J. Int. Environmental Application & Science, Vol. 10(4): 384-398 (2015)

Data-driven Modelling in River Channel Evolution Research: Review of Artificial Neural Networks Applications

Z. Rozlach*

Faculty of Geography, Taras Shevchenko National University of Kyiv, Kyiv, UKRAINE

Received July 07, 2015; Accepted November 19, 2015

Abstract: Any interaction with river systems requires detailed consideration of channel evolution. The multiplicity of physical processes occurring within catchment and channel-floodplain complex causes complicated processes in river channel. Therefore, it demands reliable and accurate methods in research, which are capable to consider exclusive and non-linear relationships in river system. In the recent years, new approaches, relied on intelligence models of machine learning are proposed. Among them artificial neural networks (ANN) method is presently widely used in the data-driven modelling for non-linear system behaviour. This paper presents a review of artificial neural network models and numerous applications of ANNs in river channel processes research.

Keywords: *artificial neural networks, intelligence models, channel processes, sediment transport, channel erosion.*

Introduction

Harmonious and stable existence of humanity with environmental hydroecosystems is not possible without deep understanding of physical processes of the latter as well as without knowledge of regularities of its changes. Rivers are of the most dynamic constituents of such systems, so the changes are inevitable part of them. River channel processes is much complicated and exacerbated under the human impact. Often, anthropogenic factor becomes dominant, changing natural tendencies of evolution.

The essence of river channel processes is the interrelations between water flow with channel and soil, which underlying it, and also in sediment migration which either appeared in channel from “outside” (tributaries inflow, catchment denudation etc.), or formed as a result of channel deformations (Makkaveev, 1955). That is why main fundamental knowledge of the river channel processes associated with amount and variability of sediment load, its transport mechanism, and in particular sediment runoff is one of the main factors of river channel processes (Chalov, 1979).

Basically, the physical change of any watercourse can be manifested in erosion or deposition of sediments. There are numerous methods to determine and assess such processes however the main purpose of this paper is review of modern approaches in river channel evolution research, in particular modern techniques using artificial neural network models. According to the stated above the structure of this review is the following: in the next Section 2 brief statements of conventional methods of river channel processes assessment are given. Section 3 describes methodology of ANN and conventional ANN structures. The hybrid ANN models will be discussed in Section 4. In Sections 5 and 6 application of neural network models for erosion and sediment assessment are given. Finally, conclusions are presented in the last Section 7.

Conventional models in river channel research

There are a quite huge amount of existing models for estimation of hydromorphological changes, erosion processes and sediment transport. These models have differences in the estimating processes, the number of data necessary for the model calibration and its further application. It is also determines the complexity of them. The choice of a model is guided by the purposes and tasks.

In general, models fall into three main categories, depending on the physical processes simulated by the model, the model algorithms describing these processes and the data dependence of the model (Merritt *et al.*, 2003): empirical or statistical; conceptual; physically-based. Often in research there is a mixture of modules from several categories mentioned above.

Corresponding: zakhar.rozlach@gmail.com; Tel: 38 066 6400330; Fax: 38 044 5213270

Empirical/statistical models are widely used because of their relatively simple structure and mathematical methods involved, and their ability to work with limited data (Zhu *et al*., 2007). They are based primarily on the analysis of observations and seek characterize response from these data (Wheater *et al*., 1993). In many cases conceptual models tend to include a general description of catchment processes without including the specific details of process interaction, which would require detailed catchment information (Sorooshian, 1991). Parameter values for conceptual models traditionally obtained through calibration against observed data (Abbott *et al*., 1986). Due to that, conceptual models tend to suffer from problems associated with the identifiability of their parameter values (Jakeman & Hornberger, 1993). Most of calibration techniques which are used for conceptual models are capable of finding only local optima at best. This type of models is considered as intermediate between empirical and physically-based (Beck, 1987). Physically-based models are based on the solution of fundamental physical equations which describes stream flow and sediment as well as surrounding processes in catchment. Standard equations employed in such models are the equations of conservation of mass and momentum for flow and the equation of conservation of mass for sediment (Bennett, 1974).

For most conventional models based on mathematical framework, the lack of understanding of physical processes is usually reimbursed by either simplifying the problem or incorporating a number of assumptions into the models. That is why many traditional models fail to simulate the complex behavior of river channel processes problems.

Artificial neural networks

Development of artificial neural networks was inspired by functioning of a human brain and its possibilities to process information. Despite the very feeble semblance to biological neural networks, ANN is a very powerful universal tool for solution of many tasks. Among the main valuable properties (Zaentsev, 1999) are:

- Educability. If to choose one of the models of ANN, develop the network and perform the training algorithm, we can teach the network to solve a task.
- Generalization ability. After training, network becomes not sensible to minor changes of input signals (noise) and gives correct result in output.
- Ability to abstraction. If submit a network with several distorted input variables, the network is able to create the perfect image in output, which has never encountered before.

The first fundamental concepts related to neural computing were developed in 1943 (McCulloch and Pitts, 1943). Research raised at the end of 80-s when Rumelhart et al. (1986) introduced the backpropagation (BP) training algorithm for feedforward ANN.

Nowadays ANN is a broad term covering a large variety of network architectures, the most common of which is a multi-layer perceptron (MLP) (Bhattacharya et al., 2005). ANNs consist of a number artificial neurons or nodes. Each neuron is an independent computational unit. Usually these neurons arranged in layers: input layer, output layer and one or more hidden layers (Figure 1). Each layer of processing elements or nodes is fully connected to the proceeding layer by interconnection strengths, or weights.

Network operation However for more than 40 years there were no significant interests in the application of ANN can be described as follows: given a set of input vectors and the associated target (output) vectors. The objective of ANN is to learn a functional relationship between the input vectors and the target vectors. Each target vector z is an unknown function f of the input vector x:

$$z = f\,(x) \qquad (1)$$

The task of the network is to learn the function f. The network includes a set of parameters (weights vector). Such weights are varied to modify the function f' which are computed by the network and should be as close to the desired value as possible. Based on the training data the weight parameters are determined during training (calibration) process (Bhattacharya *et al*., 2005).

ANN can be classified based on three main criteria: *learning type*; *network architecture* and *optimization method*. *Learning or training* is divided into supervised and unsupervised (Masters, 1993). In supervised learning, the network is presented with an observed data set of model inputs and the corresponding (true) outputs. The actual output of the network is compared with the true output and the error is calculated. According to this error connection weights between the model inputs and outputs is adjusted in order to reduce the error between the observed outputs and those predicted by the ANN. During the unsupervised learning no true outputs are given. The network itself adjusts the connection weights according to the input values. The idea of training in unsupervised networks is to cluster the input data into classes of similar features (Shahin et al., 2008).

