

# A Hybrid Double Feedforward Neural Network for Suspended Sediment Load Estimation

Xiao Yun Chen<sup>1</sup> · Kwok Wing Chau<sup>1</sup>

Received: 25 May 2015 / Accepted: 1 March 2016 /

Published online: 7 March 2016

© Springer Science+Business Media Dordrecht 2016

**Abstract** Estimation of suspended sediment loads (*SSL*) in rivers is an important issue in water resources management and planning. This study proposes a hybrid double feedforward neural network (HDFNN) model for daily *SSL* estimation, by combining fuzzy pattern-recognition and continuity equation into a structure of double neural networks. A comparison is performed between HDFNN, multi-layer feedforward neural network (MFNN), double parallel feedforward neural network (DPFNN) and hybrid feedforward neural network (HFNN) models. Based on a case study on the Muddy Creek in Montana of USA, it is found that the HDFNN model is strongly superior to the other three benchmarking models in terms of root mean squared error (RMSE) and Nash-Sutcliffe efficiency coefficient (NSE). HDFNN model demonstrates the best generalization and estimation ability due to its configuration and capability of physically dealing with different inputs. The peak value of *SSL* is closely estimated by the HDFNN model as well. The performances of HDFNN model in low and medium loads are satisfactory when investigated by partitioning analysis. Thus, the HDFNN is appropriate for modeling the sediment transport process with nonlinear, fuzzy and time-varying characteristics. It explores a practical alternative for use and can be recommended as an efficient estimation model for *SSL*.

**Keywords** Hybrid neural network · Double parallel feedforward · Suspended sediment load · River flow · Differential evolution

## 1 Introduction

The estimation of suspended sediment loads (*SSL*) is required in river restoration, stable channel design and water quality assessment. It is a difficult and sophisticated task in practice, however, since the sediment transport is highly nonlinear and

---

✉ Kwok Wing Chau  
cekwchau@polyu.edu.hk

<sup>1</sup> Department of Civil and Environmental Engineering, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, People's Republic of China

governed by many factors including strength of flow, sediment supply, river bed, etc. Conventional sediment rating curves (SRC) are incapable of providing sufficiently accurate estimates attributed to the misleading practice of using sediment load versus discharge (McBean and Al-Nassri 1988; Demirci and Baltaci 2013). Artificial intelligence techniques have been proven to be efficient tools in modeling sediment loads. Alp and Cigizoglu (2007) employed two artificial neural network (ANN) models, i.e. feedforward back propagation (FFBP) method and radial basis functions (RBF) to estimate daily *SSL*. The use of support vector machine (SVM) was investigated by Cimen (2008) for *SSL* estimation in rivers. Lafdani et al. (2013) used a combination of gamma test and genetic algorithm (GT-GA) to identify the best input of SVM and ANN models for daily *SSL* prediction. These models could capture the nonlinear behavior of sediment data without going into the details of physical processes in watershed. Nevertheless, in reverse, the totally implicit and physically meaningless features are also the major criticisms. It is still necessary to develop estimation models with conceptual ideas to reflect the characteristics of sediments.

The fuzzy nature of *SSL* series necessitates the utilization of fuzzy and highly nonlinear methods for sediment simulation. Fuzzy logic was accepted as a good procedure in suspended sediment estimation (Tayfur et al. 2003; Kisi et al. 2006; Demirci and Baltaci 2013). It yielded better results than SRC and ANN models since the degree of 'belongingness' to a set or category is described by a membership number. Neuro-fuzzy models for suspended sediment estimation were found to provide good performances as well (Cobaner et al. 2009; Kisi et al. 2009; Rajaei et al. 2009). However, these approaches suffer from difficulties in their manipulation as they need different membership functions in different cases. A flexible and transparent model which allows implementing the fuzzy concept in activation functions is appreciated. Qiu et al. (1998) introduced a fuzzy pattern-recognition activation function (from the input layer to the hidden layer) into an ANN model for annual runoff forecasting. This function classified inputs into a number of categories in terms of different patterns. In this way, the fuzziness of runoff was considered with respect to the seasonal characteristic of the river system. Zhao and Chen (2008) further applied this model for predictions in ungauged basins using hydrological data in gauged basins that were similar. In their study, the fuzzy pattern-recognition activation function was employed to connect the hidden nodes and the network output. This method offers practically significant advantage over other fuzzy-based models and is employed in this study for *SSL* estimation.

In addition, the time-varying nature of sediment transport process can be considered by adding a continuity equation in the ANN structure, as inspired by Yang et al. (1998) and Li and Gu (2003). In their works, the nodes in the hidden/output layers were regarded as storage reservoirs, and continuity equation was satisfied when river flows from upstream to downstream sections. Yang et al. (1998) successfully forecasted monthly river flow of Salford University station in Irwell River basin. Li and Gu (2003) expanded this method to the sediment yield forecasting. They obtained satisfactory results and encouraged the use of continuity equation in modeling the sediment loads. The spatial and temporal factors were taken into account in the sediment transport process by continuity equation, which can shed light on the effect of upstream sediment loads. Thus, this method can build a relationship between upstream and downstream sediment loads, and is feasible and acclaimed in an *SSL*

estimation model. It is preferred to completely physics-based approaches in which the detailed environmental data are generally not available and simplified assumptions are unrealistic (Kothyari et al. 1997; Kouassı et al. 2013).

Traditional multi-layer feedforward neural network (MFNN) has some drawbacks in its architecture and regularization. He (1993) proposed double parallel feedforward neural network (DPFNN), which involves a paratactic relationship between linear and nonlinear mapping. It is a parallel connection of a multi-layer feedforward neural network and a two-layer feedforward neural network. The multi-layer network uses its hidden nodes to adjust the solution and thus improves nonlinear mapping performance; and the two-layer network can give high learning speed for linear solution (He 1994). It was demonstrated that DPFNN has faster convergence speed and better generalization capability than MFNN (Zhong and Ding 2005; Wang et al. 2011). When using particle swarm optimization for feature selection, the DPFNN model could rectify over-fitting problem as well (Huang and He 2007). It has been used for hyperspectral data classification (He and Huang 2005), concentration estimation of gas mixture (Zhao et al. 2010) and water diversion demand estimate (Khan et al. 2014), which has been proved to be a promising method for regression and prediction.

