



Nature-inspired algorithms in sanitary engineering: modelling sediment transport in sewer pipes

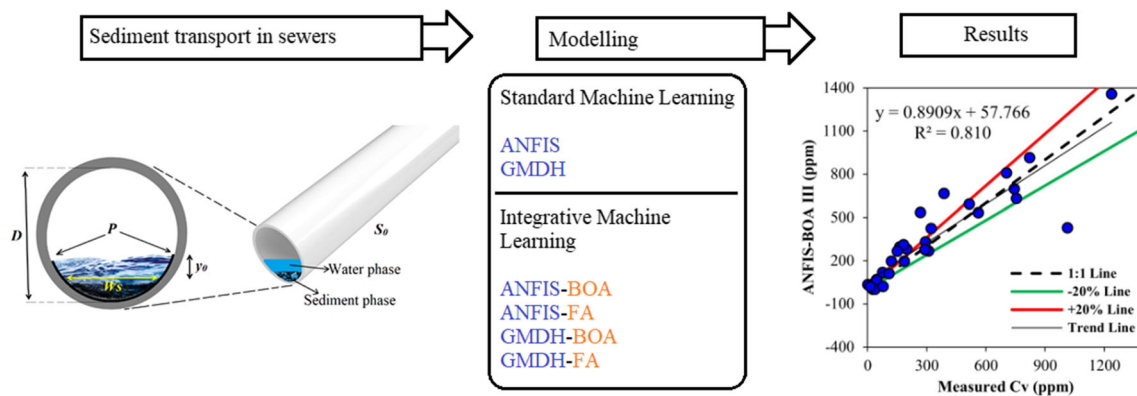
Mohammad Zounemat-Kermani¹ · Amin Mahdavi-Meymand² · Reinhard Hinkelmann³

Accepted: 26 January 2021 / Published online: 15 February 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH, DE part of Springer Nature 2021

Abstract

In this study, two nature-inspired optimization techniques such as firefly algorithm (FA) and butterfly optimization algorithm (BOA) are combined with adaptive neuro-fuzzy inference system (ANFIS) and group method of data handling (GMDH) models for optimal prediction of the complex phenomenon of volumetric concentration of sediment (C_v) in sewer systems. Three different scenarios based on the methods of dimensional analysis and forward selection are implemented for determining the input structure of ANFIS, GMDH, and regression models (multiple linear regression, MLR; stepwise regression; SR) regarding 13 independent hydraulic and geometric input variables. Several statistic criteria including the root-mean-square error (RMSE), mean absolute error (MAE), Nash–Sutcliffe efficiency (NSE), index of agreement (AI), coefficient of determination (R^2), and comprehensive synthesis index (SI) as well as Taylor diagram were used to further quantify simulating and predicting accuracies. In comparison with the regression models and two empirical equations, the results obtained by standard machine learning models (ANFIS and GMDH) were very promising. However, such integration of FA and BOA noticeably improved the performance of ANFIS (around 7% improvement in RMSE criterion) and slightly optimized the performance of GMDH (less than 1% improvement in RMSE criterion) in modelling the process of C_v prediction.

Graphic abstract



✉ Mohammad Zounemat-Kermani
zounemat@uk.ac.ir

Amin Mahdavi-Meymand
amin.mahdavi1990@gmail.com

Reinhard Hinkelmann
reinhard.hinkelmann@wahyd.tu-berlin.de

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

² Institute of Hydro-Engineering, Polish Academy of Sciences, Warsaw, Poland

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

Keywords Swarm intelligence · Heuristic algorithms · Soft computing · Hydraulics of sewers · Data mining

1 Introduction

1.1 Background and statement of the problem

Sediment transport in sewers and sewage networks have been the topics of a few studies in recent years due to some concerns associated with the pollution to watercourses, blockage, and surcharging (DeSilva et al. 2011). In general, the sediment deposition occurs occasionally in sewers as a result of the intermittent nature of flow (Ghani 1993; Azamathulla et al. 2012). Several researchers have conducted experimental studies to simulate and characterize the sediment transport in sewers and consequently presented empirical equations for modelling factors related to the sediment transport processes (May et al. 1989; Vongvisessomjai et al. 2010; Hofer et al. 2018). These equations can be taken as criteria for designing and modelling the sediment transport process in sewers. Although the extracted empirical equations are simple to use, most existing empirical equations based on experimental observations cannot properly characterize different parameters on the sediment and bedload transport (Najafzadeh et al. 2017; Safari et al. 2018). In addition, the sediment transport process in sewers is a three-dimensional multi-phase flow which makes it a highly complicated phenomenon. In consequence, many aspects of the sediment transport phenomenon in sewers cannot be captured by empirical equations.

1.2 Literature review and research hypotheses

In the past decades, in the light of the reliability and capability of machine learning models in analysing and modelling nonlinear problems, the approval and application of these models have been increased in various fields of science, especially in the hydraulics of sediment transport in aquatic areas (Rajaei et al. 2010, 2020; Zounemat-Kermani 2017). Accordingly, recently some researchers have utilized machine learning approaches in modelling sediment transport in pipes and sewers.