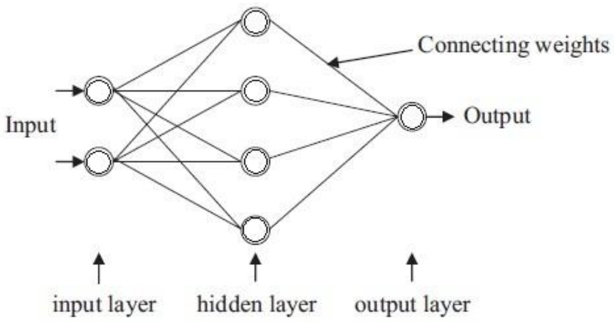

Figure 1. Structure of typical artificial neural network

Network architecture is expressed in connection type and its geometry. Determination the network architecture is one of the most important and at the same time of the most difficult tasks in the model building process. Connection type represents the connections between nodes and can be divided into feedforward and feedback networks. In feedforward networks, the connections between nodes are only in the forward direction. In case of feedback networks, connections between nodes are in both forward and backward directions. The dominant number of existing ANN models in the field of water resources has feedforward type of connections (Maier & Dandy, 2000; ASCE, 2000a,b; Shahin *et al.*, 2008; Rezapour *et al.*, 2010).

As for the network geometry, it determines the number of nodes in the input layer, one or several hidden layers with a number of nodes each one and the number of nodes in the output layer. There is no unified approach for determination of optimal neural network geometry. It is generally achieved by fixing the number of layers and choosing the number of nodes in every layer. It has been shown that one hidden layer is sufficient to approximate any continuous function provided that sufficient connection weights are given (Hornik *et al.*, 1989). However often in practice, many functions are difficult to approximate using a model with only one hidden layer (Flood & Kartam, 1994). Some researches stated that use of more than one hidden layer provides greater flexibility and allows to approximate complex functions with fever connection weights in many situations (Flood & Kartam, 1994; Sarle, 1994; Ripley, 1994).

The number of nodes in the input and output layers are restricted by the number of model inputs and outputs respectively. Concerning the number of hidden nodes, there is no distinct rule. The proper quantity usually determined based on a trial-and-error procedure. Generally, the more hidden nodes and layers, the more capabilities has the network, more non-linear relation can be between input-output, and the more slowly it trained and process. Nevertheless, if the nodes are overmuch, the network can be overfitted and has poor generalization ability. Several approaches to find an optimum solution in determining of a number of hidden nodes, such as pruning algorithms (Karnin, 1990; Sietsma & Dow, 1991; Setiono, 1997) or constructive algorithms (Hirose *et al.*, 1991; Chen *et al.*, 1997) are proposed.

Optimization method defines the specificity of a training process - how to choose the correct values of network parameters (connection weights). The most commonly used for feedforward MLP neural network is the backpropagation algorithm, which is based on the method of steepest descent (Rumelhart *et al*., 1986). The goal of any training algorithm is to minimize the global error '*E*'. The BP algorithm calculates '*E*' and distributes it backward from the output to the hidden and then to the input neurons. Using the steepest gradient principle, where the change in weight is directed towards the negative of the error gradient, does this.

Another training scheme which is not as popular as BP however is able to ensure more optimum network configuration (ASCE, 2000a) is cascade correlation (CC). At the training process with such algorithm the configuration of the network is not fixed. Hidden nodes are added one by one until the training termination criterion is reached. Convergence of an error is not achieved by transmitting it backward, but by maximizing the effect or correlation of the new hidden node’s output on the residual error.

Beside these training algorithms a number of global optimization methods, such as genetic algorithms (Goldberg, 1989) and simulated annealing (Kirkpatrick *et al.,* 1983) were proposed and their implementation is becoming more frequent in water resources research. The advantage of these methods is that they are able to escape local minima in the error surface and, thus, able to find optimal or near optimal solutions. However, they also have a slow convergence rate. Ultimately, the model performance criteria, which are problem specific, will dictate which training algorithm is the most appropriate (Shahin *et al.,* 2008). More detailed theoretical explanation of ANN can be found in (Wasserman, 1989; Zurada, 1992; Haykin, 1999).

Having significant capabilities to overcome the problem of exclusive and the non-linear relationships ANNs are being widely applied in varies scientific areas, especially in applications involving precise estimation and forecasting. A lot of works in rivers research showed the best performance of ANNs in comparison to the conventional models and approaches. That is also proved by the increasing of a number of scientific publications for last two decades.

Despite the good and accurate performance of ANNs, they have some weaknesses, the main of which are extrapolation ability and knowledge extraction. It is generally accepted that ANNs perform best when they do not extrapolate beyond the ranges of the data for which it was trained (Flood and Kartam, 1994; Tokar & Johnson, 1999). In many fields of hydrology and river morphology extreme values prediction is of high concern, such as flood forecasting, channel deformation assessment etc. That shortcoming of neural networks is needed to be solved and some propositions have already been done (Sudheer *et al.*, 2003).

Because of a limited ability of knowledge extraction ANNs are often considered as a “black box” models due to their lack of transparency as they do not consider nor explain the underlying physical process (Shahin et al., 2008). In contrast, a number of investigations were done in hydrologic engineering (Jain *et al.*, 2004; Olden *et al.*, 2004; Sudheer, 2005) which stated that the careful examination of the distributed information contained in the trained ANN can inform about the nature of the physical processes in a watershed captured by various components of the ANN model.

Hybrid ANN models

Despite popularity in application of ANN models in various directions in civil and environmental engineering, hydroecology, river morphology due to its essential advantages over conventional models, ANN has some shortcomings which sometimes deprives of its benefits. Researchers are trying to improve such models combining with different methods and algorithms, creating hybrid types of ANN. Several types of improved neural networks models are worth of attention.

Adaptive neuro-fuzzy inference system (ANFIS). ANFIS is a combination of an adaptive ANN and a fuzzy inference system (FIS). It was first introduced by Jang (1993). A basic ANFIS structure is illustrated below (Figure 2).

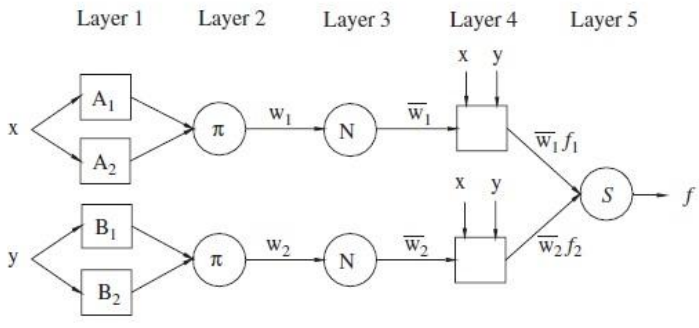

Figure 2. ANFIS network architecture (Bateni *et al.*, 2007)

The parameters of the FIS are determined by the neural network learning algorithms. During the learning procedure the set of “if-then” rules with appropriate membership functions (MF) is constructed from the specified input-output pairs. ANFIS identifies a set of parameters through using of combination of two learning methods: gradient descent and least squared error method (Agil *et al.*, 2007). Gradient descent method aims to search in a stepwise fashion for the best values of the estimate. At each step it uses a linear approximation of the function and refines this approximation by successive corrections. The method of least squares assumes that the best-fit curve of a given type is the curve that has the minimal sum of the deviations squared (least square error) from a given set of data.

