

PAPER • OPEN ACCESS

Predicting Sediment transport in sewers using integrative harmony search-ANN model and factor analysis

To cite this article: Mohammad Zounemat-Kermani *et al* 2020 *IOP Conf. Ser.: Earth Environ. Sci.* **491** 012004

Recent citations

- [Estimating hourly PM2.5 concentrations using Himawari-8 AOD and a DBSCAN-modified deep learning model over the YRDUA, China](#)
Xiaoman Lu *et al*

View the [article online](#) for updates and enhancements.



240th ECS Meeting ORLANDO, FL

Orange County Convention Center **Oct 10-14, 2021**



Abstract submission due: April 9

SUBMIT NOW

Predicting Sediment transport in sewers using integrative harmony search-ANN model and factor analysis

Mohammad Zounemat-Kermani^{1,*}, Marzieh Fadaee¹, S Adarsh² and Reinhard Hinkelmann³

¹Department of Water Engineering, Shahid Bahonar University of Kerman, Kerman, Iran

²TKM College of Engineering, Kollam 691005, India

³Chair of Water Resources Management and Modeling of Hydrosystems, Institute of Civil Engineering, Technische Universität Berlin, Berlin, Germany

*Corresponding author: e-mail: zounemat@uk.ac.ir

Abstract. This study evaluates the performance of an integrated version of artificial neural network namely HS-ANN (which is a combination of neural network and heuristic harmony search algorithm) as an alternative approach to predict the sediment transport in terms of sediment volumetric concentration (C_v) in sewer pipe systems. To overcome the complexities of choosing the optimum number of the input variables and to consider the effective parameters of the model, the factor analysis technique is utilized. In addition to the HS-ANN model, an empirical equation, as well as a multiple linear regression model, are also considered. The mean square error ($RMSE$), mean absolute percentage error ($MAPE$), and Pearson correlation coefficients (PCC) are used for evaluating the accuracy of the applied models. As the comparisons demonstrate, the HS-ANN model ($PCC = 0.97$) is more accurate than the existing empirical equation and MLR model and could be successfully employed in predicting sediment transport in sewer networks.

Keywords: Sewer systems, sediment transport, water engineering

1. Introduction

Sediment transport and movement in sewer networks have been the topics of several studies in recent years due to some pertinent concerns such as pollution to watercourses, blockage and surcharging. As a matter of fact, the deposition of solids occurs occasionally in sewers as a result of the intermittent nature of flow [1, 2]. Several studies have attempted to model the transport of sediments in sewer network using experimental studies and related empirical relations [1, 3] and numerical methods [4, 5].

In the past decades, due to the reliability and capability of soft computing methods in analyzing and modelling complex problems, the popularity and application of these methods have been increased in various fields of science, especially in water resources engineering, hydraulics and hydrology [6, 7].

Few previous works on sediment transport using soft computing techniques have been reported [2, 8-10]. Azamathulla *et al.* [2] presented the adaptive neuro-fuzzy inference system (ANFIS) as an alternative approach to predict the functional relationships of sediment transport in sewer pipe systems.



The applied ANFIS approach gave satisfactory results in comparison to the existing empirical relations and multiple linear regression (MLR) model. Ebtehaj et al. [10] challenged the capability of the hybrid support vector machine (SVM) and wavelet transform model for the prediction of the densimetric Froude number (Fr) as a sediment transport index in sewer networks. They claimed that both hybrid and standard SVM give better results than the conventional relations, however, the hybrid SVM-wavelet model performed better than the standard SVM model.

In this study, it is aimed to model the sediment transport in sewers by developing an integrative model combined with the ANN model and harmony search algorithm, so called HS-ANN. Also, the factor analysis is conducted on the entire datasets to understand the most important input variables in prediction of volumetric sediment concentration (C_v). Finally, the performance of applied methods are evaluated using different statistical criteria.