The purpose of this paper is to develop a novel estimation model with a combination of fuzzy pattern-recognition, continuity equation and double feedforward neural network. In addition to river flows, the influence of sediment loads in the upstream river sections is investigated in this study. Two sediment stations on the Muddy Creek in Montana of USA are used as case study sites.

## 2 Description of Estimation Models

### 2.1 Multi-Layer Feedforward Neural Network (MFNN)

The three-layer feedforward neural network consisting of the input, hidden and output layers, is the most widely used MFNN model. The input layer introduces input data  $\{p_1, p_2, \dots, p_k\}$  to the network. The weighted sum of inputs and bias are passed with a predetermined activation function  $f(\cdot)$  to the nodes in the hidden layer (Thirumalaiah and Deo 1998):

$$t_i = f\left(\sum_{j=1}^k p_j w_{ji} + b_i\right) \quad (1)$$

where  $t_i$  ( $i = 1, 2, \dots, s$ ) represent nodes in the hidden layer and  $p_j$  ( $j = 1, 2, \dots, k$ ) represent nodes in the input layer. The weight parameter connecting the input layer and the hidden layer is denoted by  $w_{ji}$ , and  $b_i$  is the bias value. Similarly, the node in the output layer is computed from nodes in the hidden layers (Thirumalaiah and Deo 1998):

$$y = F\left(\sum_{i=1}^s t_i \bar{w}_i + \bar{b}\right) \quad (2)$$

in which  $y$  represents a single node in the output layer and  $F(\cdot)$  is the activation function for the output layer. The weight parameters from the hidden layer to the output layer and bias are denoted by  $\bar{w}_i$  and  $\bar{b}$ , respectively. For traditional MFNN models, the activation function  $f(\cdot)$  is usually a radial basis function or sigmoid

function, and  $F(\cdot)$  is a linear function, respectively. They reveal relation of nodes between two layers, although having no physical meanings. The MFNN model for *SSL* estimation has limitations attributed to the negligence of sediment properties.

## 2.2 Double Parallel Feedforward Neural Network (DPFNN)

As can be seen in Fig. 1a, DPFNN model is developed from MFNN model in which two networks connect each other in parallel with the same  $k$  input nodes. For the three-layer neural network of DPFNN, the nodes in the hidden layer ( $t_1, t_2, \dots, t_s$ ) are computed by Eq. (1) and then connected to the output with  $\bar{w}_i$  in the same manner. Analogously for the two-layer neural network, the weight parameters directly from the input layer to the output layer are denoted by  $v_j$  ( $j=1, 2, \dots, k$ ). The node in the output layer is acquired in the following equation (Zhong and Ding 2005):

$$y = \sum_{i=1}^s t_i \bar{w}_i + \sum_{j=1}^k p_j v_j \quad (3)$$

That is, the output is a summary of two parallel neural networks. The procedure of computing  $y$  from its inputs is demonstrated in Fig. 1a, for the MFNN and DPFNN models respectively, whilst the calibration process for searching optimized parameters is shown in Fig. 1b. For a given set of training samples ( $p^n, Y^n$ ) supplied to the model, the error function is defined as:

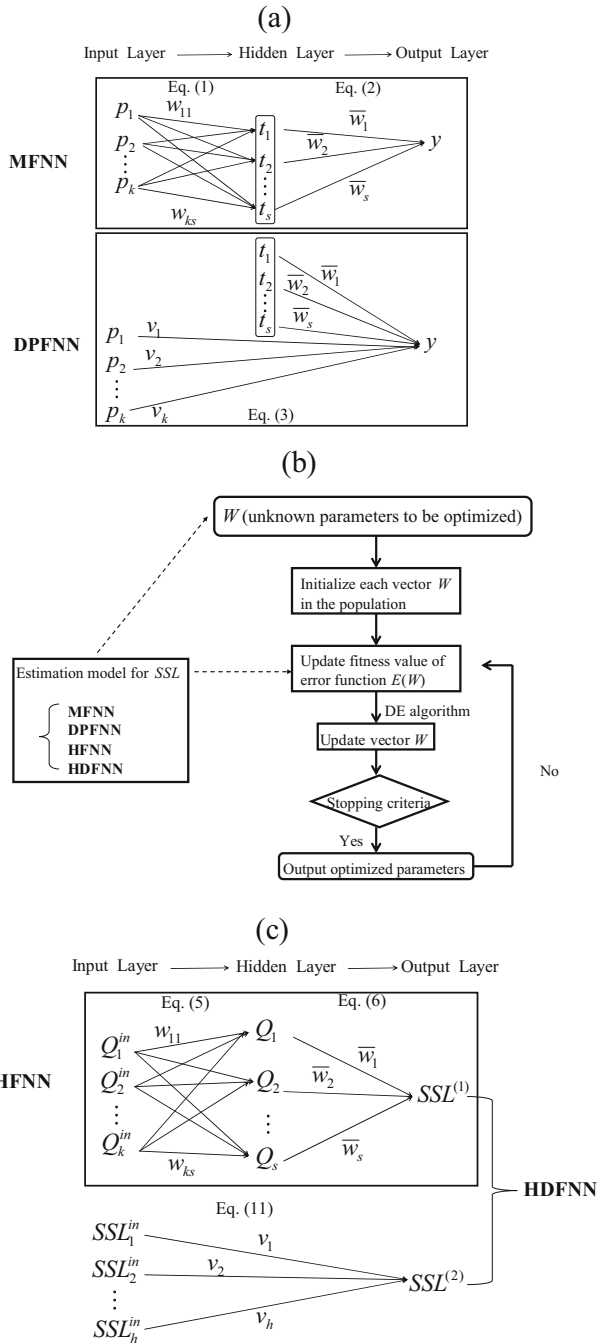
$$E(W) = \sqrt{\frac{1}{N} \sum_{n=1}^N (y^n - Y^n)^2} \quad (4)$$

where the vector  $W$  is a collection of all unknown parameters, and varies with the estimation model;  $y^n$  and  $Y^n$  are computed and desired output ( $n=1,2,\dots, N$ ), respectively, and  $N$  is the number of training samples. The objective of network training, hence, is to find  $W_{opt}$  which satisfies that  $E(W_{opt}) = \min E(W)$ . As shown in Fig. 1b, the vector  $W$  is updated with the updated fitness value of  $E(W)$  and is finally outputted if stopping criteria is satisfied. In the present paper, differential evolution (DE) is employed as an optimization technique to find the minimum value of error function and the corresponding  $W_{opt}$ . The DE is a widely used population-based optimization algorithm, which is favourable for searching parameters of non-differentiable and time-varying models (Storn and Price 1995; Rocca et al. 2011; Li et al. 2013). It conducts mutation, crossover and selection operations based on the differences of randomly sampled pairs of solutions in the population, thus avoids local optima and allows fast convergence, details of which can be found in Chen et al. (2015).