There are two main approaches for fuzzy inference systems: approaches of Mamdani (Mamdani & Assilian, 1975) and Sugeno (Takagi & Sugeno, 1985). Sugeno's system is more compact and computationally efficient. In first-order Sugeno's system, a typical rule set with two fuzzy "if-then" rules can be expressed as follows:

Rule 1 : If x is $A1$ and y is $B1$, then $f_1 = p_1x + q_1y + r_1$ (2)

Rule 2 : If x is $A2$ and y is $B2$, then $f_2 = p_2x + q_2y + r_2$ (3)

where $A1$ $A2$ and $B1$, $B2$ are the MF for inputs x and y, respectively, p_1, q_1, r_1 and p_2, q_2, r_2 are the parameters of the output function. For example, if the bell-shaped MF is employed, μ_{Ai} is given by:

$$\mu_{Ai}\,(\mathrm{x}) = \frac{1}{1 + \{\left(\frac{\mathrm{x} - \mathrm{c_i}}{\mathrm{a_i}}\right)^{2}\}^{\mathrm{b^i}}} \tag{4}$$

where a_i, b_i and c_i are the parameters of the MF, governing the bell-shaped function accordingly.

There are two adaptive layers in such ANFIS architecture (Fig. 2) – the first and the fourth layers. In the first layer three modifiable parameters $\{a_i, b_i, c_i\}$, known as premise parameters, are settled. They related to the input MFs. In the fourth layer there are also three modifiable parameters $\{p_i, q_i, r_i\}$, which are called consequent parameters. The task of the learning procedure is to tune all the modifiable parameters (premise and consequent), to make the ANFIS output match the training data. In the implementation of fuzzy logic, several types of membership functions can be used. More detailed description of ANFIS theory can be found in related literature (Jang, 1993; Jang *et. al.*, 1997).

The ANFIS is a universal approximator and as such is capable of approximating any real continuous function on a compact set to any degree of accuracy (Jang *et al.*, 1997). The difference between the common neural network and the ANFIS is that, while the former captures the underlying dependency in the form of the trained connection weights, the latter does so by establishing the fuzzy language rules (Azamathulla *et al.*, 2009). The fuzzy model is put in the framework of adaptive system, which helps to facilitate learning and adaptation. Such framework makes the ANFIS modeling more systematic and less reliant on expert knowledge. Moreover, neuro-fuzzy models in combination with ANN are able to facilitate the transparency of the latter which is important in application at fields related to earth sciences and water resources in particular.

Recently, fuzzy logic has been successfully used in river channel research. The results showed that ANFIS can be used as good alternative to traditional methods and often shows a better performance than conventional ANN models. There are still not a wide range of works in application of ANFIS in this field (Kisi, 2003, 2005; Bateni *et al.*, 2007; Lohani *et al.*, 2007; Shu & Quadra, 2008; Kisi *et al.*, 2009; Azamathulla *et al.*, 2009; Cobaner *et al.*, 2009; Rajaee *et al.*, 2009, Kisi & Shiri, 2012). Considerable researches of river sedimentation using a fuzzy logic were done by O. Kisi. Starting this his PhD research in modeling of suspended sediment yield in a river cross-sections (Kisi, 2003), he continued implementation of neuro-fuzzy networks in estimation and prediction of suspended sediment concentration. Kisi (2005) investigated the abilities of neuro-fuzzy and ANN approaches to model the daily stream flow - suspended sediment load. He compared the performance of these models with conventional - sediment rating curve and multiple linear regression. Results indicated that the neuro-fuzzy model gave better estimation than the other techniques. Kisi *et al.* (2009) investigated the accuracy of a neuro-fuzzy technique in estimation of daily suspended sediment of rivers in Turkey and compared the results with different ANN models. Kisi and Shiri (2012) implemented several soft computing techniques, among which ANFIS, ANNs and genetic programming (GP) to estimate daily suspended sediment concentration on river by using hydro-meteorological data. Bateni et al. (2007) proposed alternative approaches for estimation of time-dependent scour depth around piers. The performance of ANFIS and ANN models were compared with existing methods. The results showed a much more accurate estimation made by intelligent models (ANFIS and ANN). A usage of fuzzy logic for deriving stage-discharge-sediment concentration relationships was done by Lohani *et al*. (2007). Shu and Quadra (2008) applied ANFIS for analyzing regional flood frequency at un-gauged sites. The prediction of bed-load in rivers was also successfully done using ANFIS (Azamathulla *et al*., 2009). They recommended ANFIS model for computing bed-load transport rates, which provided much closer agreement with the measured values in comparison to existing equations. Cobaner *et al.* (2009) proposed an adaptive neuro-fuzzy approach to estimate suspended sediment concentration on rivers by climatic parameter (rainfall), stream flow and suspended sediment concentration data. The comparison results reveal that neuro-fuzzy models performed better than other models, among which tree types of ANN models and two different sediment rating curves (SRC). To the same conclusions came Rajaee *et al*. (2009) who examined neuro-fuzzy, ANN, multi linear regression (MLR) and SRC models. Comparison of the model's results indicated that the neuro-fuzzy model had more ability in predicting suspended sediment concentration in rivers.

Artificial neural networks and genetic programming (ANN-GP). Genetic programming is an evolutionary algorithm based on the concepts of natural selection and genetics. It was introduced by Koza (1992). GP algorithms firstly define an objective function as a qualitative criterion. Further, this function is used for measurements and evaluation of different solutions in a step by step manner of structural correction until suitable solution is found. The main advantages of GP are that it can be applied to area where interrelationships among relevant variables are poorly understood; theoretical analysis is constrained by assumptions and therefore their solutions are of limited use; and when there is a large amount of data in computer readable forms requiring tedious processing (Banzhaf *et al*., 1998).

GP is a widely used machine learning technique. It uses a tree-like structure, as decision trees, to represent its concepts and its interpreter as a computer program. Rao (1996) found out that GP well suited to finding global optimum with high probability. The strong point also of GP technique among others is that it can produce explicit formulations – model expressions - of the relationship that rules the physical phenomenon. Based on such expressions it is possible to interpret the inherent physical essence of the observed processes. This enables a researches to evolve new interesting formulae, which can be useful in future studies. Besides, comprehensibility of GP models is also a way to reduce the risk of over-fitting in ANNs training process and improve generalization of resulting models.

