KSCE Journal of Civil Engineering (2016) 20(2):581-589

DOI 10.1007/s12205-015-0630-7

Environmental Engineering

pISSN 1226-7988, eISSN 1976-3808
www.springer.com/12205

Bed Load Sediment Transport Estimation in a Clean Pipe using Multilayer Perceptron with Different Training Algorithms

Isa Ebtehaj* and Hossein Bonakdari**

Received October 25, 2014/Revised February 6, 2015/Accepted February 24, 2015/Published Online May 7, 2015

Abstract

Due to the presence of solid matter in the flow passing through sewer pipes, determining the minimum velocity that prevents sediment deposition is essential. In this study, the Multilayer Perceptron (MLP) network optimized with three different training algorithms, including variable learning rate (MLP-GDX), resilient back-propagation (MLP-RP) and Levenberg-Marquardt (MLP-LM) is studied in terms of ability to estimate sediment transport in a clean pipe. The results indicate that for all algorithms, model ANN(d) that uses volumetric sediment concentration (C_V), median relative size of particles (d/D), ratio of median diameter particle size to hydraulic radius (d/R) and overall sediment friction factor (λ_s) as input parameters, is more accurate than the other models. In predicting Fr, the results of MLP-LM ($R^2 = 0.98$, $RMSE = 0.02$ and $MAPE = 5.13$) are better than MLP-GDX ($R^2 = 0.96$, $RMSE = 0.03$ and $MAPE = 5.9$) and MLP-RP ($R^2 = 0.95$, $RMSE = 0.26$ and $MAPE = 5.74$). A comparison of the model selected in this study with existing equations of sediment transport in sewer pipes also indicates that ANN(d)-LM ($RMSE = 0.025$ & $MAPE = 5.78$) perform better than existing equations.

Keywords: *bed load, limit of deposition, multilayer perceptron, sediment transport, sewer*

1. Introduction

Among the essential topics regarding sediment transport in sewer pipes is the deposition of suspended sediment in the inlet to the pipe. The pipe diameter must be determined such that it results in the capability to simultaneously transfer maximum discharge and minimum discharge without deposition. Solid deposition on the channel bed increases the bed's roughness and decreases the cross sectional area of the flow. A decrease in the flow cross sectional area leads to local flooding or surcharge, thus creating a septic condition that causes problems such as odor (Bonakdari and Larrate, 2006). One of the simplest methods of designing sewer pipes is to use minimum velocity and shear stress. However, in view of the fact that this method does not consider the factors affecting flow transfer (e.g., flow and sediment characteristics), it may lead to over or under design (Ebtehaj *et al*., 2014). Therefore, different researchers have considered the parameters influencing flow transfer and presented various equations for determining the minimum velocity needed to transfer flow without deposition (Macke, 1982; Nalluri, 1985, Mayerle *et al*., 1991; May, 1993; Ota and Nalluri, 2003). In new design criteria, known as self-cleansing design, the minimum velocity required to prevent sediment deposition is considered such that it minimizes the construction, performance, and maintenance costs simultaneously so as to make the design economical (Butler *et al*., 2003). At limit of deposition, the aim is to determine the minimum velocity required to prevent sediment deposition. This concept is sometimes referred to as "no-deposition." However, for bed deposition, the essential minimum velocity is determined for conditions in which a maximum of 2% deposition is permissible (Butler *et al*., 2003).

Numerous researchers have attempted to study sediment transport in pipe channels by conducting different theoretical analyses and experimental works on non-cohesive sediment transport in channels. Ackers *et al.* (1996) assumed different hydraulic conditions to control the problems caused by sedimentation and presented a method of designing sewers. Ota and Nalluri (2003) used physical concepts and proposed a new model for economically designing large-diameter sewer pipes ($D > 500$ mm). Based on experimental results for partial and non-cohesive sediment deposits in pipes, Banasiak (2008) examined deposited sediment behavior in sewers and its effects on hydraulic performance. It was found that an increase in sediment depth from 2 to 10% results in significant transport capacity reduction (10-20%) relative to the clean pipe. For designing sewers with allowable sediment depth proposed in the new concept of self-cleansing, Butler *et al*. (2003) suggested that the appearance of bed forms must be avoided. Almedeij and Almohsen (2010) made a number of remarks on Camp's criterion in order to present a more flexible equation with respect to the required

*Ph.D. Candidate, Dept. of Civil Engineering, Razi University, Kermanshah 67149-67346, Iran (E-mail: isa.ebtehaj@gmail.com)
**Associate Professor, Dept. of Civil Engineering, Razi University, Kermanshah 67149-67346, Iran (Corresponding Author, E-mail: bonakdari@yahoo.com)

minimum velocity used in Kuwait. Vongvisessomjai *et al.* (2010) conducted a number of experiments under partially full flow conditions and presented a new criterion for the non-deposition state. Almedeij (2012) presented a new designing criterion for rectangular sewers under equal sediment mobility conditions. Ota and Perrusquia (2013) used the limit of deposition concept and offered an equation for calculating dimensionless bed shear stress in pipe channels. Bong *et al.* (2013) examined existing incipient motion equations by conducting a series of experiments and understood that as the bed roughness increases, the existing equations become less accurate. They presented an equation for incipient motion by considering the effects of bed thickness. Using dimensionless parameters and the smoothing function, Bonakdari and Ebtehaj (2014) proposed a different relation to estimate Fr in sewers using a smoother function. It was found that the smoother function leads to more accurate Fr estimation. Using a wide range of data and regression analysis, Ebtehaj *et al.* (2014) modified Vongvisessomjai's (2010) equation. Based on critical velocity and regression models, Bong *et al.* (2014) expanded the incipient motion equation and designed a chart to determine the minimum velocity for a self-cleansing design of rectangular storm water.