1.1 Datasets

In the present study, pertinent datasets were gathered from the data reported by Ghani [1] for modeling sediment transport in sewers at the Hydraulic Laboratories of the University of Newcastle, UK. In the study, all experiments were carried out under part-full uniform flow conditions. Two pipes of 154 mm and 305 mm diameter were utilized in which sediments transported as bed load. The sediments used were uniformly graded and non-cohesive ($d_{30} = 0.5 \text{ mm} - 10.0 \text{ mm}$). Figure 1 schematically depicts the experimental sewer pipes.

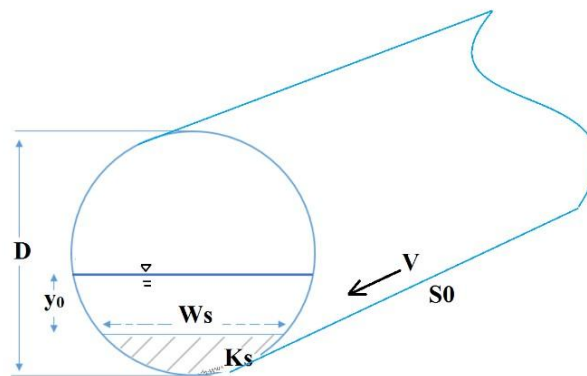


Figure 1. Schematic view of the overall geometry of a sewer pipe with deposited beds

A total number of 195 datasets were utilized for modeling the sediment transport process in sewers. Each dataset consists of 17 independent variables (input variables) and one predictable variable (target value) of the C_v . The potential input variables included median diameter of particles in a mixture (d_{50}), flow discharge (Q), mean velocity of flow (V), depth of uniform flow (y_0), internal diameter of pipe channel (D), flow Reynolds number (Re), flow Froude number (Fr), bed slope (S_0), overall friction factor with sediment (λ_s), overall equivalent sand roughness with sediment (k_s), overall Manning roughness coefficient with sediment (n_s), width of sediment spread (w_s) in pipes, temperature (T), cross-sectional area of the flow (A), wetted parameter of the flow (P), overall hydraulic radius (R) and water surface width (B). The summary of the statistical characteristics of potential factors on sediment transport in sewers is given in Table 1. From Table 1, it can be seen that the C_v factor is mostly correlated with the Fr number ($r = 0.69$) and bed slope ($r = 0.68$). On the other hand, C_v is in disagreement with the geometric parameters of the hydraulic radius of the pipe ($r = -0.55$).

Table 1. Statistical characteristics of the parameters considered in this study.

Parameters	Minimum	Maximum	Average	Standard deviation	Coefficient of variation	Correlation (r) with C_v
Q (l/s)	0.44	34.63	16.87	9.39	0.56	-0.37
V (m/s)	0.24	1.21	0.69	0.17	0.25	0.43
y_0/D	0.15	0.77	0.42	0.15	0.36	-0.45
Re	12715	276449	147704	54569	0.37	-0.18
Fr	0.3	1.47	0.8	0.27	0.34	0.69
T ($^{\circ}C$)	11	21.5	17.22	2.65	0.15	0.13
A (m^2)	0.002	0.06	0.02	0.01	0.6	-0.55
P (m)	0.05	0.65	0.38	0.13	0.69	-0.49
R (m)	0.01	0.09	0.06	0.02	0.34	-0.55
B (m)	0.11	0.31	0.26	0.06	0.23	-0.36
d_{50} (mm)	0.46	8.3	4	2.45	0.61	0.34
λ_s	0.01	0.05	0.02	0.01	0.26	0.09
k_s (mm)	-0.12	1.87	0.51	0.51	0.99	-0.16
w_s (mm)	9	80	33.94	15.37	0.45	0.44
n_s	0.01	0.01	0.01	0	0.13	-0.19
S_0	0	0.01	0	0	0.55	0.68
D (mm)	154	305	274.64	60.52	0.22	-0.28
C_v (ppm)	0.76	1450	283.27	339.11	1.2	1

2. Methods

2.1 Artificial neural network

Artificial neural networks (ANNs) are inspired by the biological nervous system of animals which are mathematically composed of some interconnected neurons. Due to their nonlinear structure by using activation functions, interconnection weights, and biases, trained ANNs can provide outputs by processing the input data. The training process of ANNs (which is normally a back-propagation process) is done using the optimization algorithm. Conventional ANNs used mathematical methods such as gradient descent algorithm for the optimization process. However, it is possible to get the advantage of heuristic methods (such as particle swarm optimization algorithm, firefly algorithm and etc.) for optimizing the ANN parameters. In this study, the harmony search algorithm is integrated with the multi-layer perceptron neural network for training the ANN [11, 12].