Successful combination of genetic programming and ANN can be found in several scientific works dedicated to river channel research (Minns, 2000; Kamp & Savenije, 2006; Guven & Gunal, 2008; Elshorbagy et al., 2010a,b; Najafzadeh & Barani, 2011; Kisi *et al.*, 2012a; Kisi & Shiri, 2012). GP technique in these works was mostly used for finding an optimal architecture of neural networks, in particular the best input variables to the model. Minns (2000) found that genetic programming may be used as a data-mining technique to discover usable relationships in measured or experimental data. He proved that superior performance of the ANN paradigm over traditional methods of data mining and analysis can also be claimed by GP, as it can supply a symbolic-algebraic relation between the measured data through a process of evolution and competition between all possible solution expressions. He stated for the hybrid approach, which uses the best characteristics of both ANNs and GPs, and can provide even more accurate solutions than just one approach on its own (Minns, 2000). Kamp and Savenije (2006) used GP to improve the results of the ANN by optimizing the original training set. Based on the optimal training set that was found by GP the authors trained the flow model again and improved the accuracy measured in root mean square error (RMSE) up to 39 %. Guven and Gunal (2008) used GP to choose an optimal structure for the proposed neural network in prediction of scour downstream of grade-control structures. Experimental investigation of the predictive capabilities of six different data driven modeling techniques in hydrology was done by Elshorbagy *et al*. (2010a,b). They implemented ANNs, GP, evolutionary polynomial regression (EPR), support vector machines (SVM), M5 model trees (M5), K-nearest neighbors (K-nn) and multiple linear regression (MLR) and found out that GP was the most successful technique due to its ability to adapt the model complexity to the modeled data (Elshorbagy et al., 2010b). In their study Najafzadeh and Barani (2011) developed group method of data handling (GMDH) network using GP to predict the pier scour depth in bridges. It was found that although GMDH-GP model is very time-consuming and more complicated, this method provided a better prediction than GMDH model developed by typical for ANN back propagation algorithm. Application of GP gave positive results in river suspended sediment estimation (Kisi *et al*., 2012a; Kisi & Shiri, 2012). It was found that models based on GP performed better than ANNs, ANFIS, SVM and conventional methods.

Wavelet artificial neural network (WANN). Wavelet transform of signals is a generalization of spectral analysis and is more effective method than Fourier transforms (Astafyeva, 1996). The basis of wavelet transform in general is implementation of two continuous, interdependent and integral in independent variable functions:

- Wavelet-function $\psi(t)$, as psi-function of time with zero integral value, and frequency Fourier image $\Psi(\omega)$. This function, which is usually called as wavelet, is used to distinguish local signal peculiarities. As a wavelet the functions which are well localized in time- and frequency range are usually selected.
- Scaling function φ (t), as a phi-function with integral unit value, which performs signal rough approximation.

Wavelet transform is a tool that cuts up data or functions or operators into different frequency components, and then studies each component with a resolution matched to its scale (Daubechies, 1992). Figuratively speaking, it is possible to see both the forest and the trees. It is successful technique to detect characteristics of target time series and to detect localized phenomena in non-stationary time series. This method is a powerful signal processing tool used in a time series analysis. As wavelet analysis provides information in both the time and frequency domains of the signal, it gives significant insight into the physical form of the data.

From recently, there has been a growing interest in combination of ANN and wavelet analysis, as it was proved that wavelet transform has a positive influence on neural network models. Such combined model enables very accurate simulation and forecasting due to data representation at many different periods by wavelet transform (Wang & Ding, 2003). In some publications hybrid WAAN models in river research are proposed (Anctil & Tape, 2004; Cannas *et al*., 2005; Partal & Cigizoglu, 2008; Rajaee, 2011). Anctil and Tape (2004) used a WANN model for rainfall-runoff forecasting. Using wavelet transform the time series was decomposed into three sub-series: short, intermediate and long wavelet periods. After that neural network models were trained for each separate sub-series and later forecasted decomposed signals were reconstructed into the original time series. Cannas et al. (2005) proposed a hybrid WANN model for monthly rainfall-runoff modeling in rivers of Italy. Partal and Cigizoglu (2008) developed WANN model for estimation and forecasting of daily suspended sediment load in rivers. The compared results with conventional ANN model and SRC method showed WANN model prediction is significantly superior. The peak sediment values and the cumulative sediment sum are closely approximated with the wavelet-ANN method. Another successful combination of wavelet analysis and ANN for sediment processes research in rivers was implemented by Rajaee (2011). He used a WANN model for daily suspended sediment load prediction. In proposed model the water discharge (Q) and suspended sediment load (SSL) signals were firstly decomposed into sub-signals with different scales in order to obtain temporal properties of the input time series. The decomposed Q and SSL time series were entered to the ANN model for prediction of SSL in one day ahead. This provided the improvement of conventional ANN method up on 46 %. The results of WANN model were also compared with the results of MLR and SRC models and it indicated that WANN model was able to predict SSL much more accurately. Furthermore, the results showed that proposed hybrid model could satisfactory simulate hysteresis phenomenon, estimate cumulative SSL and predict high SSL values.

Artificial neural network with artificial bee colony algorithm (ANN-ABC): Another not only innovative and interesting but also effective example in river channel research of ANN development employed by Kisi *et al*. (2012b). They proposed hybrid model – combination of neural network approach and artificial bee colony algorithm. Such algorithm belongs to a class of methods which are called ‘swarm intelligence’ and combine biological and social heuristic. The swarm intelligence based on positive and negative feedbacks, adaptations to environment and decentralized relations between individuals. At a computational point of view swarm algorithms are stochastic search methods, in which effective combinations of new solution search and improvement of existing solutions are enclosed. Mentioned features helps to avoid earlier convergence in a local decision and find global optimum. The idea of such algorithms based on the statement, that each individual are in the space of searching of possible task solution. Swarm behavior characterized by the following features: autonomy, distribution functions and self-organization (Karaboga, 2005, Alatas, 2010).

Algorithm of artificial bee colony (ABC) was developed by Karaboga (2005). It is an optimization algorithm which mimics the foraging behavior of honey bees. In the ABC algorithm each food source position corresponds to a possible solution to the problem optimization and the solutions are improved in the search cycle to find the global optimum. The algorithm is initialized by the random food source population and a cycle which consists of employed, onlookers and scout bees' phases is iterated until resulting criterion is satisfied.

The ABC algorithm was used in training for calculating weights of the ANN model for suspended sediment estimation in rivers (Kisi *et al*., 2012b). In comparison to traditional back-propagation algorithm which is usually employed in ANN, the ABC algorithm can work with non-differentiable and discontinuous functions (Rojas, 1996). It is able to escape from local minima of multimodal surfaces due to its balanced exploration and exploitation capability. It was found that ANN-ABC model was able to produce better results than ANN, neural differential evolution (NDE), NF and SRC models in estimation of river suspended sediment.

The examples of hybrid models revealed in this Section have demonstrated that combinations of techniques may be very useful and applicable. In many cases such combined techniques improved the models accuracy in estimation and forecast.

Application of ANN models for analysis and estimation of channel erosion

Inevitable part of evolution of any natural river channel is its morphological changes. Such reshaping of a channel is in general a reaction to flow changes. The influence of a man-made structure often may lead to significant scours in a non-distinctive river reaches. Accurate estimation and prediction of scour caused by any hydraulic engineering work is a critical and very important in many field of environmental and civil engineering.

Scour affected by a number of physical parameters such as parameters of fluid, flow, sediment and structure geometry, as well as by a wide variety of non-physical variables like turbulent boundary layer and sediment transport mechanism. This makes the assessment of this phenomenon very complex.

In general, three methods often used in estimation of a local scour: theoretical, experimental and numerical (Lee at al., 2007). Theoretical model of the scour depth often make a lot of assumptions to establish the related mathematical relationship. Besides that the parameters in the function have to be decided by the experimental data. Experimental models based on a numerous experimental studies which combine dimensional analysis with the experimental test of the model in the past. Such experimental relationships are often inadequate due to the large number of parameters affecting the scour. A number of numerical models also have been developed for scour estimation. They are based on flow physics and often fall in a specialist's area, but not readily available to the engineer.