Artificial Intelligence (AI) and especially Artificial Neural Network (ANN) perform efficiently in complex engineering problems, which has led to these methods being widely used in hydrology and water engineering (Schulz *et al.*, 2005; Nour *et al.*, 2006; Alp and Cigizoglu, 2007; Nourouzi *et al.*, 2011; Azamathulla and Zakaria, 2011; Bonakdari et al., 2011; Baghalian *et al.*, 2012; Ebtehaj and Bonakdari, 2013; Ebtehaj and Bonakdari, 2014). Nagy *et al.* (2002) used ANN and the back propagation training algorithm, and presented a model capable of predicting sediment load concentration in a river. They also evaluated the proposed model using data on different rivers. Sarangi and Bhattachayra (2005) compared two geomorphology-based (GANN) and non-geomorphology based (NGANN) types of neural networks in predicting the sediment load from the Banha Watershed in India. They observed that GANN presents better results than NGANN. Alp and Cigizoglu (2007) used Radial Basis Functions (RBF) and Feed-Forward Back-propagation (FFBP) algorithms to train ANN in order to estimate the load of suspended sediment. They showed that both algorithms present better results than multiple linear regression. Partal (2009) used Wavelet to train ANN to predict river flow. Partal compared the prediction results of this algorithm with those of the feed-forward back-propagation, Radial Basis Function (RBF), and Generalized Neural Network (GNN) and found that the presented model is superior to the rest. Melesse *et al.* (2011) compared different methods including ANN, Autoregressive Integrated Moving Average (ARIMA) and multi-linear and non-linear regression (MLR and LNLR) in predicting time series and came to the conclusion that ANN benefits from greater capability of predicting suspended sediment load compared with MLR, MNLR, and ARIMA.

Three different algorithms, namely Multi-Layer Perceptron (MLP) including variable learning rate (MLP-GDX), resilient back-propagation (MLP-RP) and Levenberg–Marquardt (MLP-LM) are used for training to predict the densimetric Froude number (Fr) in this study. The factors influencing sediment transport in sewer pipes are first examined, determined, and eventually placed in five different groups. Subsequently, the dimensionless parameters presented in these groups will be used and the Fr parameter is predicted using the dimensionless parameters presented in other dimensionless groups. The results of the model presented in this study will also be compared with existing sediment transport equations.

2. Existing Sediment Transport Equations

Equations of non-cohesive sediment transport in sewer pipes can be divided into two general groups: semi-experimental equations and dimensional analysis. The equations obtained from dimensional analysis can further be divided into three different groups depending on the use of different affective parameters. Novak and Nalluri (1975) presented the equations below by simultaneously using transport parameter φ, flow parameter ψ and the Darcy-Weisbach resistance equation ($S_0 = \lambda_s V^2/8gR$).

$$\varphi = \frac{C_V VR}{\sqrt{(s-1)gd^3}} \tag{1}$$

$$\psi = \frac{(s-1)d}{RS_0} \tag{2}$$

$$Fr = \frac{V}{\sqrt{g(s-1)d}} = 1.77 C_V^{1/3} \left(\frac{d}{R}\right)^{-1/3} \lambda_s^{-2/3} \tag{3}$$

where C_V is volumetric sediment concentration, R is hydraulic radius, V is flow velocity, λ_s is overall sediment friction factor, s is the specific gravity of sediment, d is the median diameter of particles and S_0 is the channel slope.

The second group contains equations similar to the equations presented by Novak and Nalluri (1975), except they use the dimensionless particle number D_{gr} ($=d(g(s\text{-}1)/\nu^2)^{1/3}$) parameter in addition to the parameters used in the previous group. Azamathulla *et al.*'s (2012) equation is as follows:

$$Fr = \frac{V}{\sqrt{g(s-1)d}} = 0.22 C_V^{0.16} D_{gr}^{-0.14} \left(\frac{d}{R}\right)^{-0.29} \lambda_s^{-0.51} \tag{4}$$

where λ_s is calculated as shown below using the equation presented by Nalluri and Kithsiri (1992):

$$\lambda_s = 0.851 \lambda_c^{0.86} C_V^{0.04} D_{gr}^{0.03} \tag{5}$$

where λ_c is the clear water friction factor of the channel.

The third group considers dimensionless parameters affecting Fr prediction through the following equation, as presented by Ebtehaj *et al.* (2014):

$$Fr = \frac{V}{\sqrt{g(s-1)d}} = 4.49 C_V^{0.21} \left(\frac{d}{R}\right)^{-0.54} \tag{6}$$

May *et al.* (1996) used 7 different sets of data (Macke, 1982; May *et al.*, 1989; Mayerle *et al.*, 1991; May, 1993; Ab Ghani, 1993; Nalluri *et al.*, 1994) and presented the following equations as the best known sediment transport equations at limit of deposition (Ackers *et al.*, 1996). Vongvisessomjai *et al.* (2010) have also used them to verify their own equations.

$$C_V = 3.03 \times 10^{-2}\left(\frac{D^2}{A}\right)\left(\frac{d}{D}\right)^{0.6}\left(\frac{V^2}{g(s-1)D}\right)^{1.5}\left(1-\frac{V_t}{V}\right)^4 \tag{7}$$

$$V_t = 0.125[g(s-1)d]^{0.5}\left[\frac{y}{d}\right]^{0.47} \tag{8}$$

where C_V is volumetric sediment concentration, A is cross-sectional area of the flow, D is the pipe diameter, V is the flow velocity, d is the median diameter of particle size, V_t is the velocity required for incipient motion of the sediment, and y is the depth of flow.

3. Data Used

The experimental data relating to sediment transport in non-deposition mode employed in this article include Ab Ghani's (1993) data, which were used for training the model and Vongvisessomjai *et al.*'s (2010) data, which served to verify the model. Ab Ghani employed three different pipes with 154, 305, and 450 mm diameter and constant length of 20.5 m to conduct experiments. The hydraulic parameter ranges examined by Ab Ghani (1993) are as follows: $0.24 < V$ (m/s) < 1.216; $1 < C_V$ (ppm) < 145; $0.072 < d$ (mm) < 8.3; $0.033 < R$ (m) < 0.136, $0.153 < y/D < 0.8$ and $0.0007 < S_0 < 0.0056$. Vongvisessomjai *et al.* (2010) conducted experiments with two pipes that were different from those of Ab Ghani (1993). The pipes were 16 meters long and had diameters of 100 and 150 mm. The ranges of parameters measured were: $0.237 < V$ (m/s) < 0.626; $4 < C_V$ (ppm) < 90; $0.2 < d$ (mm) < 0.43; $0.032 < R$ (m) < 0.012, $0.2 < y/D < 0.4$ and $0.002 < S_0 < 0.006$.