2.2 Harmony search algorithm

The harmony search (HS) heuristic algorithm was introduced by Geem et al. [13]. This algorithm mimics the improvisation of music players. In general, a music player would improvise the pitches of her/his instrument to obtain better harmony. Inspired by this analogy, the HS algorithm has been developed for seeking the optimum solution to a problem. For this, five steps are considered in this algorithm. Firstly the algorithm parameters such as the number of decision variables and solution vectors are initialized. Secondly, the harmony memory (a memory where all the solution vectors are stored) is initialized based on the predetermined algorithm parameters in the first step. In the third step, a new harmony is improvised normally on the basis of the three rules of memory consideration, pitch adjustment, and random selection.

In the fourth step, the harmony memory is updated. The computation is terminated if the stopping criterion is satisfied in the fifth step [14]. This algorithm is used in this study for optimizing the weights and biases of the multi-layer perceptron neural network.

2.3 Multiple linear regression

A multiple linear regression (MLR) analysis is used to predict the values of a dependent variable (C_v in this study), given a set of explanatory variables. Generally, the least-squares criterion is considered to estimate the MLR equation, so that the sum of squares of the differences between the observed and predicted values for each observation will be minimized.

2.4 Empirical relation

May et al. [15] have studied the parameters influencing the prediction of sediment concentration in sewers and presented the following equation to determine it.

$$C_v = 0.0211 \left(\frac{y_0}{D} \right)^{0.36} \left(\frac{D^2}{A} \right) \left(\frac{d}{R} \right)^{0.6} \left(\frac{V^2}{g(s-1)D} \right)^{3/2} \left(1 - \frac{V_c}{V} \right)^4 \quad (1)$$

$$V_c = 0.125(g(s-1)d)^{1/2} (y/d)^{0.47} \quad (2)$$

where C_v is volumetric sediment concentration, y is depth of uniform flow, D is internal diameter of pipe channel, A is cross-sectional area of the flow, d is diameter of particles, R is overall hydraulic radius, V is mean velocity of flow, g is the gravitational acceleration, s is the value for particle density, and V_c is the critical incipient motion velocity [16].

3. Application and results

3.1 Selection of input parameters (factor analysis)

As it was mentioned earlier, 17 input parameters can be considered for modeling and predicting C_v . Most researchers considered less number of independent variables such as the effective factors in sediment transport as flow velocity, the size of particles, the hydraulic radius, the average size of sediment, and the total friction factor [8]. In this study in order to find out which are the most important input variables in predicting C_v , the factor analysis has been conducted on the entire datasets.

Figure 2a shows the results of the scree plot for all parameters. It can be observed from the dashed line in Figure 2a, that at least seven parameters should be considered for modeling. In this respect, we consider selecting seven parameters out of 17 independent variables for predicting C_v . Hence, as plotted in the loading plot (Figure 2b), seven groups of parameters have been spotted and the seven variables S_0 , Fr , V , λ_s , y_0/D , R , and Re have been chosen as the input vector.