## 2.3 Hybrid Feedforward Neural Network (HFNN)

The above two models are incapable of distinguishing the influences of different inputs, thus  $\{p_1, p_2, \dots, p_k\}$  is employed to represent any potential inputs for *SSL* estimation. In practice, some previous studies estimated sediment based on the river

**Fig. 1** (a) Topological structure of MFNN and DPFNN models (b) Calibration process using DE algorithm for the estimation models (c) Topological structure of HFNN and HDFNN models



flow and sediment data at its own station (Aytek and Kisi 2008; Afan et al. 2014), while others focused on the estimation of downstream sediment data by using data

from both upstream and downstream stations (Kisi 2004; Partal and Cigizoglu 2008). For the case of this study, river flows  $Q$  either at the upstream or downstream stations and  $SSL$  from upstream stations are involved as inputs. When fed with various inputs, the output  $SSL$  at the downstream station is obtained in different manners.

In this section, a hybrid feedforward neural network (HFNN) is developed with respect to river flow inputs ( $Q_1^{in}, Q_2^{in}, \dots, Q_k^{in}$ ). A conceptual activation function based on fuzzy pattern-recognition is introduced as follows (Qiu et al. 1998):

$$Q_i = \frac{1}{\sum_{l=1}^c \frac{\sum_{j=1}^k [w_{ji}(Q_j^{in} - M_l)]^2}{\sum_{j=1}^k [w_{ji}(Q_j^{in} - M_l)]^2}} \tag{5}$$

where  $Q_i$  ( $i=1, 2, \dots, s$ ) are nodes in the hidden layer and  $Q_j^{in}$  ( $j=1, 2, \dots, k$ ) are nodes in the input layer. Model vector  $M=[M_i]=[M_i]$  contains a number of patterns in the hidden layer. It entertains fuzzy pattern-recognition idea in the hidden layer, since the inputs are classified into a number of categories in terms of different patterns. The parameter  $C$  refers to the number of elements in the model vector as well as the number of nodes in the hidden layer (i.e.  $C=s$ ). Generally, a higher value of  $C$  generates a higher precision for the estimation result, since it implies that there are more categories in the hidden layer and represents a higher degree of nonlinearity. We further give a general expression for the vector  $M$ : if the number of the nodes in the hidden layer equals to  $C$  ( $\geq 2$ ), then  $M = (1.0, \frac{C-2}{C-1}, \frac{C-3}{C-1}, \dots, \frac{1}{C-1}, 0)$ . The degree of membership is 1.0 for “wet” model in wet season and 0 for “dry” model in dry season, thus, the defined vector would fully cover the models ranging from “wet” to “dry” season. Meanwhile, the activation function from the hidden layer to the output layer is given as follows:

$$SSL^{(1)} = a_0 \times \left( \sum_{i=1}^s Q_i \bar{w}_i + \bar{b} \right)^{b_0} \tag{6}$$

where  $SSL^{(1)}$  represents node in the output layer;  $\bar{w}_i$  and  $\bar{b}$  denote the weight parameters and bias for the output layer, respectively. The activation function in Eq. (6) expresses an exponential relationship between river flows and sediment loads, which is generally a functional relationship representing the SRC. Values of  $a_0$  and  $b_0$  for a specific river are to be optimized in the training process of neural network. The structure of HFNN model is depicted in the framework of Fig. 1c, where  $SSL^{(1)}$  is considered as the final output with inputs ( $Q_1^{in}, Q_2^{in}, \dots, Q_k^{in}$ ). Accordingly, HFNN model examines the relationship of  $Q$  and  $SSL$  by considering the fuzzy property of sediment loads in an MFNN structure.

### 2.4 Hybrid Double Feedforward Neural Network (HDFNN)

In this section, a hybrid double feedforward neural network (HDFNN) is developed when the sediment data at the upstream river stations are included as inputs. These sediment inputs directly work on the output in a two-layer neural network. In the representation of a river system, upstream stations are regarded as nodes in the input layer and downstream station as node in the output layer. Thus, mass conservation is satisfied over the river network by the following continuity equation (Li and Gu 2003):

$$\frac{\partial SD}{\partial T} = \sum_{i=1}^h v_i Q_i^s - Q^s \tag{7}$$

where  $SD$  and  $Q^s$  are respectively the sediment deposition and sediment transport rate at the downstream station, and  $T$  is time. Meanwhile,  $Q_i^s$  is the sediment transport rate at each upstream station, wherein  $i$  (1, 2, ...,  $h$ ) refers to the index of each node in the input layer. The fraction of sediment from a node in the input layer entering into the node in the output layer is denoted by  $v_i$ . In the physical point of view, Eq. (7) implies that the rate of change of sediment deposition in the current river section is determined by the difference with the source sediment transport rate at the upstream river reaches, which reveals the sediment mass conservation over the entire river system. After discretization, the  $SD$  at time  $T + \Delta T$  is determined by the following equation:

$$SD_{(T+\Delta T)} = SD_{(T)} + \left( \sum_{i=1}^h v_i Q_{i(T)}^s - Q_{(T)}^s \right) \times \Delta T \tag{8}$$

Multiplying the sediment transport rate  $Q^s$  by a time step  $\Delta T$  produces a change in the mass during the time step, thus daily  $SSL$  could be denoted by the equation  $SSL = Q^s \times \Delta T$  when  $\Delta T = 1$  day. Accordingly, Eq. (8) could be written as follows:

$$SD_{(T+1)} = SD_{(T)} + \left( \sum_{i=1}^h v_i SSL_{i(T)} - SSL_{(T)} \right) \tag{9}$$