4. Neural Network Training Algorithms

Variable learning rate, resilient back-propagation, and Levenberg-Marquardt are the three algorithms used in this study to train the neural network. The major objective of training the neural network is to minimize the global error (E), which is defined as follows:

$$E = \frac{1}{p}\sum_{p=1}^{p} E_p \tag{9}$$

where p is the total number of training patterns and E_p is the training pattern (p) error defined as:

$$E_p = \frac{1}{2}\sum_{k=1}^{n}(o_k - t_k)^2 \tag{10}$$

where t_k and o_k are the target and network output for the k^{th} output node (respectively) and n is the total number of output nodes. The main goal of all algorithms presented next is to reduce the global error value using different weights and biases (Kisi, 2007).

4.1 Variable Learning Rate Algorithm

The training rate remains constant during the neural network training process by employing standard steepest descent. Gradient descent is expressed as follows:

$$x_{k+1} = x_k - \alpha_k g_k \tag{11}$$

where x_k is a vector of current biases and weights, α_k is the learning rate and g_k is the current gradient.

This algorithm's performance is especially sensitive to training rate selection, such that an excessively low training rate leads to algorithm convergence that takes a long time. Meanwhile, selecting an excessively large value leads to instability. It is impractical to determine optimal specifications for the learning rate before the training process begins. In fact, the optimal learning rate changes during the training process with respect to the algorithm's performance level. Using an adaptive learning rate can yield the maximum learning rate, which in turn leads to stable learning. By using an adaptive learning rate, we attempt to use the largest possible value of the learning step size until learning becomes stable.

4.2 Resilient Back-propagation Algorithm

Sigmoid activation function is commonly used within the hidden layers in multi-layer networks. Since these sorts of functions compress the infinite input range into a finite output range, they are commonly called squatting functions. Sigmoid functions are described such that their slope must get close to zero as the input gets larger.

The small magnitude of the gradient causes changes in the magnitude of biases and weights when the steepest descent is used to train a multi-layered network along with the sigmoid function. The probability of achieving optimal values of weights and biases will therefore decrease (Riedmiller and Braun, 1993). The resilient back-propagation algorithm is presented for the purpose of eliminating the effects of magnitudes of partial derivatives which prevent attaining the optimal global error. Only the sign of the derivative affects the direction of the updated weight in this algorithm and the magnitude of the derivative has no effect on it. The weight change size is determined using separately updated values. To do so, a w_{ij} with a $\Delta_{ij}(p)$ individual updated value is defined for each weight that determines the size of the updated weight only. Therefore, a secondary training rate which determines the evaluation of the updated value $\Delta_{ij}(p)$ is presented. This secondary rate is presented below based on the partial derivatives' behavior in two consecutive weight steps:

$$\Delta w_{ij}(p) = \begin{cases} -\Delta w_{ij}(p) & \text{if } \delta E(W(p-1))/(\delta w_{ij}) > 0 \\ \Delta w_{ij}(p) & \text{if } \delta E(W(p-1))/(\delta w_{ij}) < 0 \end{cases} \tag{12}$$

$$\Delta_{ij}(p) = \begin{cases} a\Delta_{ij}(p-1) \text{ if } \delta E(W(p-1))/(\delta w_{ij}) * \delta E(W(p-2))/(\delta w_{ij}) > 0 \\ b\Delta_{ij}(p-1) \text{ if } \delta E(W(p-1))/(\delta w_{ij}) * \delta E(W(p-2))/(\delta w_{ij}) < 0 \\ 0 < b < 1 < a \end{cases} \tag{13}$$

where $\Delta_{ij}(p)$ is an individual update value for each weight w_{ij}, w_{ij} is the weight from neuron j to neuron I and E is the global error function.

4.3 Levenberg–Marquardt Algorithm

The Levenberg-Marquardt algorithm is used to train the network with a second-order algorithm without the need to calculate the Hessian matrix (More, 1977). The Hessian matrix and gradient can be approximated as shown below in case the performance function is in the form of the sum of squares, as it is often found in training feed-forward networks:

$$H = J^T J \tag{14}$$

$$g = J^T e \tag{15}$$

where J is the Jacobian matrix, which includes the first derivatives of network error with regard to biases and weights, and e is the error vector of the network. The Jacobian matrix can be calculated through standard back-propagation method, which entails fewer calculations compared with calculating the Hessian matrix. From the approximation presented for the Hessian matrix, the Levenberg-Marquardt algorithm uses the Newton-like equation as follows:

$$x_{k+1} = x_k - [J^T J + \mu I]^{-1} J^T e \tag{16}$$

If the numerical value of μ is equal to zero in the above-mentioned equation, this equation will turn into a Newton equation obtained through Hessian matrix approximation. If μ is a greater number, it decreases to gradient descent with step size. The Newton method is close to the least amount of error and it is therefore faster and more accurate. So the aim is to move as fast as possible toward the Newton method. Therefore, μ decreases after each successful step (performance function decrease) and it increases only when the tentative test stage increases the performance function. Thus, the performance function decreases per interaction.

5. Application and Results

5.1 Development of Densimetric Froude Number Models

In order to determine the minimum velocity to prevent sediment deposition in sewer pipes, the hydraulic parameters affecting it must normally be examined. Research works conducted indicate that the most important hydraulic parameters for sediment transport in a pipe are flow velocity (V), volumetric sediment concentration (C_V), pipe diameter (D), dimensionless particle number (D_{gr}), median diameter of particle size (d), hydraulic radius (R), depth of flow (y), overall sediment friction factor (λ_s), cross sectional area of the flow (A) and the specific gravity of sediment (s) (Ab Ghani, 1993; May *et al.*, 1996; Ebtehaj *et al.*, 2014). Consequently, the dimensionless parameters can be classified in different groups including motion, transport, transport mode, sediment and flow resistance as illustrated in Table 1. The "motion", "transport", and "flow resistance" groups consider only one parameter each, Fr, C_V and λ_s respectively. The median relative size of particles (d/D), dimensionless particle number (D_{gr}) and specific gravity of the sediment (s) parameters are related to the "sediment" group, while the square of the pipe diameter to the cross sectional area of the flow (D^2/A), the ratios of hydraulic radius to the median diameter particle of size (R/d) and the ratio of relative depth of flow (y/d), which is usually replaced by the R/d ratio, are related to the "transport" group. The six groups are presented below for the purpose of examining the effect of different dimensionless parameters on predicting the Fr.