ANN models is a powerful tool to provide good solution for circumstances having complex system that may be poorly define or understood using mathematical equations. That is why more and more implementations of neural network models can be found in such a complex process as scour effected by the heavily anthropogenic impact such as bridge piers, piles, weirs, spillways etc.

A number of scientific works showed useful application of ANN models in assessment of scour depth around bridge piers (Bateni *et al.,* 2007; Lee *et a*l., 2007; Kaya, 2010; Rahman *et al*., 2010; Najafzadeh & Barani, 2011). The characteristics of these, and several other models which will be discussed below, are given in Table 1. Bateni *et al*. (2007) employed ANNs and ANFIS to estimate the equilibrium and time-dependent scour depth. Experimental data was used as input in two combinations: original and non-dimensional data set. Raw data produced better results than transformed. They compared the results with traditional approaches and found intelligence models are much superior. Lee *et al*. (2007) employed back-propagation neural networks trained on a field data. They choose five non-dimensional parameters to forecast the normalized scour depth. The comparisons between proposed neural model and several conventional expressions found that ANN model has good ability of forecasting the scour depth and gave more accurate results. More than 380 local scour measurements at 56 bridges in various rivers of USA were collected to develop ANN model for estimation of scour in research of Kaya (2010). Starting with 14 variables in the input layer, 4 inputs which have the main influence, namely – pier width and skew, flow depth and velocity, were chosen. The results reveal that local scour depth can be estimated with a rather appropriate accuracy, with R^2 of 0.81. Further live-bed scour and clear-water scour were assessed. Having a poor prediction ability of clear-water pier scour, due to absence of peak scour depth occurring under given conditions, the ANN model was able to accurately estimate pier scour depth for live-bed conditions. Rahman *et al*. (2010) applied several types of neural network models with single hidden layer and multiple hidden layers for estimation of scour around bridge piers with bed sill. Experimental tests were conducted under different flow conditions, distances between piers and bed sill. Authors also employed ordinary krigin and inverse distance weighting models. It was found that the ANN with two hidden layers was the optimum model to predict scour depth. Najafzadeh and Barani (2011) developed group method of data handling (GMDH) network using conventional for ANN back-propagation algorithm and genetic programming to predict the pier scour depth in bridges. They found such intelligence models can be successfully used for prediction the pier scour, which was also proved by comparison with traditional equations. The sensitive analysis indicated that non-dimensional parameter of pier diameter to flow depth (*D/y*) is the most important parameter in the modeling of scour depth. The attempt to estimate both scour depth and width using one network was done by Khosronejad *et al*. (2003). They tried two different ANNs – Model types: MLP/BP – multi layer perceptron based on back propagation algorithm; MLP/CC – multi layer perceptron based on cascade correlation algorithm; RBF/OLS – radial basis function based on orthogonal least-squares algorithm; ANFIS – adaptive neuro-fuzzy inference system; GP – genetic programming; GMDH/BP – group method of data handling network based on back propagation algorithm; GMDH/BP – group method of data handling network based on genetic programming.

Table 1. Characteristics of ANN models for erosion estimation

Reference	Erosion type	Model type	Independent variables	Dependent variables	Data used for model training and calibration	Model accuracy in R (R^2) between observed and predicted data
Azmathullah et al., 2006	Scour below spillways	MLP/BP MLP/CC	q, H	s	Field data	R = 0.92 R = 0.90
Bateni et al., 2007	Scour around bridge piers	MLP/BP RBF/OLS ANFIS	Y, U, U_c, d_{50}, D	s	Experimental data	R^2 = 0.99 R^2 = 0.84 R^2 = 0.94
Lee et al., 2007	Scour around bridge piers	MLP/BP	U/U_c, $U/\sqrt{gY}$, Y/D, D_{50}/D, σ	s/D	Field data, 1966-1993	R = 0.96
Guven and Gunal, 2008	Scour down-stream grade-control structures	MLP/BP+GP	b/z, h/H, A_{50}, d_{90}/d_{50}, b/B	s/z	Experimental data	R = 0.94
Kaya, 2010	Scour around bridge piers	MLP/BP	θ, D, U, Y	s	Field data	R^2 = 0.81
Rahman et al., 2010	Scour around bridge piers	MLP/?	g, U, ρ_w, ρ_s, q, B, d_{50}, r, D*	s	Experimental data	R^2 = 0.93-0.99 (8 experimental cases)
Najafzadeh & Barani, 2011	Scour around bridge piers	GMDH/BP GMDH/GP	Fr, D/Y, d_{50}/Y, Re, σ	s/Y	Experimental data	R = 0.89 R = 0.93

Variables: s – scour depth; q – flow discharge per unit; H – drop between upstream water level and tail water level; Y – flow depth; U – mean velocity; U_c *– critical velocity;* d_i *– bed grain size for which i% of sampled particles are finer; D – pier diameter; g – gravitational acceleration; σ – geometric standard deviation of the grain size distribution; b – weir width; z – fall high; h – tail water depth; B – channel width; θ – pier skew;* ρ_w *– flow density;* ρ_s *– particle density; r – arc distance of the circular sill; D* - sill diameter; Fr – Froude number; Re – Reynolds number.*

Multilayer perceptron with different learning rules and radial basis functions and found that a network with two output neurons – scour depth and width – was not able to provide desired mapping. Using to networks each with one output neurons can estimate depth and width of scour correctly. It was demonstrated that the designed neural networks could intelligently model nonlinear relationships between input parameters – such as wave height, water depth, maximum flow velocity, maximum shear velocity, wave period and Shields parameters; and output parameters – scour depth and width with high accuracy.

Estimation of scour below spillways with the help of neural networks was successfully done in Azmathullah *et al.* (2006). Both network models, with back-propagation and cascade correlation algorithms gave much superior accuracy than traditional formulae, with the correlation coefficient of predicted and observed results of 0.92 and 0.90 respectively.

Guven and Gunal (2008) proposed alternative method to the conventional nonlinear regression approaches for erosion process prediction downstream of grade-control structures. Using neural network model they obtained an explicit expression for prediction of maximum scour depth. Performed estimations by such model were much better than those of conventional nonlinear regression equations.

Stated above prove that using such intelligence models can be very useful and necessary in research of erosion processes in rivers. Even such a complex and multifactor phenomena as scour caused by artificial structures can be accurately estimated and predicted with the help of neural network models.

Application of ANN models in sediment transport and river sedimentation

Sediment transport is an essential part of river channel processes. Due to changes of a sediment balance in stream flow, channel may response in morphological transformations sometimes on a very large scale. It is essential to understand the principals of sediment transport in many fields one way or another related to river systems. Precise estimation of sediment in river is of high importance in water resources and hydraulic engineering, hydropower, sanitary engineering, fishery, recreation and scientific interests.

