 realistic combinations, and out of these, ten best combinations were selected for model

development. Furthermore, the data length selection was carried out using *M*-test. Input combinations and dataset length selected by these tests were then used for development of LLR models, BP neural network models, and BFGS-based ANN models for the prediction of suspended sediment load at Besham. Evaluation of these models was performed using different error assessment parameters including *R*-square, MBE, and RMSE.

The best input combination for LLR has a mask **(01101011)** [(Q_{SHA}) , (T_{SHA}) , (S_{SHA}) , (Q_{BUN}) , (T_{BUN}) , (S_{BUN}) , (Q_{BESH}) , (T_{BESH})]. For two-layer BP model, **(11111111)** combination was found to be the best by using all inputs at the same time [(Q_{SHA}) , (T_{SHA}) ,

Table 5 Testing phase: model comparison results (Box-cox transformation)

Model	LLR			BP			BFGS		
	<i>R</i> -square	BIAS	RMSE	<i>R</i> -square	BIAS	RMSE	<i>R</i> -square	BIAS	RMSE
1	0.50	23.43	1065.4	0.761	-227.02	719.86	0.756	-111.86	682.48
2	0.487	-84.5	1021.31	0.733	-160.09	704.44	0.73	-128.33	701.03
3	0.217	-168.19	1794.49	0.726	-148.06	684	0.734	-137.67	670.62
4	0.401	-77.81	<i>1116.76</i>	0.692	-177.7	729.6	0.703	-134.22	703.07
5	0.178	27.51	2103.08	0.74	<i>-140.85</i>	<i>668.57</i>	0.738	-127.18	<i>667.16</i>
6	0.428	-43.01	1090.1	0.677	-248.88	761.75	<i>0.763</i>	-60.9	692.93
7	<i>0.645</i>	-149.16	<i>768.06</i>	0.751	-144.11	701.58	0.73	-85.53	729.41
8	0.485	26.08	1007.41	<i>0.768</i>	-219.22	683.23	0.737	-172.55	694.72
9	0.557	-30.29	875.32	0.66	-150.9	759.71	0.688	-169.84	729.15
10	0.40	-68.75	1098.17	0.379	449.36	1735	0.644	-58.8	759.8

Italicized values present the highest value of each statistical parameter column-wise

$(S_{SHA}), (Q_{BUN}), (T_{BUN}), (S_{BUN}), (Q_{BESH}), (T_{BESH})$]. And finally for BFGS model, **(11101011)** combination was found to be the best for inputs $[(Q_{SHA}), (T_{SHA}), (S_{SHA}), (Q_{BUN}), (T_{BUN}), (S_{BUN}), (Q_{BESH}), (T_{BESH})]$.

The study demonstrated that the ANN-based models can be used for estimation of suspended sediment load at the upper Indus basin to a reasonably high degree of accuracy. The study could also prove to be the pioneer in initiating other research projects that involves the use of data-driven modeling within the Indus basin. It also demonstrates that Gamma test has the ability to identify the best input combination and data length for development of smooth ANN models. It is recommended that data length and no. of stations along with input parameters must be increased for more precise and accurate results.