Ghani and Azamathulla (2010) presented gene expression programming (GEP) for modelling the functional relationships of sediment transport with partial flows in sewer pipe systems. The functional GEP relation gave satisfactory results compared to classical regression analysis. Azamathulla et al. (2012) presented an adaptive neuro-fuzzy inference system (ANFIS) to predict the functional relationships of sediment transport in sewers. It was commented that the ANFIS approach provided

satisfactory results compared to the multiple linear regression (MLR) model and existing empirical relations. Ebtehaj and Bonakdari (2013) applied an artificial neural network (ANN) in predicting sediment transport in self-cleansing sewer systems. In comparison with existing empirical methods, the findings of the study resulted in the superiority of ANNs over the traditional methods. Ebtehaj et al. (2016) investigated the potential of the wavelet transform model and hybrid support vector machine (SVM) for the prediction of the densimetric Froude number (Fr_d) in sewer networks. The findings showed that both hybrid and standard SVM models gave more accurate predictions than the conventional relations. Najafzadeh et al. (2017) applied two approaches of a model tree and evolutionary polynomial regression to simulate the critical velocity of sediment deposition in sewers. They used four independent parameters (volumetric concentration, total friction factor, the ratio of the hydraulic depth of flow to pipe diameter, and non-dimensional size of particles) for predicting the target variable. It was reported that the proposed machine learning models outperformed the benchmark formulations from the literature from the accuracy point of view.

Mahdavi-Meymand and Zounemat-Kermani (2020) used the firefly algorithm (FA) as an optimization approach to optimize GMDH parameters and introduced GMDH-FA and applied this method to simulate spillways aerators air demand. The results indicated that FA increases the performance of GMDH. Accordant with the above-mentioned researches, it is hypothesized that combining novel swarm intelligence algorithms (such as butterfly optimization algorithm by Arora and Singh (2019)) and robust nonlinear machine learning models (like adaptive neuro-fuzzy inference systems) would give efficient results in predicting complex engineering problems like sediment transport in sewers.

1.3 Research objectives, contribution, and scope of the paper

This study presents two standard and four combined machine learning approaches for predicting volumetric sediment concentration (C_v) in sewers. To meet this end, at first, two prevailing machine learning models of adaptive neuro-fuzzy inference systems (ANFIS) and group method of data handling (GMDH) were utilized as standard models. Following that, two swarm intelligence heuristic optimization techniques such as firefly algorithm (FA) and butterfly optimization algorithm (BOA) were embedded into the standard ANFIS and GMDH machine learning

models (ANFIS-FA, ANFIS-BOA, GMDH-FA, and GMDH-BOA). The BOA is a new optimization algorithm proposed by Arora and Singh (2019). They used 30 benchmark functions and three engineering problems to analyse the BOA performance. The results indicated that the BOA performance is better than the other well-known algorithms (e.g. particle swarm optimization (PSO) and genetic algorithm (GA)). So, in this research, the BOA was selected to optimize the ANFIS and GMDH parameters. The FA, as another well-known and capable heuristic algorithm, was also selected to compare and challenge the results of the BOA. It is worth mentioning that the application of the FA in many engineering optimization problems—especially as a hybrid method with ANFIS—has been confirmed in many pieces of research (Yaseen et al. 2017; Sihag et al. 2019; Roy et al. 2020).

For each model, three modelling scenarios based on two-dimensional input vectors (taking into account all the effective variables as well as forward selection method) and one non-dimensional input vector (using dimensional analysis technique) were put into practice. Afterwards, the efficiency of FA and BOA was evaluated in comparison with the standard ANFIS and GMDH models as well as two empirical equations, multiple linear regression (MLR) and stepwise regression (SR) models.

On the basis of the methodology used, the major contribution of this study lies in the general and comprehensive evaluation of FA and BOA heuristic algorithms and their reliability and capability in modelling complex problems in engineering. To the best knowledge of the authors, no similar study has ever reported the combination of BOA with ANFIS (ANFIS-BOA) and both FA and BOA with GMDH model (GMDH-FA and GMDH-BOA). In other words, this paper presented a novel application of ANFIS-BOA, GMDH-FA, and GMDH-BOA for the first time.

The remainder of the paper is categorized as follows. The next section will express the methods employed in this study. Following that, the application of the machine learning models constructed on three input scenarios will be explained. The results of the standard and combined machine learning models will be assessed and will also be compared with existing sediment transport equations and regression models (MLR and SR) in Sect. 4. In Sect. 5, the performance of the employed heuristic methods (FA and BOA) will be evaluated. Eventually, the principal findings of this research will be summarized in Sect. 6.

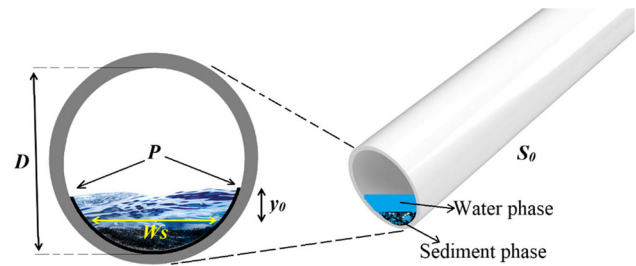


Fig. 1 Schematic view of the overall geometry of a sewer pipe with deposited beds; D : internal diameter of the pipe channel, P : the wetted perimeter of the flow, W_s : width of sediment spread, y_0 : depth of uniform flow, S_0 : the longitudinal slope of the pipe

2 Materials and methods

2.1 Data sets

In the present study, a data set was collected from the data reported by Ghani (1993) for modelling sediment transport in sewers at the Hydraulic Laboratories of the University of Newcastle, UK. In the study conducted, all experiments were done under part–full uniform flow conditions. Two pipes of 154 mm and 305 mm diameter were employed to study the bedload sediment transportation. The particles used were uniformly graded and non-cohesive ($d_{30} = 0.5\text{--}10.0$ mm). Figure 1 illustrates schematically the experimental sewer pipes.