Table 1. Dimensionless Sediment Transport Parameters in Clean Pipes

Parameter type	Dimensionless groups
Movement	Fr
Transport	C_V
Sediment	D_{gr}, d/D, s
Transport mode	d/R, D^2/A, d/y, R/D
Flow resistance	λ_s

ANN (a): $Fr = f(C_V; D_{gr}, d/R, \lambda_s)$
ANN (b): $Fr = f(C_V; D_{gr}, D^2/A, \lambda_s)$
ANN (c): $Fr = f(C_V; D_{gr}, R/D, \lambda_s)$
ANN (d): $Fr = f(C_V; d/D, d/R, \lambda_s)$
ANN (e): $Fr = f(C_V; d/D, D^2/A, \lambda_s)$
ANN (f): $Fr = f(C_V; d/D, R/D, \lambda_s)$

In order to apply a neural network, MATLAB software is used for coding. One of the most important issues with ANNs is selecting proper architecture. It is essential to use a particular type of architecture for each specific problem in ANN. Considering the complexities of a problem, this architecture must directly affect the computational complexity and must be capable of generalizing that network (Jain *et al.*, 2008). The number of input and output parameters depends on the type of problem. In this case there is no specific method to determine the number of hidden layers and the number of nodes in them. Therefore, the network architecture is determined through trial and error (Sudheer and Jain, 2004; Shahin *et al.*, 2008; Ebtehaj and Bonakdari, 2013). The input layer used in this study to estimate Fr contains 4 input parameters that are presented in models ANN(a) to ANN(f). Considering that the topology of the network employed is of special importance in the calculations, determining an adequate architecture for the intended problem can significantly affect the prediction results. In view of the fact a multilayer neural network can have more than one hidden layer, but using one hidden layer often presents desirable results in complex nonlinear problems (Hornik *et al.*, 1989; Jalili-Ghazi Zade and Noori 2008; Noori *et al.*, 2010), one hidden layer was used in this study.

Taking into account that using actual and non-normalized data may lead to obtaining undesirable results when predicting the intended parameter, the utilized data (test and train) are normalized before modeling. The data will ultimately be anti-normalized after modeling ends. The following equation is used to normalize the data in this study:

$$x_{normal_i} = a\frac{x_i - x_{min}}{x_{max} - x_{min}} + b \tag{17}$$

where x_{max} and x_{min} are the maximum and minimum values of the data respectively; x_i is the data intended to normalize; and the a and b parameters are fixed values that may differ depending between which numbers the normalization range rests. Because there is no standard method for normalizing data (Dawson and Wilby, 1998), the normalization range of parameters in this study considered is [0.2 0.8] as suggested by Cigizoglu's (2003). Therefore, the values of a and b are 0.6 and 0.2 respectively.

Eighty percent (109 data) of Ab Ghani's (1993) data, which were randomly selected from among 137 data, served to train the ANN using the three algorithms presented in this article, and the remaining 20% of data (28data) were used to test the model. Then Vongvisessomjai *et al.*'s (2010) experimental data set was used to validate the presented models in order to examine the *Fr* prediction accuracy of the different models.

There are 4 neurons in the input layer of each model presented in this study (ANN(a) to ANN(f)). First, the number of hidden layer neurons was assumed to be 1, after which this number was increased up to 20. It is clear that compared with other states, 10 hidden layer neurons is optimum for all models. Since the sigmoid activation function is highly accurate in predicting sediment transport at limit of deposition (Ebtehaj and Bonakdari, 2013), the MLP neural network structure is constructed with this activation function. The results indicate the superior performance of tangent sigmoid function over the logistic sigmoid. This has been approved in recent research conducted by Rezaeian Zadeh *et al.* (2010) and Yonaba *et a*l. (2010).

5.2 Discussion

Coefficient of determination (R^2), root mean absolute error (*RMSE*) and mean absolute percentage error (*MAPE*) statistical indexes are used to examine the performance of the ANN(a) to A NN(f) models for all three algorithms presented in this study:

$$R^2 = \left[\frac{\sum_{i=1}^{n}(Fr_{EXP_i} - \overline{Fr_{EXP_i}})(Fr_{EST_i} - \overline{Fr_{EST_i}})}{\sqrt{\sum_{i=1}^{n}(Fr_{EXP_i} - \overline{Fr_{EXP_i}})^2 \sum_{i=1}^{n}(Fr_{EST_i} - \overline{Fr_{SET_i}})^2}}\right]^2 \tag{18}$$

$$MAPE = \left(\frac{100}{n}\right)\sum_{i=1}^{n}\left(\frac{|Fr_{EXP_i} - Fr_{EST_i}|}{Fr_{EXP_i}}\right) \tag{19}$$

$$RMSE = \sqrt{\left(\frac{1}{n}\right)\sum_{i=1}^{n}(Fr_{EXP_i} - Fr_{EST_i})^2} \tag{20}$$

The above-mentioned statistical indexes are used to quantitatively examine the predictions carried out by the MLP neural network. The MLP neural network served to predict the *Fr* value for each different model. Three different algorithms, including variable learning rate (MLP-GDX), resilient back-propagation (MLP-RP), and Levenberg-Marquardt (MLP-LM) were employed to train this type of ANN. Therefore, 18 different models are presented in this study. The results obtained from the predictions carried out by all models using the statistical indexes are given in Table 2. As demonstrated in this table, compared with the other models, ANN(d) is the most accurate for the three learning algorithms presented in the current study. In this model, the effective parameters in *Fr* evaluation considered are volumetric sediment concentration (C_V), proportion of particles' mean diameter to pipe diameter (d/D), relative depth of flow (R/d) and overall sediment friction factor (λ_s). As the table shows, in training MLP networks the LM algorithm yields better results for ANN(d) (the best model) than the other algorithms (R^2 = 0.98, *RMSE* = 0.02 and *MAPE* = 5.13%). Therefore, it can be stated that the d/D parameter produces better results than other dimensionless parameters in the sediment group and the d/R parameter produces better results in the transport mode dimensionless group.