Equation (9) in its simplified form is given by:

$$SD_{(T+1)} = \lambda_{(T)} \times (SD_{(T)} + P_{(T)}) \tag{10}$$

wherein  $\lambda_{(T)} = 1 - \frac{SSL_{(T)}}{SD_{(T)} + P_{(T)}}$  and  $P_{(T)} = \sum_{i=1}^h v_i SSL_{i(T)}$ . Here  $\lambda$  could be regarded as a recession coefficient, which is assumed to be independent of time (Yang et al. 1998). An initial value of sediment deposition  $SD_0$  is given in advance, and the value of  $SD$  at each time step could be computed from Eq. (10). The  $SSL$  in the output layer is evaluated as a nonlinear function of sediment deposition (Li and Gu 2003)

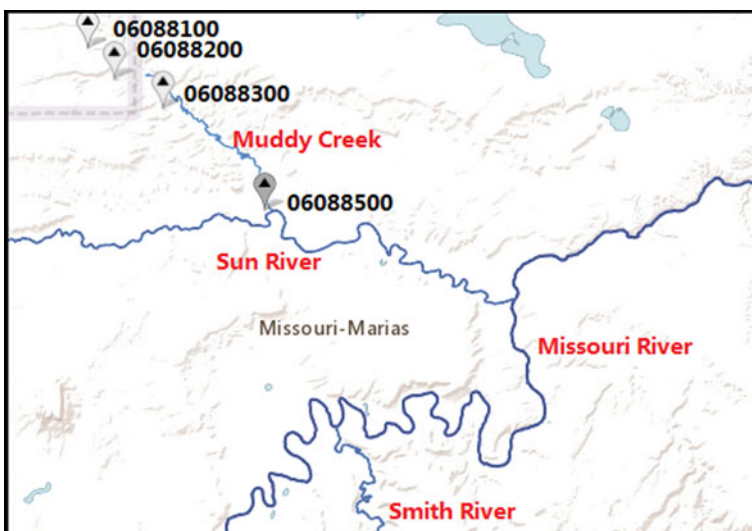
$$SSL_{(T)}^{(2)} = \frac{1}{1 + \exp[-(SD_{(T)} + P_{(T)})]} \tag{11}$$

The HDFNN model adopts two separate neural networks with different influences of river flows and upstream sediment loads on downstream  $SSL$ , which is different from the DPFNN model using the same input variables in two parallel networks. This is tantamount to say that two neural networks with respect to  $(Q_1^{in}, Q_2^{in}, \dots, Q_k^{in})$  and  $(SSL_1^{in}, SSL_2^{in}, \dots, SSL_h^{in})$  are involved, as shown in Fig. 1c. The final output  $SSL$  is a summary of  $SSL^{(1)}$  and  $SSL^{(2)}$ . Accordingly, HDFNN model allows for dealing with two separate inputs due to the double networks used. Besides, the inclusion of fuzzy pattern-recognition and continuity equation in the neural networks enables consideration of fuzzy and time-varying feature of sediment loads.

### 3 Study Area

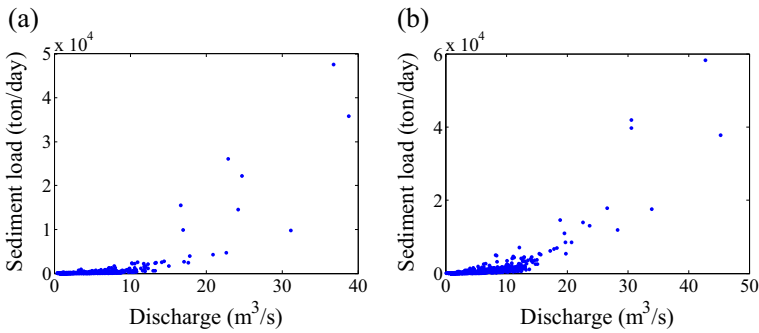
The time series of daily river flow and suspended sediment data used in this study belong to two stations on the Muddy Creek near Vaughn in Montana, USA. The drainage areas at these sites are  $730.377 \text{ km}^2$  for the upstream station (station No. 06088300) and  $813.256 \text{ km}^2$  for the downstream station (station No. 06088500), as shown in Fig. 2. These two stations have been studied in several works (Browning et al. 2005; Kisi Ö and Fedakar 2014), which ensures the reliability of our collected data. The objective of this work is to estimate the suspended sediment loads at the downstream station ( $SSL_d$ ) based on river flows either at the upstream or downstream station ( $Q_u$  or  $Q_d$ ) and sediment loads at the upstream station ( $SSL_u$ ).

The daily dataset was collated from US Geological Survey (USGS), covering a time period of 4 years from 1st January 1977 to 31st December 1980. The discharge and sediment data for the upstream and downstream stations are plotted in Fig. 3. It can be seen that there is a highly nonlinear relationship between discharge and



**Fig. 2** Locations of stations on the Muddy Creek near Vaughn in Montana, USA





**Fig. 3** Scatter plots of (a) upstream and (b) downstream data between sediment load and discharge

sediment data for both stations. The presence of outliers is detected as well, particularly for the sediment data. In the downstream dataset, only four values above or near 40,000 ton/day are observed while the others are below 20,000 ton/day. These outliers of data may give difficulty to the estimation models.

For the purpose of calibration and estimation, data of years 1977 and 1978 are chosen in the training period, whilst those of year 1980 are chosen in the testing period. The remaining data of year 1979 (around 25 % of the whole data) are used for validation, which is an indispensable process to avoid over-fitting. The statistical parameters of river flow and sediment data for the two stations are summarized in Table 1, in which  $X_{mean}$ ,  $X_{median}$ ,  $S_X$ ,  $X_{max}$  and  $X_{min}$  denote the mean, median, standard deviation, maximum and minimum, respectively. A noticeable difference between  $X_{mean}$  and  $X_{median}$  is detected for the sediment data, which provides supporting evidence for the existence of outliers. The high values of  $S_X$  indicate the complexity of the sediment data, and this may have a negative effect on the estimation