A total of 195 data sets were utilized for modelling the sediment transport process in sewers. Each data set consisted of several independent variables (see Table 1) and one predictable variable (target value) of volumetric sediment concentration (C_v). The potential input variables included median diameter of particles in a mixture (d), flow discharge (Q), the mean velocity of flow (V), depth of uniform flow (y_0), the internal diameter of pipe channel (D), flow Froude number (Fr), the longitudinal slope of sewer (S_0), overall friction factor (λ_s), overall equivalent sand roughness with sediment (K_s), overall Manning's roughness coefficient with sediment (n), the width of sediment spread (W_s) in pipes, ambient temperature (T), cross-sectional area of the flow (A), the wetted perimeter 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 which, 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
Fr	0.30	1.47	0.80	0.27	0.34	0.69
T (°C)	11.00	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
B (m)	0.11	0.31	0.26	0.06	0.23	- 0.36
d (mm)	0.46	8.30	4.00	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)	9.00	80.00	33.94	15.37	0.45	0.44
n (s/m ^{1/3})	0.01	0.01	0.01	0.00	0.13	- 0.19
S_0	0.00	0.01	0.00	0.00	0.55	0.68
D (mm)	154	305	274.64	60.52	0.22	- 0.28
C_v (ppm)	0.76	1450	283.3	339.1	1.2	1.00

2.2 Dimensional analysis and empirical equations

As it was mentioned earlier, several independent factors affect the volumetric sediment concentration in sewers. However, not all the independent variables will have a significant effect on the result (Azamathulla et al. 2012). Hence, a sensitivity test of forward selection and dimensionless analysis were implemented to investigate the effect of each dimensionless parameter on C_v (May et al. 1989). From the dimensional analysis using Buckingham Π Theorem, the function for volumetric concentration can be obtained. Based on the data available, the values of C_v can be expressed as a function of the following parameters:

$$C_v = f\left(S_0, \frac{y_0}{D}, \frac{D^2}{A}, \frac{d}{R}, Fr_d\right), Fr_d = \frac{V^2}{g(G_s - 1)D} \quad (1)$$

where G_s is the specific gravity of sediment, Fr_d stands for the densimetric Froude number, and g denotes the gravitational constant. Regarding the dimensional analysis and Eq. (1), five dimensionless parameters will be taken into account for predicting the C_v values. In addition to the predictive models, two well-known nonlinear regression equations presented by May et al. (1989) and May et al. (1996) are also considered for better evaluation of the machine learning and regression performances. May et al. (1989) presented Eq. (2) for estimating the values for volumetric sediment concentration (C_v) in sewers:

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

where V_i denotes the critical incipient motion velocity of the sediment. In a later study, May et al. (1996) introduced the following equations based on different data sets of experimental laboratory sets for volumetric bedload transport:

$$C_v = 0.0303 \left(\frac{D^2}{A}\right) \left(\frac{d}{D}\right)^{0.6} (Fr_d)^{3/2} \left(1 - \frac{V_i}{V}\right)^4 \quad (3)$$

$$V_i = 0.125(g(S_0 - 1)d)^{0.5} (y_0/d)^{0.47}. \quad (4)$$

The present study implemented both mentioned empirical equations (Eqs. 2 and 3) for simulating C_v values.

2.3 Multiple linear and stepwise regression models

Regression models such as multiple linear regression (MLR) and stepwise regression (SR) models can be established to estimate the level of correlation between the independent variables and target value and explore the forms of relationships between them (Zounemat-Kermani 2012). MLR forms a relationship taking into account all the individual independent data points with the target value (C_v). Here, thirteen individual dimensional hydraulic and geometric predictor parameters were used for generating

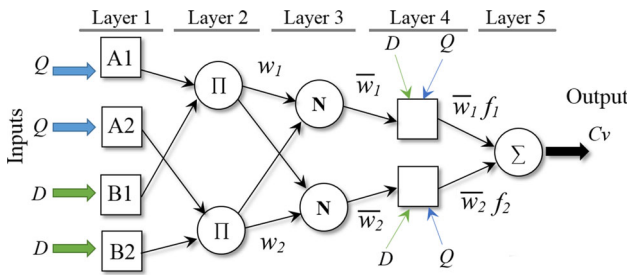


Fig. 2 Scheme of an ANFIS with two input parameters (Q and D) and two fuzzy rules

the general form of a multiple linear regression model as follows:

$$Cv = a_0 + a_1Q + a_2D + a_3S_0 + a_4\lambda s + a_5Ks + a_6n + a_7T + a_8Y + a_9d + a_{10}A + a_{11}P + a_{12}B + a_{13}Ws \quad (5)$$

where a_i are partial regression coefficients. In the stepwise regression, the selection procedure for recognizing significant input variables is automatically performed. By applying the forward selection procedure, the following stepwise regression (SR) model was employed for simulating Cv :

$$Cv = a_0 + a_1Q + a_2S_0 + a_3Ks + a_4d + a_5B. \quad (6)$$

2.4 Adaptive neuro-fuzzy inference system, ANFIS

ANFIS is known as a powerful and efficient machine learning approach which is a combination of adaptive multi-layer feedforward neural networks (ANN) and fuzzy inference system (Jang 1993; Zounemat-Kermani et al. 2020).