Table 2. Validation of ANN Models for Different Training Algorithms using Statistical Indexes

Models	R^2	RMSE	MAPE
ANN (a) - GDX	0.94	0.0314	6.85
ANN (b) - GDX	0.94	0.0431	6.56
ANN (c) - GDX	0.92	0.0423	6.96
ANN (d) - GDX	<u>0.96</u>	<u>0.0298</u>	<u>5.90</u>
ANN (e) - GDX	0.90	0.0569	7.21
ANN (f) - GDX	0.87	0.0428	9.26
ANN (a) - RP	0.95	0.0353	6.95
ANN (b) - RP	0.95	0.0471	7.12
ANN (c) - RP	0.93	0.0557	8.05
ANN (d) - RP	<u>0.95</u>	<u>0.0264</u>	<u>5.74</u>
ANN (e) - RP	0.91	0.0579	8.52
ANN (f) - RP	0.88	0.0631	9.84
ANN (a) - LM	0.97	0.0288	6.54
ANN (b) - LM	0.96	0.0317	6.95
ANN (c) - LM	0.96	0.0302	6.54
ANN (d) - LM	**<u>0.98</u>**	**<u>0.0200</u>**	**<u>5.13</u>**
ANN (e) - LM	0.96	0.0325	6.99
ANN (f) - LM	0.96	0.0318	8.53

Accordingly, compared to the two other algorithms, i.e. variable learning and resilient back-propagation, the Levenberg-Marquardt (LM) algorithm is more accurate in estimating the *Fr*. First-order methods use a linear local approximation of error space while second-order methods use quadratic approximation. The LM training algorithm is a second-order gradient method. As previously mentioned, it utilizes second-order derivatives (Hessian matrix) and may be effective under certain conditions. Not only is the Levenberg-Marquardt algorithm more accurate than the other two algorithms, but its convergence speed is also higher so much so that the number of epochs the algorithm uses in this study is equal to 100 while the other two algorithms use 1000.

Figure 1 illustrates the results of examining the *Fr* predicted by the models using LM. It is evident that the error of model

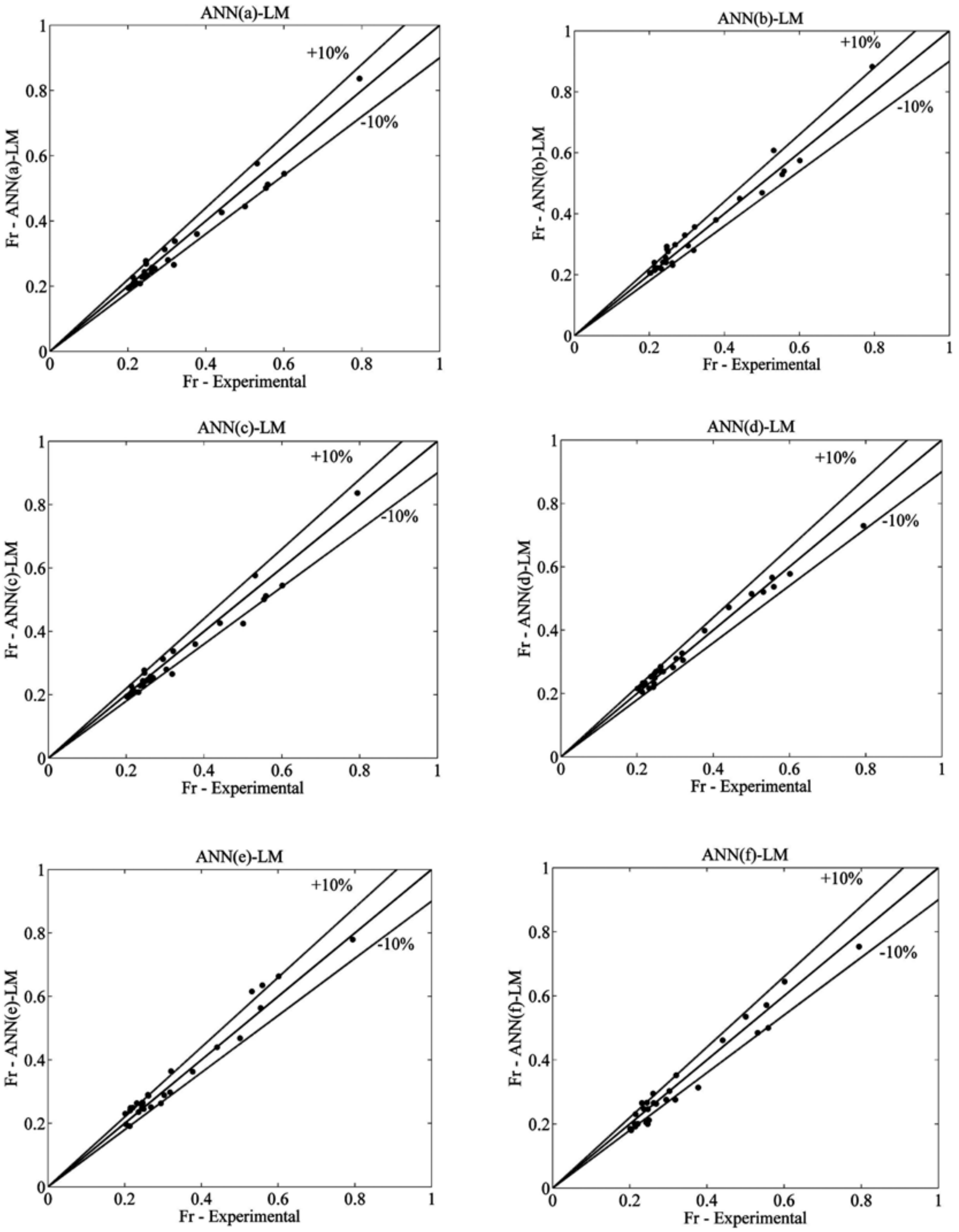

Fig. 1. Validation of the *Fr* Predicted by the Different ANN-LM Models using Observed Values (Ab Ghani's data)

ANN(b) is greater than 10%. This entails overestimated values that lead to a non-economical design due to the large error. In almost all samples, model ANN(e) predicts higher *Fr* than the observed value; therefore, it causes a condition similar to that of model ANN(b). Models ANN(c), ANN(a), and ANN(f) have mean relative errors of 6.54%, 6.54%, and 8.53% (respectively) and make better predictions than ANN(e) and ANN(b). Although their mean relative errors are greater, they bear a relative error greater than 10% at some points while ANN(d) estimates *Fr* values with a relative error less than 10% while its least mean relative error is approximately 5.13%.