Motivated by successful implementation in a modeling of a complex, nonlinear system behavior in a broad range of area, artificial neural networks have been used in research of river sediment transport. A number of scientific publications demonstrate successful application examples of useful and reliable ANN models proposed for estimation and prediction of bed-load and suspended sediments load in rivers. In such models the former was usually assessed by the various parameters of flow, sediment characteristic and channel morphology data (Bhattacharya *et al.*, 2007a,b; Caamano *et al.*, 2006; Arrifin *et al.*, 2008; Azamathulla *et al*., 2009; Sasal *et al*., 2009), while the later – mostly by the water discharge and existed data concerning suspended sediment concentration (Cigizoglu & Kisi, 2006; Cigizoglu & Alp, 2006; Partal & Cigizoglu, 2008; Rajaee *et al*., 2009; Kisi, 2010; Rajaee, 2011; Kisi et al., 2012a,b). In several publications river suspended sediment estimation has been carried out using additionally climatic variables (Zhu *et al*., 2007; Alp & Cigizoglu, 2007; Cobaner *et al.*, 2009; Kisi & Shiri, 2012).

Bhattacharya et al. (2005) proposed a model to predict total sediment transport based on ANNs. This model outperformed conventional models of Van Rijn and Engelund-Hansen both for flume and field data estimation. Advantages in predictive accuracy of intelligence models in comparison to conventional methods in sediment transport were also shown in the work of the same group of authors (Bhattacharya *et al*., 2007a). Caamano *et al*. (2006) employed a simple neuron network which provided appropriate results in bed-load sediment transport by using just 4 input parameters: grain Froude and Reynolds numbers, characteristics of the particle size distribution of the transport sediment and relative roughness. Even less input parameters were used in neural network to successfully predict bed-load transport in a different channel types, including sand-bed and gravel-bed rivers (Sasal et al., 2009). As input variables in network only dimensionless particle diameter D* and transport stage velocity T were sufficient in accurate prediction. Arrifin *et al*. (2008) used general regression neural network for development sediment transport model applicable both for natural and artificial channels. They used almost half of thousand hydraulic and sediment field data extracted from various rivers in Malaysia and USA and a data from irrigation canal in Pakistan for model development, testing and validation. The results showed that proposed model predicts more accurately sediment transport for local and foreign rivers than presently available methods. The prediction of river bed-load was also successfully done using ANFIS (Azamathulla *et al*., 2009). These authors recommended ANFIS model for computing bed-load transport rates in moderately sized rivers, which provided much closer agreement with the measured values in comparison to existing equations.

Cigizoglu and Alp (2006) used only flow data to estimate daily suspended sediment. They found the best accuracy (with determination coefficient of 0.96) using as inputs the flow data for a time period of three days (Q_t, Q_{t-1} and Q_{t-2}). Partal and Cigizoglu (2008) developed ANN and WANN model for estimation and forecasting of daily suspended sediment load in rivers. They manipulated with different time combinations of water discharge and suspended sediment data measured in upstream and downstream gauging station to estimate suspended sediments. The compared results with conventional ANN model and SRC method showed to be WANN model prediction significantly superior. The peak sediment values and the cumulative sediment sum are closely approximated with the wavelet-ANN method. The combinations of flow data and suspended sediment concentration for different time period were also tested as inputs parameters by Rajaee *et al*. (2009) who examined neuro-fuzzy, ANN, multi linear regression (MLR) and SRC models in simulation of daily suspended sediment concentration. Comparison of the model's results indicated that the neuro-fuzzy model had more ability in predicting suspended sediment concentration in rivers. Later Rajaee developed WANN model for daily suspended sediment load prediction (Rajaee, 2011). In the proposed model the water discharge (Q) and suspended sediment load (SSL) signals were firstly decomposed into sub-signals with different scales in order to obtain temporal properties of the input time series. The decomposed water discharge and suspended sediment load time series by wavelet analysis were entered to the ANN model for prediction of SSL in one day ahead. This provided the improvement of conventional ANN method up on 46 %. The results of WANN model were also compared with the results of MLR and SRC models and it indicated that WANN model was able to accurately predict SSL. Moreover, such hybrid model could satisfactory simulate hysteresis phenomenon, estimate cumulative SSL and predict high SSL values. Significant contribution in suspended sediment modeling was done by Kisi (Kisi, 2010; Kisi et al., 2012a, b) who used various soft computing techniques to develop robust models.

River sediment transport is a very complicated process governed by a number of factors. Despite such complexity, artificial neural networks which are imitates in certain aspects a human brain functioning are capable to understand a very complex phenomena. That is why it gained popularity among researches to develop neural network models in estimation and prediction of sediment transport. The above stated examples showed superiority of ANN models over conventional methods.

Conclusions

Any interaction with river system requires detailed consideration of channel evolution. The complexity of physical processes occurs not only within channel-floodplain complex, but as well in catchment causes complicated processes in river channel evolution. Many traditional methods and approaches are not able to satisfactory analyze the main principals of erosion and sediment processes. This has motivated researches to develop alternative methods. Complicated non-linear relations and physical essence of such processes induced developers to seek a solution directly in complex creations of nature. For the basis was taken a functioning of a human brain, what have led to the development of intelligence models based on artificial neural networks. Such models are fast teachable on examples without necessity of simplification or assumption in the problem. They have great generalization and abstraction abilities. It enables to receive a precious result even if to provide a network with several distorted input variables. The computations of the components are independent, what makes ANNs parallel computational models. Such networks are able to overcome the stochastic nature of river channel evolution. From recently, new approaches, relied on intelligence models are proposed. The main attention in this paper was focused on a review of artificial neural network models and its applications in river channel research. Such models proved themselves as accurate and robust in analysis of river erosion and sediment transport. They may serve as a reliable tool for channel evolution research. Besides conventional neural network models, hybrid models such as adaptive neuro-fuzzy inference system (ANFIS), ANN developed with genetic programming (ANN-GP), wavelet artificial neural networks (WANN) and artificial neural network with artificial bee colony algorithm (ANN-ABC) were considered. In many cases such combinations of techniques improved the models accuracy in estimation and forecast. It was found that intelligence models are more robust and precise in comparison to traditional models, and capable to understand complex substance of river channel processes.

Acknowledgement: *This study became possible thanks to Program №411 - 13.04.2011 of the Ukraine Cabinet of Ministers. The author would like to extend his appreciation to Prof. Dr. Peter Rutschmann and Dr. Minh Duc Bui from the Institute of Hydraulic and Water Resources Engineering, Technical University of Munich, Germany, for their kind support.*

References

Abbott MB, Bathurst JC, Cunge JA, O'Connell PE, Rasmussen J, (1986) An introduction to the European hydrological system – Systeme Hydrologique European, SHE. 1. History and philosophy of a physically-based, distributed modeling system. *J. Hydrology*, **87**, 45-59.

Agil M, Kita I, Yano A, Nishiyama S, (2007) Analysis and prediction of flow from local source in a river basin using a Neuro-fuzzy modelling tool. *J. Environ. Manag*., **85**, 215-223.

Alatas B, (2010) Chaotic bee colony algorithms for global numerical optimization. *Expert Systems with Applic*., **37**, 5682-5687.

Alp M, Cigizoglu HK, (2007) Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. *Environ. Model. & Software*, **22**, 2-13.

Anctil F, Tape DG, (2004) An exploration of artificial neural network rainfall-runoff forecasting combined with wavelet decomposition. *J. Environ. Eng. & Science*, **3**, 121-128.