As shown in Fig. 2, the ANFIS network consists of five interconnected layers with some functional nodes in each layer. Assume that O_i^j is a functional node; it denotes the output of the i th node in the j th layer. For the sake of simplicity in describing the network architecture, here the ANFIS under consideration has two input variables for the network (flow discharge, Q , and pipe diameter, D) and one output (Cv). The output of layer 1 will be calculated as follows:

$$\begin{aligned} O_i^1 &= \mu A_i(Q), & i &= 1, 2 \\ O_i^1 &= \mu B_{i-2}(D), & i &= 3, 4 \end{aligned} \quad (7)$$

μA_i and μB_i are the membership functions which are normally chosen to be bell-shaped with maximum equal to unity and minimum equal to zero such as:

$$\begin{aligned} \mu A_i(Q) &= \exp\left(\frac{-(Q - c_i)}{(a_i)^2}\right), & i &= 1, 2 \\ \mu B_i(D) &= \exp\left(\frac{-(D - c_i)}{(a_i)^2}\right), & i &= 3, 4 \end{aligned} \quad (8)$$

a_i and c_i are premise parameters which have to be tuned during the training of the network. As can be seen in Fig. 2, every node in the second layer is marked with a circle node labelled Π which multiplies the incoming signals from the first layer (O_i^1) and sends the product out. For instance,

$$O_i^2 = \omega_i = \mu A_i(Q) \cdot \mu B_i(D), \quad i = 1, 2. \quad (9)$$

The outputs of the second layer represent the firing strength of a rule. The nodes in the third layer are labelled N in Fig. 2 which computes the ratio of i th firing strength of the i th rule to the sum of firing strength of all rules:

$$O_i^3 = \bar{\omega}_i = \frac{\omega_i}{\omega_1 + \omega_2}, \quad i = 1, 2. \quad (10)$$

In ANFIS, the Takagi–Sugeno fuzzy inference system is used. Therefore, the consequent part of the ANFIS network in terms of $\{p_i, q_i, r_i\}$ parameter set can be written as follows:

$$O_i^4 = \bar{\omega}_i \cdot f_i = \bar{\omega}_i \cdot (p_iQ + q_iD + r_i), \quad i = 1, 2. \quad (11)$$

Finally, in the fifth layer, the single node Σ computes the output (Cv) as the summation of the previous layer’s incoming signals (Kisi and Zounemat-Kermani 2014; Keshtegar et al. 2018).

$$Cv = O_1^5 = Cv = \bar{\omega}_i \cdot f_i = \bar{\omega}_i \cdot (p_iQ + q_iD + r_i), \quad i = 1, 2. \quad (12)$$

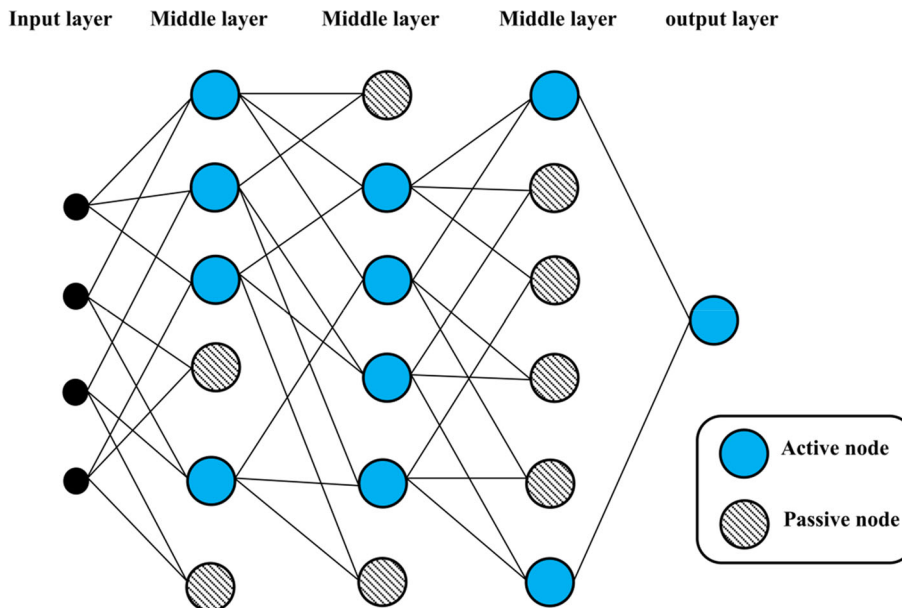
2.5 Group method of data handling, GMDH

The group method of data handling (GMDH) can be introduced as a self-organizing version of a multi-layer feedforward neural network in which layers and nodes are generated based on the input vector as shown in Fig. 3. In other words, the connection of the neurons between the network’s layers is not predetermined and fixed but are chosen and tuned during the training process to optimize the network. Hence, similar to multi-layer neural networks, in GMDH the neurons are interconnected using a polynomial through synapses. In this study, the general connection between the input and output variables is expressed by the quadratic Ivakhnenko polynomial in the form of:

$$f(x_i, x_j) = a_0 + \sum_{i=1}^{\vartheta} a_i x_i + \sum_{i=1}^{\vartheta} \sum_{j=1}^{\vartheta} a_{ij} x_i x_j \quad (13)$$

where ϑ is the number of input variables, x_i are the input variables, and a_i are the coefficients (weights). Henceforward, considering Q and D as the input variables for