Model ANN(d)-LM, which is more accurate in predicting *Fr* than other models presented in this study, is compared with existing equations in Fig. 2. Vongvisessomjai *et al.*'s experimental results (2010), which had no role in training ANN, were used to do so. The value of *Fr* predicted by model ANN(d)-LM has a relative error of less than 10% in almost all cases. The examinations carried out using the statistical indexes signify this model performed well (*RMSE* = 0.025 & *MAPE* = 5.78). The values of these indexes are almost equal to their values when we used Ab

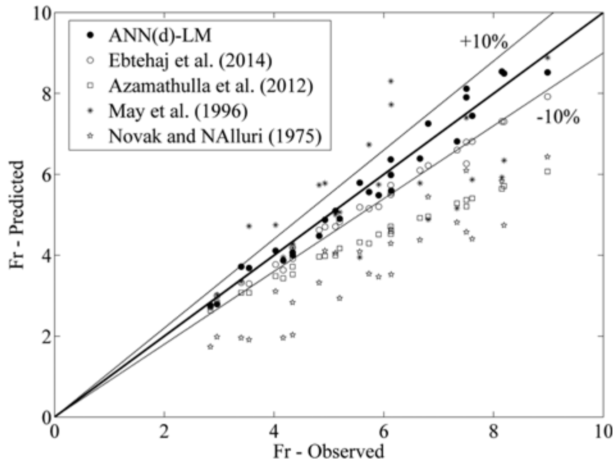

Fig. 2. Comparison of *Fr* Predicted by the ANN(d)-LM Model with Existing Sediment Transport Equations using Vongvisessomjai *et al.,* 2010's Data

Table 3. Comparison of Existing Sediment Transport Equations and ANN(d)-LM Model using Vongvisessomjai *et al.*'s (2010) Data

	RMSE	MAPE (%)
ANN(d) - LM	**0.025**	**5.78**
Novak and Nalluri (1975)	0.118	34.30
May *et al.* (1996)	0.183	13.74
Azamathulla *et al.* (2012)	0.039	21.70
Ebtehaj *et al.* (2014)	0.032	7.63

Ghani's data (Table 1). The equations presented by May *et al.* (1996) and Novak and Nalluri (1975) have low accuracy, with mean relative error of 21.7% and 34.3% (respectively) while the maximum relative error value of the ANN(d)-LM model presented in this study is nearly 10%. Azamathulla *et al.*'s (2012) equation often overestimates the *Fr*; therefore, using this equation results in sediment deposit in the sewer. Thus, considering the explanations in Fig. 2 and Table 2, it can be stated that compared with other existing methods, the model presented in this study (ANN(d)-LM) can predict *Fr* fairly accurately.

Since ANN(d)-LM had the best results, *Fr* can be calculated with the following equation:

$$Fr = purelin((tansig(input \times iw + b_1)) \times lw + b_2) \tag{21}$$

$$input = [C_V,\ d/D,\ d/R,\ \lambda_s] \tag{22}$$

$$purelin(x) = x \tag{23}$$

$$tansig(x) = \frac{2}{1+e^{-2x}} - 1 \tag{24}$$

$$iw = \begin{bmatrix} 3.03 & 3.10 & 3.36 & -0.97 & -2.82 & -0.72 & 0.42 & 1.95 & 0.89 & 0.75 \\ -0.52 & 0.71 & 7.55 & -0.87 & 0.52 & 1.02 & -1.97 & -1.21 & -1.03 & -1.72 \\ 5.06 & 1.85 & -4.47 & 3.06 & -1.51 & -1.38 & 0.02 & 3.05 & 1.11 & -0.56 \\ -0.86 & -1.77 & -0.25 & 3.52 & -1.23 & -2.79 & -0.41 & 1.30 & 3 & -1.61 \end{bmatrix} \tag{25}$$

$$lw = \begin{bmatrix} -5.21 \\ -0.16 \\ 4.29 \\ -3.59 \\ -0.09 \\ 1.07 \\ 0.22 \\ 0.38 \\ 2.12 \\ 0.96 \end{bmatrix} \tag{26}$$

$$b_1 = [7.75\ \ -3.43\ \ 7.14\ \ -2.48\ \ 0.73\ \ -2.31\ \ 0.33\ \ 0.98\ \ -2.96\ \ 3] \tag{27}$$

$$b_2 = [-0.85] \tag{28}$$

6. Conclusions

The capability of a Multilayer Perceptron (MLP) network optimized with three training algorithms, including variable learning rate (MLP-GDX), resilient back-propagation (MLP-RP) and Levenberg–Marquardt (MLP-LM) to predict sediment transport in sewer pipes is investigated. In this study, a multilayer perceptron (MLP) network optimized with three training algorithms was used to estimate sediment transport at limit of deposition. The training algorithms utilized were variable learning rate (MLP-GDX), resilient back-propagation (MLP-RP) and Levenberg–Marquardt (MLP-LM). Different models were used to examine the effect of various dimensionless parameters on predicting the *Fr*. The parameters related to the "transport" (C_V) and "flow resistance" (λ_s) groups were constant in all these models and the effects of the parameters related to the "sediment" (*Dgr*, *d*/*D*) and "transport mode" (*d*/*R*, *R*/*D*, D^2/A) groups were examined. Thus, 6 different groups were presented.

The results of the examinations carried out using different statistical indexes indicate that the ANN(d) model, which considers the influential parameters to be volumetric sediment concentration (C_V), relative depth of flow (*R*/*d*), proportion of median diameter of particles to pipe diameter (*d*/*D*) and overall sediment friction factor (λ_s) predicted the *Fr* more accurately than the other models (ANN(d)-LM). The values of the statistical indexes for this model are $MAPE = 5.13\%$ and $RMSE = 0.02$. Also, the results of comparing all three algorithms used indicate that MLP-LM mostly produced better results for all models compared with MLP-GDX and MLP-RD. Subsequently, a set of data that had no role in the current model was used from Vongvisessomjai *et al.*'s (2010) data set in order to examine the accuracy of the presented model. The results indicate there was no significant change in prediction accuracies, which shows the flexibility of this model in different data sets. Comparing the present model with existing sediment transport equations signifies that ANN(d)-LM is more accurate.

Acknowledgements

The authors would like to express their appreciation to anonymous reviewers for their helpful comments and to Maya Binder for final editing of the English text.

References

Ab Ghani, A. (1993). *Sediment transport in sewers*, PhD Thesis, University of Newcastle Upon Tyne, UK.