**Table 1** Summary of statistical parameters for data at two stations

Station no.	Data type	Data period	$X_{mean}$	$X_{median}$	$S_X$	$X_{max}$	$X_{min}$
06088300	$Q_u$ (m <sup>3</sup> /s)	Training	2.97	1.53	3.53	38.79	0.62
		Validation	3.57	1.81	3.57	31.15	0.71
		Testing	2.65	1.84	2.32	22.94	0.42
	$SSL_u$ (ton/day)	Training	291.91	10.0	2450.3	47,600	2.30
		Validation	175.84	13.0	666.67	9740	2.00
		Testing	175.55	13.5	1589.9	26,100	1.40
06088500	$Q_d$ (m <sup>3</sup> /s)	Training	3.79	1.93	4.33	45.31	0.14
		Validation	4.59	2.21	4.56	33.98	0.82
		Testing	3.64	2.07	3.34	30.58	0.45
	$SSL_d$ (ton/day)	Training	610.28	25	3219.6	58,300	0.96
		Validation	539.36	58	1413.7	17,500	2.80
		Testing	384.04	30	2351.4	41,900	1.70

performance. Besides, the  $X_{min}$  value in the training set is higher than that in the corresponding testing set, both for  $Q_u$  and  $SSL_u$ . This may cause extrapolation difficulties in estimation of low sediment values. In short, the sediment data to be estimated are irregular and ambiguous, and a model, which can fit the highly nonlinear relationship between  $SSL_d$  and the inputs, is in urgent need.

## 4 Results and Discussion

In order to undertake the comparison of performances by different inputs and models, two evaluation criteria are employed in the present study, i.e. root mean square error (RMSE) and Nash-Sutcliffe efficiency coefficient (NSEC). They are determined by the following equations:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - \hat{Y}_i)^2} \quad (12)$$

$$\text{NSEC} = 1 - \frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^N (Y_i - \bar{Y})^2} \quad (13)$$

where  $Y_i$  and  $\hat{Y}_i$  are respectively observed and computed values ( $i=1,2,\dots, N$ ),  $\bar{Y}$  is the averaged observed data, and  $N$  is the number of observations. The RMSE represents the [sample standard deviation](#) of the differences between computed and observed values. The NSEC exhibits the relative magnitude of the residual variance compared to the observed data variance. Both of them are used to assess the predictive power of [hydrological](#) models. The RMSE could also be employed as the error function in the calibration period, while NSEC is not capable due to its sensitiveness to extreme values with large outliers. However, the NSEC statistics is effective when evaluating the performances of models with different sets of data since it is a relative criterion. As can be seen from the above two equations, lower value of RMSE and higher value of NSEC indicate a better performance of estimation model.

### 4.1 Selection of Input Vectors

Six input combinations are evaluated to estimate current downstream sediment load value  $SSL_{d(t)}$ . In all cases,  $SSL_{u(t-1)}$  indicating a one-step ahead is adopted as the only sediment input. Since the travel time of flow from upstream to downstream river section is within 1 day,  $Q_{u(t)}$  and  $Q_{d(t)}$  for the current day as well as  $Q_{u(t-1)}$  and  $Q_{d(t-1)}$  for the previous one day are chosen to constitute the inputs. Table 2 provides the estimation performances of DPFNN and HDFNN models by six input combinations in the testing period. It is noted that the DPFNN model is relatively insensitive to inputs since the RMSE and NSEC values by different input combinations are comparable. The configuration of  $[Q_{u(t)}, Q_{d(t)}, Q_{d(t-1)}, SSL_{u(t-1)}]$  is the most valid input for DPFNN model, yielding the smallest RMSE and highest NSEC values. HDFNN model has the best accuracy with input  $[Q_{u(t-1)}, Q_{u(t)}, Q_{d(t)}, Q_{d(t-1)}, SSL_{u(t-1)}]$ , where there is a

**Table 2** Performances by various input combinations for DPFNN and HDFNN models

Input vector	DPFNN		HDFNN	
	RMSE (ton/day)	NSEC	RMSE (ton/day)	NSEC
$[Q_{u(t-1)}, Q_{d(t)}, SSL_{u(t-1)}]$	840.1555	0.8723	564.0116	0.9425
$[Q_{u(t)}, Q_{d(t)}, SSL_{u(t-1)}]$	835.7967	0.8737	481.9468	0.9580
$[Q_{u(t-1)}, Q_{u(t)}, Q_{d(t)}, SSL_{u(t-1)}]$	837.7312	0.8731	385.2324	0.9732
$[Q_{u(t-1)}, Q_{d(t)}, Q_{d(t-1)}, SSL_{u(t-1)}]$	817.2461	0.8792	485.8397	0.9573
$[Q_{u(t)}, Q_{d(t)}, Q_{d(t-1)}, SSL_{u(t-1)}]$	813.1171	0.8804	394.6522	0.9718
$[Q_{u(t-1)}, Q_{u(t)}, Q_{d(t)}, Q_{d(t-1)}, SSL_{u(t-1)}]$	814.1384	0.8801	362.8466	0.9762

35.67 % reduction in RMSE and 3.58 % improvement in NSEC when compared with the case of input  $[Q_{u(t-1)}, Q_{d(t)}, SSL_{u(t-1)}]$ . In general, the estimation models would attain better performance when more effective information is provided by the inputs. This explains the worse accuracy with input  $[Q_{u(t-1)}, Q_{d(t)}, SSL_{u(t-1)}]$  for both DPFNN and HDFNN models. In addition,  $Q_{u(t-1)}$  is a valid input variable in the HDFNN model since the potential discharge inputs are more likely to perform their efficiency when considered alone by combining fuzzy pattern-recognition. The above results also verify the importance of the upstream sediment and discharges to the downstream sediment in this particular study site.

## 4.2 Model Performances

To draw an effective comparison between four estimation models, performances in the training and testing periods with two input combinations  $[Q_{u(t-1)}, Q_{d(t)}, SSL_{u(t-1)}]$  and  $[Q_{u(t-1)}, Q_{u(t)}, Q_{d(t)}, Q_{d(t-1)}, SSL_{u(t-1)}]$  are presented in Table 3. The configurations of neural network for each model are provided as well. Take the cases for  $[Q_{u(t-1)}, Q_{d(t)}, SSL_{u(t-1)}]$  as an example, (3,6,1) for MFNN model implies that there are 3 nodes in the input layer, 6 nodes in the hidden layer and 1 node in the output layer. For the DPFNN model, the structures for multi-layer and two-layer are respectively (3,5,1) and (3,1). HFNN model only considers the discharge inputs, thus input nodes are  $Q_{u(t-1)}$  and  $Q_{d(t)}$ . The inputs in HDFNN model for multi-layer and two-layer are different: one contains discharge data ( $Q_{u(t-1)}$  and  $Q_{d(t)}$ ) and the other one contains sediment data  $SSL_{u(t-1)}$ .