Fig. 3 Schema of a quadratic polynomial GMDH with two input parameters (Q & D) and four layers



predicting C_v , the quadratic form may be expressed as follows:

$$f(Q, D) = a_0 + a_1 \cdot Q + a_2 \cdot D + a_{11} \cdot Q^2 + a_{12} \cdot Q \cdot D + a_{21} \cdot D \cdot Q + a_{22} \cdot D^2. \tag{14}$$

In the GMDH model, each layer produces new neurons for the next layer and when a neuron delivers an external input, synapses determine the contribution at the response of that neuron. On that account, a neuron might be eliminated from the network as a passive neuron (Mrugalski and Witczak 2002; De Giorgi et al. 2016; Mo et al. 2018).

The general architecture of a developed GMDH with four inputs and five layers is shown in Fig. 3. In the first layer, input variables are fed into the model as quadratic transfer functions (see Eq. 14) and then some candidate nodes are generated. The number of generated nodes is calculated as the following equation:

$$N_n^{l+1} = \binom{N_n^l}{2} = \frac{N_n^l!}{2 \times (N_n^l - 2)!} \tag{15}$$

where N_n^{l+1} is the number of nodes of the next layer and N_n^l is the number of nodes of the previous layer. This equation shows that the number of neurons would increase from layer to layer. It is necessary to apply a strategy to prevent the excessive growth of the network. In this study, a five-layer GMDH network was developed. Also, the maximum neurons of each layer were considered ten (except for the one layer before the last layer, since this layer needs two neurons, based on the quadratic polynomial function). Subsequently, based on the least square methods, some of the neurons will be eliminated and considered as

passive nodes. Detailed information about the GMDH model can be found at Farlow (1984) and Ivakhnenko and Ivakhnenko (1995).

2.6 Firefly optimization algorithm, FA

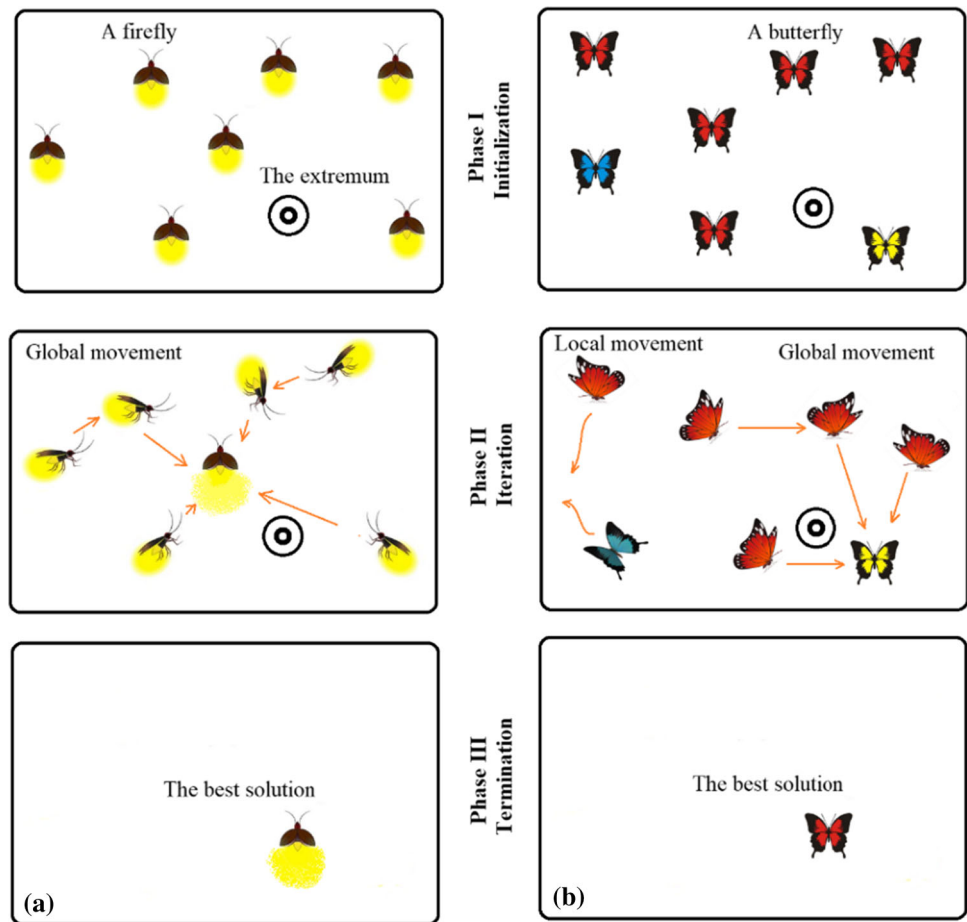
The firefly optimization algorithm (FA) is a global nature-inspired swarm intelligence algorithm inspired by collective behaviour, such as insects and birds, in nature which is based on the social behaviour of fireflies. With respect to the recent bibliography, the FA is very efficient and can outperform conventional heuristic algorithms such as genetic algorithms, for solving complicated optimization problems (Yang 2009).