Ackers, J. C., Butler, D., and May, R. W. P. (1996). *Design of sewers to control sediment problems*, Report No. CIRIA 141, Construction Industry Research and Information Association, London.

Almedeij, J. (2012). "Rectangular storm sewer design under equal sediment mobility." *American Journal of Environmental Science*, Vol. 8, No. 4, pp. 376-384.

Almedeij, J. and Almohsen, N. (2010). "Remarks on Camp criterion for self-cleansing storm sewer." *Journal of Irrigation and Drainage Engineering*, Vol. 136, No. 2, pp. 145-148. DOI: 10.1061/(ASCE) IR.1943-4774.0000129.

Alp, M. H. and Cigizoglu, K. (2007). "Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data." *Environmental Modeling and Software*, Vol. 22, No. 1, pp. 2-13, DOI: 10.1016/j.envsoft.2005.09.009.

Azamathulla, H. Md., Ab Ghani, A., and Fei, S. Y. (2012). "ANFIS – based approach for predicting sediment transport in clean sewer." *Journal of Applied Soft Computing*, Vol. 12, No. 3, pp. 1227-1230, DOI: 10.1016/j.asoc.2011.12.003.

Azamathulla, H. Md., and Zakaria, N. A. (2011). "Prediction of scour below submerged pipeline crossing a river using ANN." *Water Science and Technology*, Vol. 63, No. 10, pp. 2225-2230, DOI: 10.2166/wst.2011.459.

Baghalian, S., Bonakdari, H., Nazari, F., and Fazli, M. (2012). "Closed-form solution for flow field in curved channels in comparison with experimental and numerical analyses and Artificial Neural Network." *Engineering Application of Computational Fluid Mechanics*, Vol. 6, No. 4, pp. 514-526.

Banasiak, R. (2008). "Hydraulic performance of sewer pipes with deposited sediments." *Water Science and Technology*, Vol. 57, No. 11, pp. 1743-1748, DOI: 10.2166/wst.2008.287.

Bonakdari, H., Baghalian, S., Nazari, F., and Fazli, M. (2011). "Numerical analysis and prediction of the velocity field in curved open channel using Artificial Neural Network and Genetic Algorithm." *Engineering Application of Computational Fluid Mechanics*, Vol. 5, No. 3, pp. 384-396.

Bonakdari, H. and Ebtehaj, I. (2014). "Verification of equation for non-deposition sediment transport in flood water canals." *7th International Conference on Fluvial Hydraulics*, RIVER FLOW 2014, Lausanne, Switzerland, pp. 1527-1533.

Bonakdari, H. and Larrarte, F. (2006). "Experimental and numerical investigation on self-cleansing and shear in sewers." *2nd International IWA Conference*, Sewer Operation and Maintenance, SOM 06, Vienna, Austria, pp. 19-26.

Bong, C. H. J., Lau, T. L., and Ab Ghani, A. (2013). "Verification of equations for incipient motion studies for a rigid rectangular channel." *Water Science and Technology*, Vol. 67, No. 2, pp. 395-403, DOI: 10.2166/wst.2012.580.

Bong, C. H. J., Lau, T. L., and Ab Ghani, A. (2014). "Self-cleansing design of rectangular open storm sewer." *13th International Conference on Urban Drainage, Sarawak, Malaysia*.

Butler, D., May, R. W. P., and Ackers, J. (2003). "Self-cleansing sewer design based on sediment transport principles." *Journal of Hydraulic Engineering*, Vol. 129, No. 4, pp. 276-282, DOI: 10.1061/(ASCE) 0733-9429(2003)129:4(276).

Ebtehaj, I. and Bonakdari, H. (2013). "Evaluation of sediment transport in sewer using artificial Neural Network." *Engineering Applications of Computational Fluid Mechanics*, Vol. 7, No. 3, pp. 382-392.

Ebtehaj, I. and Bonakdari, H. (2014). "Performance evaluation of adaptive neural fuzzy inference system for sediment transport in Sewers." *Water Resource Management*, Vol. 28, No. 13, pp. 4765:4779, DOI: 10.1007/s11269-014-0774-0.

Ebtehaj, I, Bonakdari, H., and Sharifi, A. (2014). "Design criteria for sediment transport in sewers based on self-cleansing concept." *Journal of Zhejiang University Science-A*, Vol. 15, No. 11, pp. 914-924, DOI: 10.1631/jzus.A1300135.

Cigizoglu, H. K. (2003). "Estimation, forecasting and extrapolation of flow data by artificial neural networks." *Hydrological Sciences Journal*, Vol. 48, No. 3, pp. 349-361, DOI: 10.1623/hysj.48.3.349.45288.

Dawson, W. C. and Wilby, R. (1998). "An artificial neural network approach to rainfall-runoff modelling." *Hydrological Sciences Journal*, Vol. 43, No. 1, pp. 47-66, DOI: 10.1080/02626669809492102.

Hornik, K., Stinchcombe, M., and White, H. (1989). "Multilayer feed forward networks are universal approximators." *Neural Networks*, Vol. 2, No. 5, pp. 359-366, DOI: 10.1016/0893-6080(89)90020-8).

Jain, S. K., Nayak, P. C., and Sudheer, K. P. (2008). "Models for estimating evapotranspiration using artificial neural networks, and their physical interpretation." *Hydrological Processes,* Vol. 22, No. 13, pp. 2225-2234, DOI: 10.1002/hyp.6819.

Jalili-Ghazi Zade, M. and Noori, R. (2008). "Prediction of municipal solid waste generation by use of artificial neural network: A case study of Mashhad." *International Journal of Environmental Research*, Vol. 2, No. 1, pp. 13-22.

Kisi, O. (2007). "Streamflow forecasting using different artificial neural network algorithms." *Journal of Hydrologic Engineering*, Vol. 12, No. 5, pp. 532-539, DOI: 10.1061/(ASCE)1084-0699(2007)12:5(532).

Macke, E. (1982). "About sediment at low concentrations in partly filled pipes." *Mitteilungen, Leichtweiss institut fur Wasserbau der technischen Universitat Braunschweig*, Heft, Vol. 71, pp. 1-151.

May, R. W. P. (1993). *Sediment transport in pipes and sewers with deposited beds*, Hydraulic Research Ltd., Wallingford, England, Report SR 320.