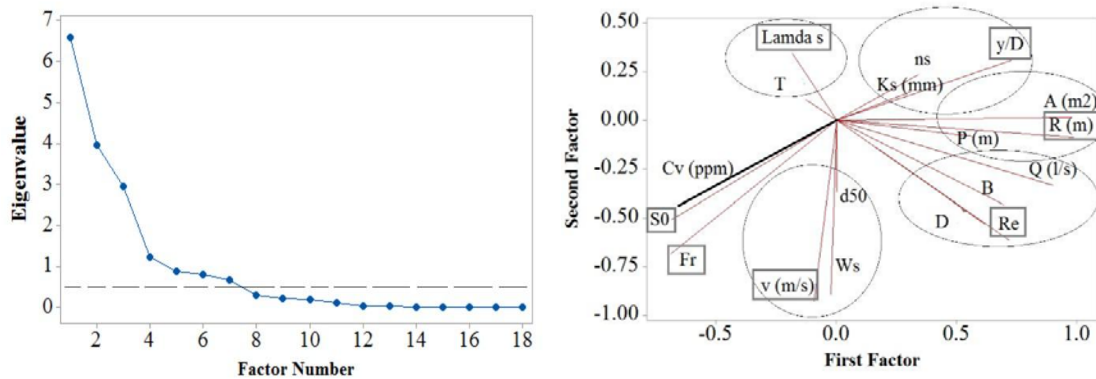


Figure 2. a) Scree plot and b) loading plots for the whole data set using factor analysis

3.2 Models set up

Several feed-forward back-propagation multi-layer perceptron neural networks with one hidden layer were developed for modeling sediment transport. The optimum number of hidden nodes in the hidden layer was achieved by trial and error procedure (the optimum structure of the final HS-ANN is tabulated in Table 2). The harmony search algorithm was embedded with the standard version of ANN for the optimization process. To train and validate the HS-ANN model using the sets of data, 85% of tests were randomly selected from 195 available data. The same training set was also chosen for constructing the multiple linear regression model. Moreover, to evaluate the efficiency of the presented HS-ANN, MLR and May et al. [15] empirical model, the remaining 15% data were used to test the accuracy.

3.3 Results and discussion

Table 2 presents the findings of different statistical indexes of the Pearson correlation coefficient (*PCC*), root mean square error (*RMSE*) and mean absolute percentage error for the *Cv* predictions using the HS-ANN and MLR models and the empirical equation. As for the MLR model, the low values of the correlation coefficient ($PCC < 0.7$) and greater values of *RMSE* and *MAPE* indicates that this MLR model could not predict the sediment transport in sewers. On the other hand, the complex phenomenon and nonlinear behavior of sediment transport in sewers cannot be captured by linear regression models.

It can be noticed that the predictions of both HS-ANN and May et al. [15] empirical equation are highly favorable so that the values of *PCC* exceeded 0.9 in the testing phase. The HS-ANN model outperformed the empirical equation in both the training and testing phases.

Table 2. Validation of the applied models in terms of statistical measures.

Phase	Model	PCC	RMSE (<i>ppm</i>)	MAPE (%)
Training	HS-ANN (7,9,1)	0.93	124.64	34.82
	MLR	0.77	218.87	179.74
	May et al. 1989	0.82	200.28	35.13
Testing	HS-ANN (7,9,1)	0.97	88.73	19.10
	MLR	0.59	284.41	359.58
	May et al. 1989	0.91	156.28	24.54

Figure 3 provides scatter plots of the results obtained from the predictions made by the HS-ANN for both the training and testing phases. It could be seen that the HS-ANN model has been fairly trained well in such a manner that most of the simulated and predicted values are in the 20% bands from the line of agreement. Figure 4 also shows similar scatter plots for the MLR model and May et al. [15] empirical equation. The results of these plots indicate that both the MLR model and empirical equation tended to under-predict higher C_v values. Besides, a general over-prediction at many data points can be seen in the MLR model. In general, it can be observed that none of these methods perform as well as the HS-ANN since the results are more scattered. However, the empirical equation could predict sediment transportation better than the MLR model.

Figure 5 depicts the distribution of the residuals of each model using dot-plot. It is obvious that the residuals related to the HS-ANN are not scattered, while the residuals of the MLR model are more distributed alongside the X-axis which is an indication of considerable difference between the observed C_v values and MLR model results. In addition, it can be concluded from Figures 3, 4 and 5 that the empirical equation and especially the MLR model tend to under-predict the C_v values, whereas the HS-ANN model is reluctant in terms of over/under-prediction issue which is a beneficial factor for the capability of HS-ANN model.