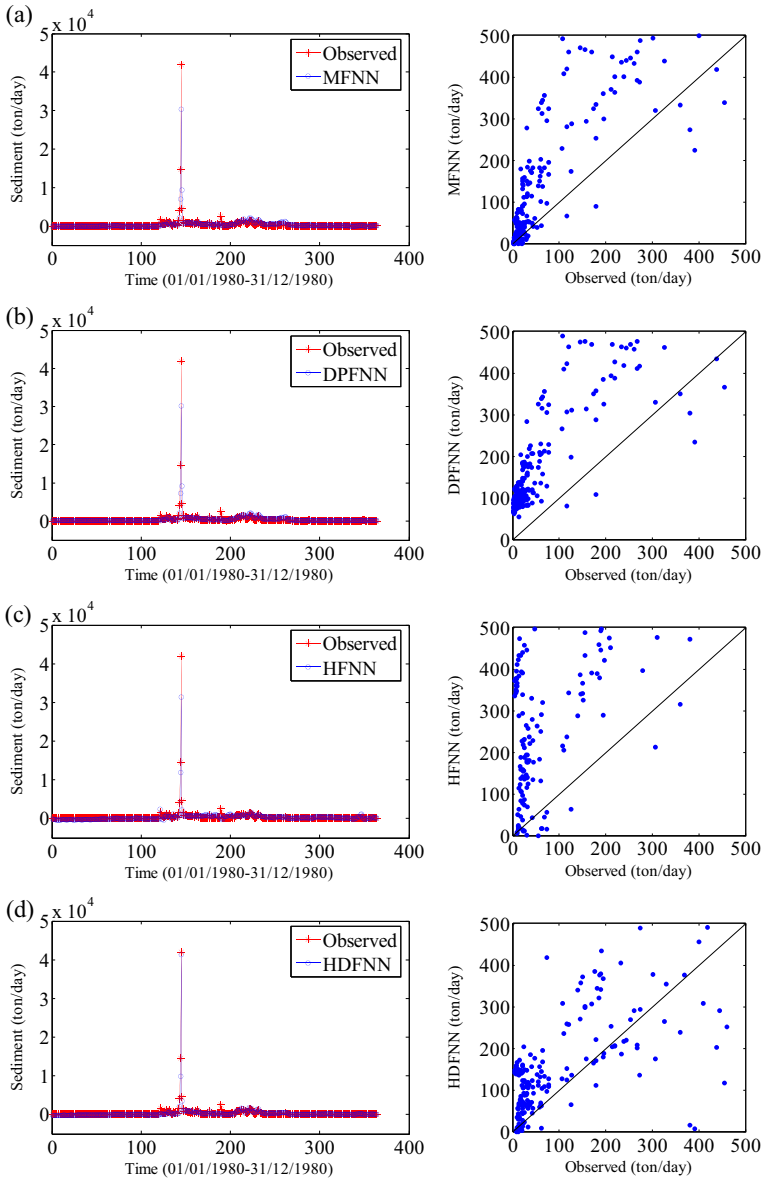
It can be found in Table 3 for the case of  $[Q_{u(t-1)}, Q_{d(t)}, SSL_{u(t-1)}]$  that DPFNN model demonstrates better generalization capability and estimation ability than MFNN model, as indicated by RMSE and NSEC values in both training and testing periods. This may attribute to its capacity of mapping both nonlinear and linear relationship with the double parallel networks. HFNN model is found to be superior to MFNN model as well, because it is capable of providing information about different patterns. However, the influence of upstream sediment loads is not included, which results in deficient estimation. HDFNN model draws the advantages of both DPFNN and HFNN models, hence, gives the best performances amongst the four models. When compared with MFNN model, there is a 49.47 % and 42.58 % reduction in RMSE value for the training and testing stages, respectively. The superiority of HDFNN over DPFNN and HFNN models is apparent, particularly for the NSEC values in the testing period.

**Table 3** Performances by various models with input  $[Q_{u(t-1)}, Q_{d(t)}, SSL_{u(t-1)}]$  and  $[Q_{u(t-1)}, Q_{u(t)}, Q_{d(t)}, Q_{d(t-1)}, SSL_{u(t-1)}]$

Input	Model	Configuration	Training		Testing	
			RMSE (ton/day)	NSEC	RMSE (ton/day)	NSEC
$[Q_{u(t-1)}, Q_{d(t)}, SSL_{u(t-1)}]$	MFNN	(3,6,1)	773.6575	0.9480	982.1818	0.8255
	DPFNN	(3,5,1) and (3,1)	681.4953	0.9551	840.1555	0.8723
	HFNN	(2,5,1)	547.4661	0.9710	813.0243	0.8804
	HDFNN	(2,7,1) and (1,1)	390.9476	0.9852	564.0116	0.9425
$[Q_{u(t-1)}, Q_{u(t)}, Q_{d(t)}, Q_{d(t-1)}, SSL_{u(t-1)}]$	MFNN	(5,4,1)	665.8922	0.9572	815.4227	0.8797
	DPFNN	(5,5,1) and (5,1)	692.0350	0.9537	814.1384	0.8801
	HFNN	(4,5,1)	402.7692	0.9843	697.1047	0.9121
	HDFNN	(4,7,1) and (1,1)	284.9112	0.9922	362.8466	0.9762

The above conclusions can be strengthened by results in Table 3 and Fig. 4 with input combination  $[Q_{u(t-1)}, Q_{u(t)}, Q_{d(t)}, Q_{d(t-1)}, SSL_{u(t-1)}]$ . MFNN model has a fairly high RMSE value (815.4227 ton/day), and shows inferior results due to its total “black-box” operation. The RMSE value of HDFNN model is respectively 55.50, 55.43 and 48.24 % lower than that of MFNN, DPFNN and HFNN model in the testing period. Meanwhile, the NSEC value attained by HDFNN model is 0.9762, which is superbly high to reveal the ability of HDFNN model for sediment estimation. It can also be observed that the improvement of DPFNN over MFNN model is not significant with input  $[Q_{u(t-1)}, Q_{u(t)}, Q_{d(t)}, Q_{d(t-1)}, SSL_{u(t-1)}]$ . A possible explanation may be that the two-layer neural network in DPFNN is redundant and invalid when more input variables are included. The two-layer neural network in HDFNN model is not a “parallel” one as the DPFNN model, since its input is different from the multi-layer neural network. Therefore, the HDFNN model is still able to yield accurate results and capture the effective inputs in this study case.