May, R. W. P., Ackers, J. C., Butler, D., and Johnt, S. (1996). "Development of design methodology for self-cleansing sewers." *Water Science and Technology*, Vol. 33, No. 9, pp. 195-205, DOI: 10.1016/0273-1223(96)00387-3.

May, R. W. P., Brown, P. M., Hare, G. R., and Jones, K. D. (1989). *Self-cleansing conditions for sewers carrying sediment*, Hydraulics Research Ltd., Wallingford, England, Report SR 221.

Mayerle, R., Nalluri, C., and Novak, P. (1991). "Sediment transport in rigid bed conveyance." *Journal of Hydraulic Research*, Vol. 29, No. 4, pp. 475-495, DOI: 10.1061/(ASCE)1084-0699(2007)12:5(532).

Melesse, A. M., Ahmad, S., McClain, M. E., and Lim, Y. H. (2011). "Suspended sediment load prediction of river systems: An artificial neural network approach." *Agricultural Water Management*, Vol. 98, No. 5, pp. 855-866, DOI: 10.1016/j.agwat.2010.12.012.

More, J. J. (1977). *The Levenberg—Marquardt algorithm: Implementation and theory*, Numerical Analysis. In: Watson GA (ed) Lecture notes in mathematics 630, Springer, New York, pp. 105-116.

Nagy, H., Watanabe, K., Hirano, M. (2002). "Prediction of sediment load concentration in rivers using artificial Neural Network Model." *Journal of. Hydraulic Engineering*, Vol. 128, No. 6, pp. 588-595, DOI: 10.1061/(ASCE)0733-9429(2002)128:6(588).

Nalluri, C. (1985). *Sediment transport in rigid boundary channels*, Proceeding. Euromech 192: Transport of Suspended Solids in Open channels, Neubiberg, Germany, pp. 101-104.

Nalluri, C., Ab Ghani, A., and El-Zaemey, A. K. S. (1994). "Sediment transport over deposited beds in sewers." *Water Science and Technology*, Vol. 29, Nos. 1-2, pp. 125-133.

Nalluri, C. and Kithsiri, M. M. A. U. (1992). "Extended data on sediment transport in rigid bed rectangular channel." *Journal of Hydraulic Research*, Vol. 30, No. 6, pp. 851-856. DOI: 10.1080/00221689209498914.

Novak, P. and Nalluri, C. (1975). "Sediment transport in smooth fixed bed channels." *Journal of Hydraulic Division,* ASCE, Vol. 101, pp. 1139-1154.

Noori, R., Hoshyaripour, G., Ashrafi, K., and Nadjar-Araabi, B. (2010). "Uncertainty analysis of developed ANN and ANFIS models in prediction of carbon monoxide daily concentration." *Atmospheric Environment*, Vol. 44, No. 4, pp. 476-482. DOI: 10.1016/j.atmosenv.2009.11.005.

Nour, M. H., Smith, D. W. and. Prepas, E. E. (2006). "Neural networks modelling of streamflow, phosphorus, and suspended solids: Application to the Canadian Boreal forest." *Water Science and Technology*, Vol. 53, No. 10, pp. 91-99.

Nourouzi, M. M., Chuah, T. G., and Choong, T. S. Y. (2011). "Optimisation of reactive dye removal by sequential electrocoagulation–flocculation method: Comparing ANN and RSM prediction." *Water Science and Technology*, Vol. 63, No. 5, pp. 985-995. DOI: 10.2166/wst.2011.280.

Ota, J. J. and Nalluri, C. (2003). "Urban storm sewer design: Approach in consideration of sediments." *Journal of Hydraulic Engineering*, Vol. 129, No. 4, pp. 291-297, DOI: 10.1061/(ASCE)0733-9429(2003)129:4(291).

Ota J. J. and Perrusquia, G. S. (2013). "Particle velocity and sediment transport at the limit of deposition in sewers." *Water Science and Technology*, Vol. 67, No. 5, pp. 959-967, DOI: 10.2166/wst.2013.646.

Partal, T. (2009). "River flow forecasting using different artificial neural network algorithms and wavelet transform." *Canadian Journal of Civil Engineering*, Vol. 36, No. 1, pp. 26-38. DOI: 10.1139/L08-090.

Rezaeian Zadeh, M., Amin, S., Khalili, D., and Singh, V. P. (2010). "Daily outflow prediction by multilayer perceptron with logistic sigmoid and tangent sigmoid activation functions." *Water Resource Management*, Vol. 24, No. 11, pp. 2673-2688, DOI: 10.1007/s11269-009-9573-4.

Riedmiller, M. and Braun, H. (1993). "A direct adaptive method for faster back propagation learning: The RPROP algorithm." *Proceedings of IEEE international joint conference on neural networks*, San Francisco, pp. 586-591.

Sarangi, A. and Bhattacharya, A. K. (2005). "Comparison of Artificial Neural Network and regression models for sediment loss prediction from Banha watershed in India." *Agricultural Water Management*, Vol. 78, pp. 195-208, DOI: 10.1016/j.agwat.2005.02.001.

Schulz, F. H., Wolf, H., Jansen, H. W., and Van der Veer, P. (2005). "Applications of artificial neural networks in integrated water management: fiction or future?." *Water Science and Technology*, Vol. 52, No. 9, pp. 21-31.

Shahin, M. A., Jaksa, M. B., and Maier, H. R. (2008). "State of the art of artificial neural networks in geotechnical engineering." *Electronic Journal of Geotechnical Engineering,* Vol. 8, pp. 1-26.

Sudheer, K. P., and Jain, A. (2004). "Explaining the internal behavior of artificial neural network river flow models." *Hydrological Processes*, Vol. 18, No. 4, pp. 833-844, DOI: 10.1002/hyp.5517.

Vongvisessomjai, N., Tingsanchali, T., and Babel, M. S. (2010). "Non-deposition design criteria for sewers with part-full flow." *Urban Water Journal*, Vol. 7, No. 1, pp. 61-77, DOI: 10.1080/15730620903242824.

Yonaba, H., Anctil, F., and Fortin, V. (2010). "Comparing sigmoid transfer functions for neural network multistep ahead streamflow forecasting." *Journal of Hydrologic Engineering,* ASCE, Vol. 15, No. 4, pp. 275-283, DOI: 10.1061/(ASCE)HE.1943-5584.0000188.