Ariffin J, Kamal NA, Sa'adon MS, Taib MN, Abdul-Talib S, Abd-Ghani A, Zakaria NA, Yahaya AS, (2008) Sediment model for natural and man-made channels using general regression neural network. *J. The Inst. of Eng., Malaysia*, **69**, 44-58.

ASCE Task Comitte on Application of Artificial Neural Networks in Hydrology, (2000a) Artificial neural networks in hydrology. I: preliminary concepts. *J. Hydrol. Eng. ASCE* **5**, 115-123.

ASCE Task Comitte on Application of Artificial Neural Networks in Hydrology, (2000b) Artificial neural networks in hydrology. I: hydrologic applications. *J. Hydrol. Eng. ASCE* **5**, 124-137.

Astafyeva NM, (1996) Wavelet-analysis: Theory and application examples (in Russian). Success of physical sciences, **166**, 1145-1170.

Azmathullah HMd, Chang CK, Ghani AA, Ariffin J, Zakaria NA, Hasan ZA, (2009) An ANFIS-based approach for predicting the bed load for moderately sized rivers. *J. Hydro-Environ. Res*, **3**, 35-44.

Azmathullah HMd, Deo MC, Deolalikar PB, (2006) Estimation of scour bellow spillways using neural networks. *J. Hydraulic Research*. **44**, 61-69.

Banzhaf W, Nordin P, Keller P.E, Francone FD, (1998) *Genetic programming. Morgan Kaufmann*, San Francisco, CA, 512 pp.

Bateni S.M, Borghei S.M, Jeng D-S, (2007) Neural network and neuro-fuzzy assessment for scour depth around bridge piers. Eng. Appl. Artif. Intell. 20, 401-414.

Beck MB, (1987) Water quality modeling: a review of uncertainty. *Water Resources Res*, **23**, 1393-1442.

Bennett J.P, (1974) Concepts of mathematical modelling of sediment yield. *Water Resources Res*, **10**, 485-492.

Bhattacharya B, Deibel IK, Karstens SAM. Solomatine DP, (2007) Neural networks in sedimentation modeling for the approach channel of the port area of Rotterdam. *Estuarine & Coastal Fine Sedim. Dynam*., **8**, 477-492.

Bhattacharya B, Price RK, Solomatine DP, (2005) Data-driven modeling in the context of sediment transport. *Physics & Chem. Earth*, **30**, 297-302.

Bhattacharya B, Price R.K, Solomatine D.P, (2007) Machine learning approach to modeling sediment transport. *J. Hydraulic Eng, ASCE*, **133**, 440-450.

Caamano D, Goodwin P, Manic M, (2006) Derivation of a bedload sediment transport formula using artificial neural networks. 7th International Conference on Hydroinformatics HIC, Nice, France.

Canas B, Fanni A, Sias G, Tronei S, Zedda MK, (2005) *River flow forecasting using neural networks and wavelet analysis.* EGU, European Geosciences Union, Vienna: Austria, p. 24-29.

Chalov RS, (1979) *Geographical investigation of river channel processes* (in Russian). Press of Moscow State University, 232 p.

Chen K, Yang L.P, Yu X, Chi H.S, (1997) A self-generating modular neural network architecture for supervised learning. *Neurocomputing* **16**, 33-48.

Cigizoglu H.K, Alp M, (2006) Generalized regression neural network in modeling river sediment yield. *Advances in Eng. Software*, **37**, 63-68.

Cigizoglu HK, Kisi O, (2006) Methods to improve the neural network performance in suspended sediment estimation. *J. Hydrology*, **317**, 221-238.

Cobaner M, Unal B, Kisi O, (2009) Suspended sediment concentration estimation by an adaptive neuro-fuzzy and neural networks approaches using hydro-meteorological data. *J. Hydrology*, **367**, 52-61.

Daubechies I, (1992) Ten lectures on Wavelets. SIAM: Society for Industrial and Applied Mathematics. ISBN-10: 0898712742.

Elshorbagy A, Corzo G, Srinivasulu S, Solomatine DP, (2010) Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology – Part 1: Concepts and methodology. *Hydrology Earth Syst. Sci*., **14**, 1931-1941.

Elshorbagy A, Corzo G, Srinivasulu S, Solomatine DP, (2010) Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology – Part 2: Application. *Hydrology Earth Syst. Sci*., **14**, 1943-1961.

Flood I, Kartam N, (1994) Neural networks in civil engineering. I: Principles and understanding. *J. Computing & Civil Eng*., **8**, 131-148.

Goldberg D, (1989) *Genetic Algorithms*. Addison-Wesley, Reading, MA.

Guven A, Gunal M, (2008) Prediction of scour downstream of grade-control structures using neural networks. *J. Hydraul. Eng*. **134**, 1656-1660.

Haykin S, (1999) *Neural Networks: A comprehensive Foundation*. Prentic-Hall, Upper Saddle River, New Jersey, 842 pp.

Hirose Y, Yamashita K, Hijiya S, (1991) Back-propagation algorithm which varies the number of hidden units. *Neural Networks* **4**, 61-66.

Hornik K, Stinchcombe M, White H, (1989) Multilayer feedforward networks are universal approximators. *Neural networks*, **2**, 359-366.

Jain A, Sudheer KP, Srinivasulu S, (2004) Identification of physical process inherent in artificial neural network rainfall runoff models. *Hydrological Processes*, **18**, 571-581.

Jakeman AJ, Hornberger GM, (1993) How much complexity is warranted in a rainfall-runoff model? *Water Resources Research*, **29**, 2637-2649.

Jang J-SR, (1993) ANFIS: adaptive-network-based fuzzy inference systems. IEEE Transactions on Systems Man and Cybernetics 23(03), 665-685.

Jang J.-SR, Sun CT, Mizutani E, (1997) *Neurofuzzy and soft computing: a computational approach to learning and machine intelligence*. Prentice-Hall, New Jersey.

Kamp RG, Savenije HHG, (2006) Optimising training data for ANNs with genetic algorithms. *Hydrology & Earth Syst. Sci*, **10**, 603-608.

Karaboga D, (2005) *An idea based on honey bee swarm for numerical optimization*. Technical report TR06, Erciyes University, Eng. Faculty, Computer Eng. Depart.

Karnin ED, (1990) A simple procedure for pruning back-propagation trained neural networks. *IEEE Transactions on Neural Networks*, **1**, 239-242.

Kaya A, (2010) Artificial neural network study of observed pattern of scour depth around bridge piers. *Computers & Geotechnics*, **37**, 413-418.

Khosronejad A, Montazer G.A, Ghodsian M, (2003) Estimation of scour hole properties around vertical pile using ANNs. Int. J. of Sediment Research, **18**, 290-300.

Kirkpatrick S, Gilatt CD, Vecchi MP, (1983) Optimization by simulated annealing. *Sci*. **220**, 671-680.

Kisi O, (2003) Modeling of suspended sediment yield in a river cross-section using fuzzy logic. Ph.D. Thesis, Istanbul Technical University, Institute of Science and Technology, Istanbul, Turkey.

Kisi O, (2005) Suspended sediment estimation using neuro-fuzzy and neural network approaches. *Hydrol. Sciences J*. **50**, 683-696.