There are mainly three modified rules based on flashing characteristics of fireflies for setting up the FA as follows: (1) All fireflies are assumed to be unisex, and they move towards brighter ones regardless of their sex, and (2) attractiveness of a firefly is proportional to its brightness. The brighter a firefly, the more attractive it is to the other ones. However, the brightness decreases as the distance from the other firefly increases. The firefly moves randomly if it cannot find and distinguish a brighter firefly in its surroundings.

(3) The brightness or light intensity of an agent (firefly) in the FA is pertinent to the objective function of an optimization problem. For instance, in maximization problems, the light intensity is in line with the value of the objective function.

Generally, there are three distinct phases in FA: (i) Initialization phase, (ii) Iteration phase, and (iii) termination phase (see Fig. 4a). In the initiation phase, a population of several fireflies is generated. Afterwards, the

Fig. 4 General view for the three phases of initiation, iteration, and termination in **a** firefly algorithm and **b** butterfly optimization algorithm



following decreasing function is introduced for determining the attractiveness of a firefly (β):

$$\beta(r) = \beta_0 \exp(-\gamma r^2) \tag{16}$$

where $r_{ij} = \|z_i - z_j\|$ denotes the Euclidean distance between any two z_i and z_j fireflies, β_0 is the initial attractiveness ($r = 0$), and γ is an absorption coefficient of a firefly's light intensity.

In the second phase, in an iterative procedure, the movement of a firefly i towards a brighter and more attractive firefly j is given by the following equation:

$$z_i^{updated} = z_i + \beta_0 \cdot \exp(-\gamma r_{ij}^2) (z_j - z_i) + \alpha \epsilon_i \tag{17}$$

where α is a randomization parameter and ϵ_i is a random variable from a Gaussian distribution. Finally, when the stopping criterion is met, the fireflies are ranked and the best solution is found (Apostolopoulos and Vlachos 2010; Wang et al. 2018).

2.7 Butterfly optimization algorithm, BOA

The novel optimization technique, namely butterfly optimization algorithm (BOA), is a nature-inspired algorithm

that was introduced by Arora and Singh (2019). In BOA, butterflies are substituted as search agents in the solution process. In order to find the optimum solution to a problem (e.g. a source of nectar), butterflies use sense receptors which are nerve cells on butterflies' body surface and are called chemoreceptors. During the search process, butterflies generate fragrances with some intensity which are correlated with their fitness. As butterflies move from one location to another, their fitness varies accordingly. Their fragrance spreads and propagates over distance, and other butterflies can recognize it, and in this way, the butterflies can share their information with other butterflies and create a social knowledge network.

Similar to fireflies' general movement towards the brighter firefly, a butterfly moves towards the fragrance of another butterfly. Whenever a butterfly cannot sense fragrance from its surrounding, it will move randomly and this phase is named as local search in the BOA (see Fig. 4b). The BOA takes into account three modified rules for the optimization process: (1) All butterflies are supposed to produce some scent and fragrance which make them attract each other. (2) Each butterfly moves randomly or towards the butterfly producing more fragrance. (3) The stimulus

intensity of each butterfly is influenced by the objective function. In BOA, each fragrance has its own unique fragrance which is calculated as the following:

$$fr = cI^\phi \tag{18}$$

where fr represents the perceived magnitude of the butterfly’s fragrance, I is the stimulus intensity, c is the sensory modality, and ϕ is the power exponent dependent on the varying degree of absorption. Like FA, there are three main phases in BOA: (i) initialization phase, (ii) iteration phase, and (iii) termination phase (see Fig. 4b). In the initiation phase, all butterflies are positioned randomly in the search space, with their fragrance and fitness values calculated and stored. During the next phase, the algorithm starts the iteration phase. In each iteration step, all butterflies in the solution space move to their new positions (globally or locally), and afterwards their fitness values are evaluated. In the global search phase, the butterfly takes a step towards the best (fittest) butterfly given in Eq. (19):

$$z_i^{updated} = z_i + (\varepsilon^2 \cdot z_j^* - z_i)fr_i; \quad r < p \tag{19}$$

where z_i and z_j are i th and j th butterflies in the solution space, z_j^* represents the current best solution, and ε is a random variable. Parameter p is a fraction between zero and unity which is affected by the environmental factors (e.g. wind and rain) in the searching process. The local movement of the butterfly i can be represented as:

$$z_i^{updated} = z_i + (\varepsilon^2 \cdot z_j - z_k)fr_i; \quad r \geq p. \tag{20}$$

The iteration phase will be continued until the stopping criteria (e.g. maximum epoch or convergence criterion) are met. Then, we reach the final phase with the optimum solution of the objective problem.