References

- Agalbjorn, S., Koncar, N., & Jones, A. J. (1997). A note on the gamma test. *Neural Computing and Applications*, 5, 131–133.
- Agarwal, A., Mishra, S. K., Ran, S., & Singh, J. K. (2006). Simulation of runoff and sediment yield using artificial neural network. *Biosystems Engineering*, 94(4), 597–613.
- Akrasi, S. A., & Ansa-Asare, O. D. (2003). Assessing sediment and nutrient transport in the Pra basin of Ghana. *West African Journal of Applied Ecology*, 13.
- Alam, U. Z. (1998). *Water rationality: mediating the Indus waters treaty*. University of Durham: Unpublished thesis.
- Ali, K. F., & De Boer, D. H. (2003). Construction of sediment budgets in large-scale drainage basins: the case of the upper Indus river. *International Association of Hydrological Sciences Publication*, 279, 206–215.
- Alizadeh, M. J., Joneyd, P. M., Motahhari, M., Ejlali, F., & Kiani, H. (2015). A wavelet ANFIS model to estimate sedimentation in dam reservoir. *International Journal of Computer Applications*, 114(9), 19–25.
- Alp, M., & Cigizoglu, H. K. (2005). Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. *Environmental Modelling & Software*, 22(1), 2–13.
- Azamatulla, H. M., Ab, G. A., Chang, C. K., Abu, H. Z., & Zakaria, N. A. (2010). Machine learning approach to predict sediment load—a case study. *Clean-Soil Air Water*, 38(10), 969–976.
- Bhattacharya, B., Price, R. K., & Solomatine, D. P. (2007). Machine learning approach to modelling sediment transport. *Journal of Hydraulic Engineering*, 133(4), 440–450.
- Cigizoglu, H. K. (2004). Estimation and forecasting of daily suspended sediment data by multi-layer perceptrons. *Advances in Water Resources*, 27, 185–195.
- Cobaner, M., Unal, B., & Kisi, O. (2009). Suspended sediment concentration estimation by an adaptive neuro-fuzzy and neural network approaches using hydrometeorological data. *Journal of Hydrology*, 367, 52–61.
- Environmental Water Resources Institute, U.W.R.R.C., Gross Solids Technical Committee. (2007). *ASCE guideline for monitoring storm water gross solids*.
- Evans, D. (2002). Database searches for qualitative research. *Journal of the Medical Library Association*, 90(3), 290–293.
- Fine T. L. (1999) “Feedforward neural network methodology” (statistics for engineering and information science). 340 pp. Springer
- Fletcher, R. (1987). *Practical methods of optimization*. New York, N.Y.: Wiley.
- Tayfur, G. (2002). Artificial neural networks for sheet sediment transport. *Hydrological Sciences Journal*, 47(6), 879–892.
- Habib-ur-Rehman, M. (2001). *Regional scale soil erosion and sediment transport modeling*. Ph.D. Thesis, University of Tokyo, pp. 183.
- Hassan, M., Shamim, M. A., Hashmi, H. N., Ashiq, S. Z., Ahmed, I., Pasha, G. A., Naeem, U. A., Ghumman, A. R., & Han, D. (2014). Predicting streamflows to a multipurpose reservoir using artificial neural networks and regression techniques. *Earth Science Informatics*, 8(2), 337–352.
- Haykin, S. (1999). *Neural networks: a comprehensive foundation* (2nd ed.,). USA: Prentice-Hall.
- Hays, D., & Wu, P. (2001). “Simple approach to TSS source strength estimates” *proceedings, 21st annual meeting of the western dredging association (WEDA XXI) and 33rd annual Texas A&M dredging seminar*. Texas: Houston.
- Heng, S., & Suetsugi, T. (2013). Using artificial neural network to estimate sediment load in ungauged catchments of the tonle sap river basin, Cambodia. *Journal of Water Resource and Protection*, 5, 111–123.
- Jain, A., Sudheer, K. P., & Srinivasulu, S. (2004). Identification of physical processes inherent in artificial neural network rainfall runoff models. *Hydrological Processes*, 18, 571–581.
- Jason W. Osborne, (2010) Improving your data transformations; applying the box-Cox transformation Practical Assessment, Research & Evaluation 15(12)
- Jie L.C., Yu S.T., (2011) “Suspended sediment load estimate using support vector machines in Kaoping River basin” 978–1–61284–459–6/11 2011 IEEE
- Jones, A. J. (2004). New tools in non-linear modelling and prediction. *Computational Management Science*, 1, 109–149.
- Kisi, Ö. (2004a). River flow modeling using artificial neural networks. *Journal of Hydrologic Engineering*, 9(1), 60–63.
- Kisi, O. (2004b). Multi-layer perceptrons with levenberg–Marquardt training algorithm for suspended sediment concentration prediction and estimation. *Hydrological Sciences Journal*, 49(6), 1025–1040.
- Kisi, O. (2005). Suspended sediment estimation using neuro-fuzzy and neural network approaches. *Hydrological Sciences Journal*, 50(4), 683–696.
- Kisi, O., Yuksel, I., & Dogan, E. (2008). Modeling daily suspended sediment of rivers in Turkey using several data-driven techniques. *Hydrological Sciences Journal*, 53(6), 1270–1285.
- Kux, D. (2006). India Pakistan negotiations: is past still prologue? In *Perspective Series*. Washington, DC: United States Institute of Peace.
- Lane, L. J., & Nicholas, M. H. (1997). A hydrologic method for sediment transport and yield. In S. Y. Wang, E. J. J.