Kisi O, Haktanir T, Ardiclioglu M, Ozturk O, Yalcin E, Uludag S, (2009) Adaptive neuro-fuzzy computing technique for suspended sediment estimation. *Adv.In Eng.Software*, **40**, 438-444.

Kisi O, (2010) River suspended sediment concentration modeling using a neural differential evolution approach. *J. Hydrology*, **389**, 227-235.

Kisi O, Dailr AH, Cimen M, Shiri J, (2012) Suspended sediment modeling using genetic programming and soft computing techniques. *J. Hydrology*, **450**-41, 48-58.

Kisi O, Ozkan C, Akay B, (2012) Modeling discharge-sediment relationship using neural networks with artificial bee colony algorithm. *J. Hydrology*, **428**-429, 94-103.

Kisi O, Shiri J, (2012) River suspended sediment estimation by climatic variables implication: Comparative study among soft computing techniques. *Comput. & Geosci*. **43**, 73-82.

Koza JR, (1992) *Genetic programming: On the programming of computers by means of natural sections.* The MIT Press, Cambridge, MA.

Lee TL, Jeng DS, Zhang GH, Hong JH, (2007) Neural network modeling for estimation of scour depth around bridge piers. *J. Hydrodynamics. Ser. B*, **19**, 378-386.

Lohani AK, Goel NK, Bhatia KK, (2007) Deriving stage-discharge-sediment concentration relationships using fuzzy logic. *Hydrol. Sciences J*., **52**, 793-807.

Maier HR, Dandy GC, (2000) Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications. Envir. Model. & Software, **15**, 101-124.

Makkaveev NI, (1955) *River channel and erosion in its catchment (in Russian). Moscow.* Academy of science of USSR Press, 347 p.

Mamdani EH, Assilian S, (1975) An experiment in linguistic synthesis with a fuzzy logic controller. *Int. J. Man Machine Studies* **7**, 1-13.

Masters T, (1993) Practical neural network recipes in C++, Academic Press, San Diego, California.

McCulloch WS, Pitts WS, (1943) A logical calculus of the ideas imminent in nervous activity. *Bulletin & Math. Biophy*., **5**, 115-133.

Merritt WS, Letcher RA, Jakeman AJ, (2003) A review of erosion and sediment transport models. *Envir. Model. & Software*, **18**, 761-799.

Minns AW, (2000) Subsymbolic methods for data mining in hydraulic engineering. *J. Hydroinformatics*, **2**, 3-13.

Najafzadeh M, Barani Gh.-A, (2011) Comparison of group method of data handling based genetic programming and back propagation systems to predict scour depth around bridge piers. *Scientica Iranica, Transactions A: Civil Eng*. **18**, 1207-1213.

Olden JD, Joy MK, Death RG, (2004) An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. *Ecological Modeling*, **178**, 389-397.

Partal T, Cigizoglu HK, (2008) Estimation and forecasting of daily suspended sediment data using wavelet-neural networks. *J. Hydrology*, **358**, 317-331.

Rahman HS, Alireza K, Reza G, (2010) Application of artificial neural network, kriging, and inverse distance weighting models for estimation of scour depth around bridge pier with bed sill. *J. Software Eng. And Appl.* **3**, 944-964.

Rajaee T, (2011) Wavelet and ANN combination model for prediction of daily suspended sediment load in rivers. *Sci. Total Environ*. **409**, 2917-2928.

Rajaee T, Mirbagheri SA, Zounemat-Kermani M, Nourani V, (2009) Daily suspended sediment concentration simulation using ANN and neuro-fuzzy models. *Sci. Total Enviro*n. **407**, 4916-4927.

Rao SS, (1996) Engineering optimization. 3rd Ed, Wiley, New York, 806-823.

Rezapour OM, Shui LT, Ahmad DB, (2010) Review of artificial neural network model for suspended sediment estimation. Australian J. of Basic and Appl. Sci, 4(8), 3347-3353.

Ripley BD, (1994) Neural networks and related methods of classification. *J. Royal Statist.Sociaty, B* **56**, 409-456.

Rojas R, (1996) *Neural networks - a systematic introduction. The backpropagation algorithm.* Springer Verlag, Berlin (Chapter 7).

Rumelhart DE, Hinton GE, Williams RJ, (1986) Learning internal representations by error propagation. In: Rumelhart D.E, McClelland J.L. (Eds), *Parallel Distributed Processing*. MIT press, Cambridge.

Sarley WS, (1994) Neural networks and statistical models. In: *Proceedings of the Nineteenth Annual* SAS Users Group Int. Conf, pp. 1538-1550. SAS Institute.

Sasal M, Kashyap S, Rennie CD, Nistor I, (2009) Artificial neural networks for bedload estimation in alluvial rivers. *J. Hydraulic Res.*, **47**, 223-232.

Setiono R, (1997) A penalty-function approach for pruning feedforward neural networks. *Neural Computation* **9**, 185-204.

Shahin MA, Jaksa MB, Maier HR, (2008) State of the art of artificial neural networks in geotechnical engineering. *Electr. J. Geot. Eng*. ISSN: 1089-3032.

Shu C, Quadra TBMJ, (2008) Regional flood frequency analysis at ungauged sites using the adaptive neuro-fuzzy inference system. *J. Hydrology*, **349**, 31-43.

Sietsma J, Dow RJF, (1991) Creating artificial neural networks that generalize. *Neural Networks* **4**, 67-79.

Sorooshian S, (1991) "Parameter estimation, model identification and model validation: Conceptual-type models," In: *Recent advantages in the modeling of hydrologic systems*. Eds.: Bowles D.S. and O'Connell P.E, Kluwer Academic Publ, pp. 443-467.

Sudheer KP, (2005) Knowledge extraction from trained neural network river flow models. *J. Hydrologic Engineering*, **10**, 264-269.

Sudheer KP, Nayak PC, Ramasastri KS, (2003) Improving peak flow estimates in artificial neural network river flow models. *Hydrological Processes*, **18**, 833-844.

Takagi T, Sugeno M, (1985) Fuzzy identification of systems and its application to modeling and control. IEEE Transactions on System, *Man and Cybernetics* **15**, 116-132.

Tokar SA, Johnson PA, (1999) Rainfall-runoff modeling using artificial neural networks. *J. Hydrologic Engineering*, **4**(3), 232-239.

Wang D, Ding J, (2003) Wavelet network model and its application to the prediction of hydrology. *Nature & Science*, **1**, 67-71.

Wasserman PD, (1989) Neural computing: theory and practice. Van Nostrand Reinhold, New York.

Wheater HS, Jakeman AJ, Beven KJ, (1993) Progress and directions in rainfall-runoff modelling. In: Jakeman A.J, Beck M.B, McAleer M.J. (Eds.), *Modelling change in environmental systems*. John Wiley and Sons, Chichester, pp. 101-132.

Zaentsev IV, (1999) *Neuron networks: Main models* (in Russian). Voronezhskiy State University.

Zurada J.M, (1992) Introduction to artificial neural systems. West Publishing Company, St. Paul.

Zhu Y-M, Lu XX, 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.