2.8 Integration of the machine learning models and nature-inspired algorithms

As stated earlier in the text, in addition to the standard versions of ANFIS and GMDH models, four integrated models (ANFIS-FA, ANFIS-BOA, GMDH-FA, and GMDH-BOA) are proposed to evaluate the potential enhancement in the training (simulation) and testing (prediction) performance of standard ANFIS and GMDH models. Hence, the nature-inspired algorithms such as BOA and FA are combined with the standard ANFIS and GMDH for constructing hybrid models. These nature-inspired algorithms can be used for optimizing the premise or consequent parts of the ANFIS (Zounemat-Kermani and Mahdavi-Meymand 2019; Mahdavi-Meymand et al. 2019). In this research, the BOA and FA are employed for optimizing the Gaussian membership function parameters of

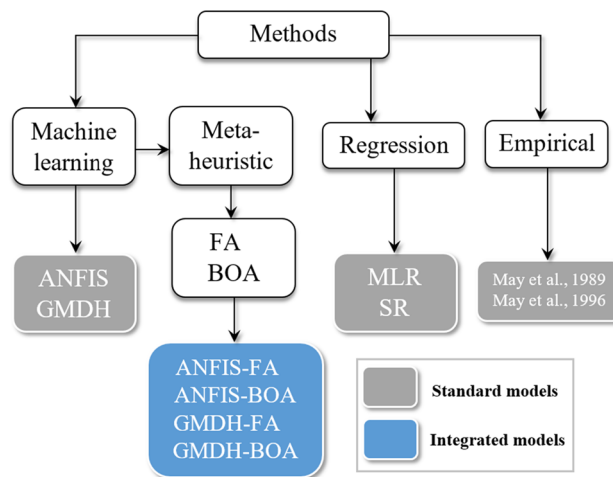


Fig. 5 General framework of the applied methods used in this study

the inputs and linear membership function parameters of the outputs.

In addition, the BOA and FA might also be integrated with the GMDH for potential optimization of either weights of the neurons in the network or the architecture of the network (the number of neurons in each layer and the number of layers). In this research, weights of the polynomial function of the network (Eq. 14) were optimized by applying the FA and BOA.

The general framework for modelling chosen for this applied research is shown in Fig. 5.

BOA and FA, like most of the nature-inspired optimization methods, have some initial values and optimizing coefficients. These parameters were chosen based on previous studies (Arora and Singh 2019; Mahdavi-Meymand and Zounemat-Kermani 2020). In Table 2, the general characteristics and initial values for the applied machine learning models are given.

2.9 Models’ evaluation

In this study, to assess the suitability of the applied models, five statistical statistics (RMSE, MAE, R^2 , IA, NSE), as well as a comprehensive index (SI), are calculated. Two deviance measures including the root-mean-square error (RMSE) and mean absolute error (MAE) in addition to two similarity measures including coefficient of determination (R^2) and index of agreement (IA) as well as the Nash–Sutcliffe model efficiency coefficient (NSE) are used. Also, a synthesis index (SI) based on RMSE, MAE, $(1-R^2)$, $(1-IA)$, and $(1-NSE)$ is calculated to obtain a comprehensive performance criterion (Chou et al. 2014). The mathematical formulae of the mentioned statistics, as well as evaluation values for comparing the results, are given in Table 3.

Table 2 Initial and tuning parameters for the applied machine learning models

Model/algorithm	Parameter	Value/type
ANFIS	FIS type	Takagi–Sugeno
	MF shapes	Gaussian MFs in the first to the fourth layers, linear MF in the fifth layer
	Derivation of data	Subtractive clustering method
	Number of MFs for inputs and output	5
GMDH	Number of layers	5
	Maximum number of neurons in each layer	10
	Polynomial	Quadratic
FA	Attraction coefficient	2
	Light absorption coefficient	1
	Mutation coefficient	0.2
	Coefficient damping ratio	0.98
BOA	Probability switch	0.8
	Sensory modality	0.01
	Power exponent	0.1

Table 3 Descriptions of the used statistical measures for the evaluation of the applied models in this research

Type	Statistical measure	Formula	Range and ideal value(s)
Deviance measure	RMSE (mm)	$\sqrt{\frac{\sum_{i=1}^N (Cv_p - Cv_o)^2}{N}}$	[0, + ∞) Ideal = 0
	MAE (mm)	$\frac{1}{N} \sum_{i=1}^N Cv_p - Cv_o $	[0, + ∞) Ideal = 0
Similarity measure	R^2	$\frac{\left(\sum_{i=1}^N (Cv_p - \bar{Cv}_p)(Cv_o - \bar{Cv}_o)\right)^2}{\sum_{i=1}^N (Cv_p - \bar{Cv}_p)^2 \sum_{i=1}^N (Cv_o - \bar{Cv}_o)^2}$	[- 1,1] Ideal = 1
	IA	$1 - \frac{\sum_{i=1}^N (Cv_p - Cv_o)^2}{\sum_{i=1}^N \left(Cv_o - \bar{Cv}_o + Cv_p - \bar{Cv}_p \right)^2}$	[0,1] Ideal = 1
Efficiency measure	NSE	$1 - \frac{\sum_{i=1}^N (Cv_p - Cv_o)^2}{\sum_{i=1}^N (Cv_o - \bar{Cv}_o)^2}$	(- ∞,1] Ideal = 1 Rejected < 0
Comprehensive measure	SI	$\frac{1}{M} \sum_{j=1}^M \left(\frac{P_j - P_{min,j}}{P_{max,j} - P_{min,j}} \right)$	[0,1] Desirable = 0

Where Cv_o is the observed value of volumetric sediment concentration, \bar{Cv}_o is the average of the observed volumetric sediment concentration, Cv_p denotes the predicted volumetric sediment concentration in the testing set as well as the simulated volumetric sediment concentration in the training set, P_j denotes the j th performance measure, \bar{Cv}_p is the average value of the predicted volumetric sediment concentration, N is the number of data samples, and $M = 5$ is the number of performance measures and the bar indicates the mean value.