- Langendean, & F. D. Shields, Jr. (Eds.), *Management of Landscapes Disturbed by Channel Incision* (pp. 365–370). University of Mississippi, Oxford, MS: Center for Comp. Hydrosoci. And Eng.
- Lin, B., & Namin, M. M. (2005). Modelling suspended sediment transport using an integrated numerical and ANNs model. *Journal of Hydraulic Research*, 43(3), 302–310.
- 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. *Journal of Hydroinformatics*, 3, 153–164.
- Maier, H. R., & Dandy, G. C. (1998). Understanding the behaviour and optimising the performance of back-propagation neural networks: an empirical study. *Environmental Modelling and Software*, 13, 179–191.
- Minsky, M., & Papert, S. (1969). *Perceptrons*. Cambridge, MA: MIT Press.
- McCulloch, W. S., & Pitts, W. H. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 7, 115–133. Reprinted in McCulloch 1964, pp. 16–39.
- Mustafa, M. R., Isa, M. H., & Rezaur, R. B. (2011). Comparison of artificial neural network for prediction of suspended sediment discharge in river—a case study in Malaysia. *World Academy of Science, Engineering and Technology*, 81, 372–376.
- Nagy, H. M., Watanabe, K., & Hirano, M. (2002). Prediction of sediment load concentration in rivers using artificial neural network model. *Journal of Hydraulic Engineering (ASCE)*, 128(6), 588–595.
- Nakai O. (1978) “Turbidity generated by dredging projects, management of bottom sediments containing toxic substances” Proceedings of the third United States-Japan experts meeting, EPA-600/3-78-084, 1–47
- Nourani, V. (2009). “Using artificial neural networks (ANNs) for sediment load forecasting of talkherood river mouth”. *Journal of Urban and Environmental Engineering*, 3(1), 1–6.
- Nourani, V., Parhizkar, M., Vousoughi, F. D., & Amini, B. (2014). Capability of artificial neural network for detecting hysteresis phenomenon involved in hydrological processes. *Journal of Hydrologic Engineering*, 19, 896–906.
- Olyaie, E., Banejad, H., Chau, K. W., & Melesse, A. M. (2015). “A comparison of various artificial intelligence approaches performance for estimating suspended sediment load of river systems: a case study in United States”. *Environmental Monitoring and Assessment*, 187(4), 189.
- Painter, D. J. (1981). Steeplands erosion processes and their assessment. In T. R. H. Davis, & A. J. Pearce (Eds.), *Erosion and sediment transport in Pacific Rim steepplands. IAHS_AISH publ. 132* (pp. 2–20). Washington, DC: International Association of Hydrological Sciences.
- Partal, T., & Cigizoglu, H. K. (2008). Estimation and forecasting of daily suspended sediment data using wavelet-neural networks. *Journal of Hydrology*, 358, 317–331.
- Pennekamp, J. G. S., Epskamp, R. J. C., Rosenbrand, W. F., Mullie, A., Wessel, G. L., Arts, T., & Deibel, I. K. (1996). Turbidity caused by dredging; viewed in perspective. *Terra et Aqua*, 64, 10–17.
- Rajae, T., Mirbagheri, S. A., Zounemat-kermani, M., & Nourani, V. (2009). Daily suspended sediment concentration simulation using ANN and neuro-fuzzy models. *The Science of the Total Environment*, 407, 4916–4927.
- Remesan, R., Shamim, M. A., & Han, D. (2008). Model data selection using gamma test for daily solar radiation estimation. *Hydrological Processes*, 22, 4301–4309.
- Remesan, R., Shamim, M. A., Han, D., & Mathew, J. (2009). Runoff prediction using an integrated hybrid modelling scheme. *Journal of Hydrology*, 372(1), 48–60.
- Remesan, R., Shamim, M. A., Ahmadi, A., & Han, D. (2010). Effect of data time interval on real-time flood forecasting. *Journal of Hydroinformatics*, 12(4), 396–407.
- Roesner, L., Qian, Y., Criswell, M., Stromberger, M., & Klein, S. (2006). *Long-term effects of landscape irrigation using household graywater— literature review and synthesis*. Water Environment Research Foundation: Resource Document.
- Rosenblatt, F. (1962). *Principles of neurodynamics; perceptrons and the theory of brain mechanisms*. Spartan Books: Washington.
- Shamim, M. A., Hassan, M., Ahmad, S., & Zeeshan, M. (2015). A comparison of artificial neural networks (ANN) and local linear regression (LLR) techniques for predicting monthly reservoir levels. *Journal of Korean Society for Civil Engineering*(in Press). doi:10.1007/s12205-015-0298-z.
- Shamim, M. A., Bray, M., Remesan, R., & Han, D. (2014). A hybrid modelling approach for assessing solar radiation. *Theoretical and Applied Climatology*(in Press). doi:10.1007/s00704-014-1301-1.
- Shamim, M. A., Remesan, R., Han, D., & Ghumman, A. R. (2010). Solar radiation estimations in ungauged catchments. Proceedings of the ICE. *Water Management*, 163(7), 349–359.
- Shamim, M. A., Ghumman, A. R., & Ghani, U. (2004). Forecasting groundwater contamination using artificial neural networks. In *International conference on water resources and arid environments, December 2004*. Riyadh: Saudi Arabia: King Saud University.
- Shamseldin, A. Y. (2010). Artificial neural network model for river flow forecasting in a developing country. *Journal of Hydroinformatics*, 12, 22–35.
- Stefánsson, A., Končar, N., & Jones, A. J. (1997). A note on the gamma test. *Neural Computing Applications*, 5, 131–133.
- Tayfur, G., & Guldal, V. (2006). Artificial neural networks for estimating daily total suspended sediment in natural streams. *Nordic Hydrology*, 37, 69–79.
- Weaver, W., & Madej, M. A. (1981). Erosion control techniques used in Redwood National Park, Northern California 1978–1979. In *Proceedings of a Symposium on Erosion and Sediment Transport in Pacific Rim Steeplands, Christchurch, New Zealand* (pp. 640–645). Washington, DC: International Association of Hydrological Sciences Publication Number 132.
- Zhu, Y. M., Lu, X. X., & Zhou, Y. (2007). “Suspended sediment flux modeling with artificial neural network: an example of the Longchuanjiang River in the Upper Yangtze Catchment. China” *Geomorphology*, 84, 111–125.