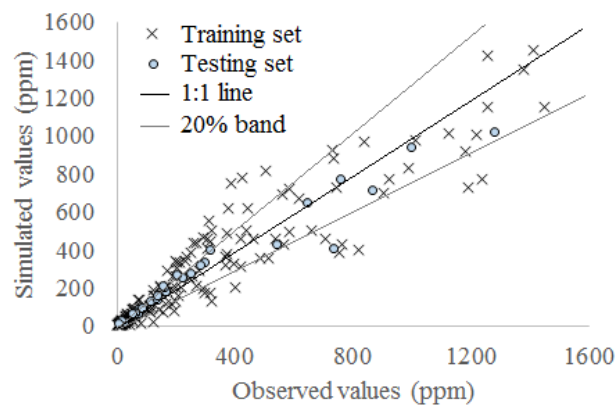


Figure 3. Scatter plots for the modeled/predicted values of C_v for the HS-ANN model

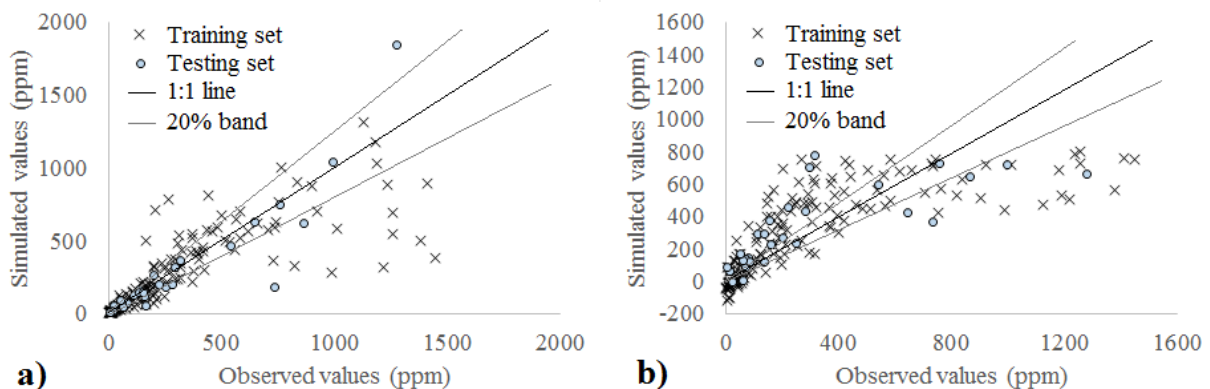


Figure 4. Scatter plots for the modeled/predicted values of C_v for the a) empirical equation, b) MLR model