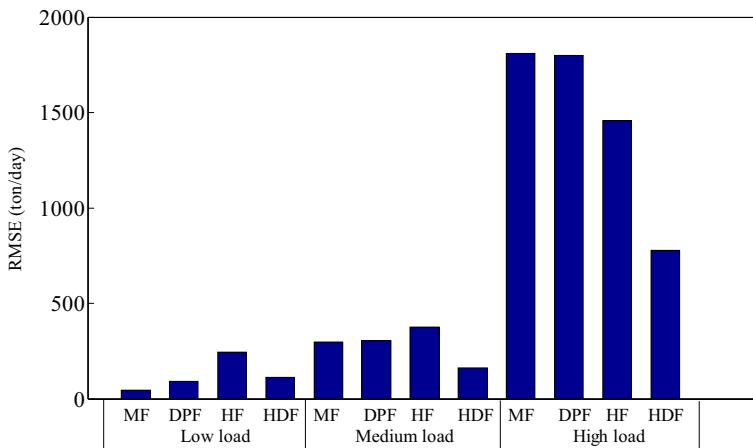
The time series of observed and computed *SSL* as well as the scatter plots by various models are demonstrated in the left and right hand side of Fig. 4, respectively. It is observed that the high values estimated by the HDFNN model are closer to the observed values than the three benchmarking models. In particular, it perfectly fits the peak sediment load as exhibited in Fig. 4d, while the other three models underestimate the peak value. The data applied to scatter plots are below 500 ton/day, which take up around 83 % of all data and are used to present the performances of relatively medium and low values. As seen from the figures, HDFNN model estimates are less scattered in relation to the other three models. MFNN and HFNN over-estimate most of the values, whilst DPFNN could not model the observations lower than 50 ton/day. It can be concluded that the HDFNN model is more adequate than the others for *SSL* estimation since it can simulate the sediments characterized by fuzziness, nonlinearity and time variety.



**Fig. 4** The observed and estimated suspended sediments by (a) MFNN (b) DPFNN (c) HFNN (d) HDFNN models in the testing period with input  $[Q_{u(t-1)}, Q_{u(t)}, Q_d(t), Q_d(t-1), SSL_{u(t-1)}]$

### 4.3 Partitioning Analysis for Low, Medium and High Loads

The statistics RMSE and NSEC scale the mean squared error of estimation models, therefore particularly reflect the performance on high values. Thus the above discussions on evaluation criteria and plots of estimated data could not provide explicit performances on different intervals of values. To address this problem, partitioning analysis is undertaken with regard



**Fig. 5** RMSE statistics of MFNN, DPFNN, HFNN and HDFNN models of low, medium and high values with input  $[Q_{u(t-1)}, Q_{u(t)}, Q_{d(t)}, Q_{d(t-1)}, SSL_{u(t-1)}]$

to the performances of four models in this study. It is performed by finding threshold values of dataset and partitioning the data into several intervals (Goyal 2014). Threshold values are determined based on the observed *SSL* data in the testing period. Median and mean of the dataset are considered as two threshold values. Values lower than median (30 ton/day) are regarded as a “low load”; values higher than mean (384.04 ton/day) as a “high load”; values higher than median and lower than mean as a “medium load”.

The RMSE statistics of four models with respect to low, medium and high loads are illustrated in Fig. 5. For the ‘low load’, HFNN model performs worse than its counterparts with a largest RMSE value due to the irrespective of upstream sediment data. HDFNN model is completely adequate in estimating low *SSL* values. Four models are able to mimic the ‘medium load’ with comparable performances, in which HDFNN model achieves the best result. The RMSE obtained by HDFNN model is much smaller than the other three for the ‘high load’. In overall, the performances of HDFNN model on low, medium and high loads are consistently excellent, which corroborates the use of this approach in *SSL* estimation. This can mainly be attributed to the introduction of continuity equation which reveals the time-varying characteristic of sediment loads.

## 5 Conclusions

This study is concerned with the application of HDFNN model for suspended sediment load estimation. The fuzzy and time-varying characteristics of the sediment data are considered in HDFNN model, while a structure of double neural networks is employed with respect to river discharge and sediment inputs. The estimates based on HDFNN models are compared with three models (i.e. MFNN, DPFNN and HFNN) with dataset from two stations on the Muddy Creek in Montana, USA. Results confirm the generalization and estimation ability of HDFNN model with the lowest RMSE and highest NSEC values. The high and peak observed values are estimated successfully by HDFNN model as well since it could address the highly nonlinear and fuzzy sediment data. Furthermore, it can perfectly fit the low and medium

values, as indicated by partitioning analysis. Conclusions can be drawn that the HDFNN model provides a superior alternative for *SSL* estimation. It overcomes the drawback of physically meaningless representation of MFNN model, enjoys the advantages of DPFNN and HFNN models by double neural networks and fuzzification process, and includes the influence of upstream sediment loads. The present work is the first application of considering the physics embedded with the structure of double neural networks, for modeling suspended sediment series in the downstream river section. This idea for modeling *SSL* can be referred and extended for other hydrological models. Nevertheless, the stability of the proposed HDFNN model has not been verified by uncertainty analysis, which should be performed as a future work. Besides, the HDFNN model only considers the inputs of upstream discharges and sediment, and their corresponding physical mechanism. The influences of other inputs and hydrological process regarding the *SSL* can be explored additionally to improve the estimation model.

**Acknowledgments** This research was supported by Central Research Grant of Hong Kong Polytechnic University (4-ZZAD).

**Compliance with Ethical Standards**

**Conflict of Interest** No conflict of interest

## References

- Afan HA, El-Shafie A, Yaseen ZM et al. (2014) ANN based sediment prediction model utilizing different input scenarios. *Water Resour Manag* 1–15
- Alp M, Cigizoglu HK (2007) Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. *Environ Model Softw* 22:2–13
- Aytek A, Kisi Ö (2008) A genetic programming approach to suspended sediment modelling. *J Hydrol* 351:288–298
- Browning LS, Bauder JW, Hershberger KE, Sessoms H (2005) Irrigation return flow sourcing of sediment and flow augmentation in receiving streams: a case study. *J Soil Water Conserv* 60:134–141
- Chen XY, Chau KW, Busari AO (2015) A comparative study of population-based optimization algorithms for downstream river flow forecasting by a hybrid neural network model. *Eng Appl Artif Intell* 46:258–268
- Cimen M (2008) Estimation of daily suspended sediments using support vector machines. *Hydrol Sci J* 53:656–666
- Cobaner M, Unal B, Kisi Ö (2009) Suspended sediment concentration estimation by an adaptive neuro-fuzzy and neural network approaches using hydro-meteorological data. *J Hydrol* 367:52–61
- Demirci M, Baltaci A (2013) Prediction of suspended sediment in river using fuzzy logic and multilinear regression approaches. *Neural Comput & Applic* 23:145–151
- Goyal MK (2014) Modeling of sediment yield prediction using M5 model tree algorithm and wavelet regression. *Water Resour Manag* 28:1991–2003
- He M (1993) Theory, application and related problems of double parallel feedforward neural networks. Xidian University, Xi'an
- He MY (1994) Double parallel feedforward neural networks with application to simulation study of flight fault inspection. *Acta Aeronautica ET Aeronautica Sinica* 15:877–881
- He MY, Huang R (2005) Feature selection for hyperspectral data classification using double parallel feedforward neural networks. In: *Fuzzy Systems and Knowledge Discovery*. Springer, pp 58–66
- Huang R, He MY (2007) Feature selection using double parallel feedforward neural networks and particle swarm optimization. In: *Evolutionary computation, 2007. CEC 2007. IEEE Congress on. IEEE, 2007*, pp 692–696
- Khan A, Yang J, Wu W (2014) Double parallel feedforward neural network based on extreme learning machine with L 1/2 regularizer. *Neuro Comput* 128:113–118

- Kisi Ö (2004) Multi-layer perceptrons with Levenberg-Marquardt training algorithm for suspended sediment concentration prediction and estimation. *Hydrol Sci J* 49:1025–1040
- Kisi Ö, Fedakar Hİ (2014) Modeling of suspended sediment concentration carried in natural streams using fuzzy genetic approach. In: *Computational intelligence techniques in earth and environmental sciences*. Springer, pp 175–196
- Kisi Ö, Karahan ME, Şen Z (2006) River suspended sediment modelling using a fuzzy logic approach. *Hydrol Process* 20:4351–4362
- Kisi Ö, Haktanir T, Ardiclioglu M, Ozturk O, Yalcin E, Uludag S (2009) Adaptive neuro-fuzzy computing technique for suspended sediment estimation. *Adv Eng Softw* 40:438–444
- Kothyari UC, Tiwari AK, Singh R (1997) Estimation of temporal variation of sediment yield from small catchments through the kinematic method. *J Hydrol* 203:39–57
- Kouassi KL, Kouame KI, Konan KS, Angulo MS, Deme M (2013) Two-dimensional numerical simulation of the hydro-sedimentary phenomena in Lake Taabo, Côte d'Ivoire. *Water Resour Manag* 27:4379–4394
- Lafdani EK, Nia AM, Ahmadi A (2013) Daily suspended sediment load prediction using artificial neural networks and support vector machines. *J Hydrol* 478:50–62
- Li YT, Gu RR (2003) Modeling flow and sediment transport in a river system using an artificial neural network. *Environ Manag* 31:122–134. doi:10.1007/s00267-002-2862-9
- Li XT, Liu HW, Yin MH (2013) Differential evolution for prediction of longitudinal dispersion coefficients in natural streams. *Water Resour Manag* 27:5245–5260
- McBean EA, Al-Nassri S (1988) Uncertainty in suspended sediment transport curves. *J Hydraul Eng* 114:63–74
- Partal T, Cigizoglu HK (2008) Estimation and forecasting of daily suspended sediment data using wavelet-neural networks. *J Hydrol* 358:317–331
- Qiu L, Chen SY, Nie XT (1998) A forecast model of fuzzy recognition neural networks and its application. *Adv Water Sci* 9:258–264
- Rajae T, Mirbagheri SA, Zounemat-Kermani M, Nourani V (2009) Daily suspended sediment concentration simulation using ANN and neuro-fuzzy models. *Sci Total Environ* 407:4916–4927
- Rocca P, Oliveri G, Massa A (2011) Differential evolution as applied to electromagnetics. *Antennas Propag Mag, IEEE* 53:38–49
- Storn R, Price K (1995) Differential evolution—a simple and efficient adaptive scheme for global optimization over continuous spaces. vol 3. International Computer Science Institute, Berkeley
- Tayfur G, Ozdemir S, Singh VP (2003) Fuzzy logic algorithm for runoff-induced sediment transport from bare soil surfaces. *Adv Water Resour* 26:1249–1256
- Thirumalaiah K, Deo MC (1998) River stage forecasting using artificial neural networks. *J Hydraul Eng* 3:26–32
- Wang J, Wu W, Li ZX, Li L (2011) Convergence of gradient method for double parallel feedforward neural network. *Int J Numer Anal Model* 8:484–495
- Yang RF, Ding J, Liu GD (1998) Preliminary study of hydrology-based artificial neural network. *J Hydraul* 8:23–27
- Zhao RH, Chen SY (2008) A hybrid fuzzy and neural network model for hydrological forecasting in ungauged basins. *IAHS-AISH Public* 39–48
- Zhao HQ, Qi JQ, Wand J, Zheng J, Wu W (2010) Concentration estimation of gas mixture using a wavelet-based DPFNN. *Chin J Sensors Act* 23:744–747
- Zhong SS, Ding G (2005) Research on double parallel feedforward process neural networks and its application. *Contrl Decision* 20:764–768