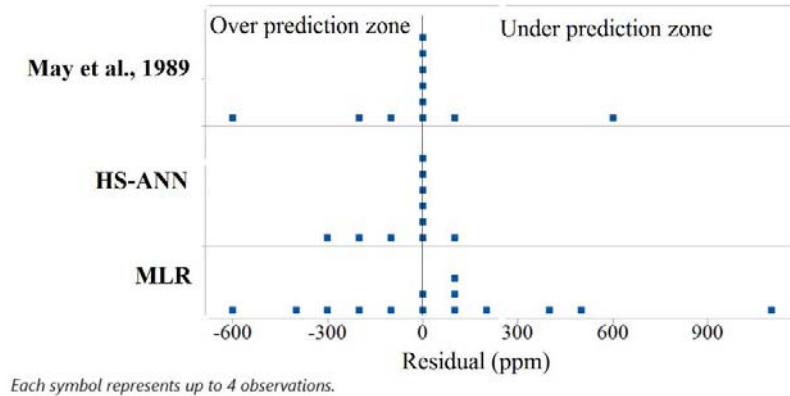


Figure 5. Dot-plot representation for the applied models' residuals

4. Conclusion

Sediment transport in sewer networks is known as a complex phenomenon that cannot be easily modeled. In this study, an integrative soft computing technique integrated from the harmony search algorithm and multi-layer perceptron was proposed for predicting the volumetric concentration of sediment in sewer pipes. Besides the HS-ANN model, an empirical equation suggested by May et al. [15] as well as a multiple linear regression model was applied. It was found that the HS-ANN model with the highest values of the Pearson correlation coefficient ($PCC = 0.93$ in the training set and 0.97 in the testing phase) and lowest values for the $RMSE$ and $MAPE$ showed better performance over the empirical and MLR model. Although the empirical equation modeled and predicted the C_v values properly ($PCC = 0.82$ in the training set and 0.91 in the testing phase), the MLR model failed to predict the C_v values.

Acknowledgment

The first author would like to acknowledge the Georg Forster Fellowship for Experienced Researcher of the Alexander von Humboldt Foundation for accomplishing this research.

References

- [1] Ghani A A 1993 *Sediment transport in sewers*. (Civil Engineering Department, Newcastle University, PhD thesis)
- [2] Azamathulla H M, Ghani A A and Fei S Y 2012 ANFIS-based approach for predicting sediment transport in clean sewer *Appl. Soft Comput.* **12** 1227-30
- [3] Seco R I 2014 *In-sewer organic sediment transport: study of the release of sediments during wet-weather from combined sewer systems in the Mediterranean region in Spain* (Doctoral dissertation, Universitat Politècnica de Catalunya)
- [4] Schlütter F 1999 *Numerical modelling of sediment transport in combined sewer systems* (Doctoral dissertation, The Hydraulics and Coastal Engineering Group, Dept. of Civil Engineering, Aalborg University)
- [5] Creaco E and Bertrand-Krajewski J L 2009 Numerical simulation of flushing effect on sewer sediments and comparison of four sediment transport formulas *J. Hydraul. Res.* **47** 195-202

- [6] Rajaei T, Nourani V, Zounemat-Kermani M and Kisi O 2010 River suspended sediment load prediction: application of ANN and wavelet conjunction model *J. Hydrol. Eng.* **16** 613-27
- [7] Zounemat-Kermani M and Mahdavi-Meymand A 2018 Hybrid meta-heuristics artificial intelligence models in simulating discharge passing the piano key weirs *J. Hydrol.* **569** 12-21
- [8] Ebtehaj I and Bonakdari H 2013 Evaluation of sediment transport in sewer using artificial neural network *Eng. Appl. Comp. Fluid* **7** 382-92
- [9] Ebtehaj I and Bonakdari H 2014 Performance evaluation of adaptive neural fuzzy inference system for sediment transport in sewers *Water Resour. Manag.* **28** 4765-79
- [10] Ebtehaj I, Bonakdari H, Shamshirband S and Mohammadi K 2016 A combined support vector machine-wavelet transform model for prediction of sediment transport in sewer *Flow Meas. Instrum.* **47** 19-27
- [11] Neaupane K M and Adhikari N R 2006 Prediction of tunneling-induced ground movement with the multi-layer perceptron *Tunn. Undergr. Sp. Tech.* **21** 151-9
- [12] Zounemat-Kermani M, Kisi O and Rajaei T 2013 Performance of radial basis and LM-feed forward artificial neural networks for predicting daily watershed runoff *Appl. Soft Comput.* **13** 4633-44
- [13] Geem Z W, Kim J H and Loganathan G V 2001 A new heuristic optimization algorithm: harmony search *Simulat.* **76** 60-8
- [14] Mahdavi M, Fesanghary M and Damangir E 2007 An improved harmony search algorithm for solving optimization problems *Appl. Math. Comput.* **188** 1567-79
- [15] May R W P, Brown P M, Hare G R and Jones K D 1989 Self-cleansing condition for sewers carrying sediment *Hydraul. Res. Ltd (Wallingford)* Report SR 221
- [16] Bonakdari H and Ebtehaj I 2015 Comparison of two data-driven approaches in estimation of sediment transport in sewer pipe *E-proceedings of the 36th IAHR World Congress* 1-11