3 Implementation

3.1 Data preparation

At first, all the available data set is shuffled randomly. Then, data are standardized between the ranges of zero and unity. Afterwards, by using the hold-out method, the original data set is separated into the training and testing sets. The training data set is used for the training process of the machine learning models as well as estimation of the partial descriptions of the nonlinear system of sediment transport in sewers, and the testing data set is used for the final assessment and complete description of the model. Thereupon, from the total 194 data sets used in this study,

Table 4 General structures of the applied predictive models in terms of input selection strategy

Scenario	Input selection method	Types of input parameters	Input parameters	Predictive models
I	Hydraulics of sediment transport in sewers	Dimensional	$Q, D, S_0, \lambda s, Ks, n, T, Y, d, A, P, B, Ws$	ANFIS[I]; ANFIS-FA[I]; ANFIS-BOA[I]; GMDH[I]; GMDH-FA[I]; GMDH-BOA[I]; MLR[I]
II	Forward selection procedure	Dimensional	Q, S_0, Ks, B, d	ANFIS[II]; ANFIS-FA[II]; ANFIS-BOA[II]; GMDH[II]; GMDH-FA[II]; GMDH-BOA[II]; SR[II]
III	Dimensionless analysis	Dimensionless	$S_0, \frac{v_0}{D}, \frac{D^2}{A}, \frac{d}{R}, Fr$	ANFIS[III]; ANFIS-FA[III]; ANFIS-BOA[III]; GMDH[III]; GMDH-FA[III]; GMDH-BOA[III]; SR[III]
			$\frac{v_0}{D}, \frac{D^2}{A}, \frac{d}{R}, Fr, 1 - \frac{V_c}{V}$	May et al. (1989)
			$\frac{D^2}{A}, \frac{d}{D}, Fr, 1 - \frac{V_c}{V}$	May et al. (1996)

164 data sets were used for the training process (nearly 85%) and 30 data sets (15%) were considered for the testing set.

3.2 Input selection and modelling scenarios

In this study, three different scenarios are considered in order to select the subset of 13 input candidates for sediment prediction in sewers. In the first scenario (I), all the available dimensional hydraulics and geometric factors are taken into account. In this case, the predictive models (ANFIS, ANIFS-FA, ANFIS-BOA, GMDH, GMDH-FA, GMDH-BOA, and MLR) will have 13 input factors (see Table 4) which makes the models’ structure complicated. However, in the second and third scenarios (II, III) the forward selection procedure is used for choosing the best dimensional and non-dimensional subset of the independent variables using ANFISs, GMDHs, and stepwise regression (SR) models.

Based on the results of the forward selection in the second scenario, a five-dimensional input vector (Q, S_0, Ks, B, d) with the most significant effect on the target (C_v) is selected and the other variables are removed. Similar to scenario II, five dimensionless independent parameters ($S_0, \frac{v_0}{D}, \frac{D^2}{A}, \frac{d}{R}, Fr$) are opted for simulating C_v values. The

final forms of all machine learning (ANFISs, GMDHs), regression (MLR and SR), and empirical equation models (May et al. 1989, 1996) with respect to the three input scenarios are given in Table 3.

4 Application and results

Tables 5 and 6 report the statistical measures for the proposed machine learning and MLR models considering all of the hydraulic and geometric input parameters (scenario I) for the training and testing sets, respectively. Looking at the accuracy of training and testing data (Tables 5 and 6), both ANIFS and GMDH machine learning models (ANFIS[I], ANFIS-FA[I], ANFIS-BOA[I], GMDH[I], GMDH-FA[I], and GMDH-BOA[I]) thoroughly outperformed the MLR model, with a considerable enhancement in the averaged RMSE equal to 43% for the training set and 24% for the testing set.

Table 5 also indicates that the synthesis index of the ANFIS-BOA[I] model is less than the other applied models for both the training ($SI = 0.017$) and testing sets ($SI = 0.047$) for modelling and predicting C_v values, which implies the better performance of this integrated model.

The results of the applied models considering forward selection for the dimensional (scenario II) and

Table 5 Evaluation performance of the applied models considering scenario (I) for the training set (bold font implies the best value)

Model	RMSE (ppm)	R^2	MAE (ppm)	NSE	IA	SI	Rank
ANFIS[I]	113.845	0.890	71.632	0.886	0.972	0.062	2
ANFIS-BOA[I]	106.573	0.904	71.541	0.899	0.974	0.017	1
ANFIS-FA[I]	121.958	0.874	77.048	0.867	0.965	0.128	6
GMDH[I]	121.545	0.875	68.705	0.870	0.966	0.103	5
GMDH-BOA[I]	116.924	0.884	64.793	0.879	0.969	0.063	3
GMDH-FA[I]	120.954	0.876	67.059	0.873	0.966	0.096	4
MLR[I]	203.833	0.647	146.170	0.644	0.882	1.000	7

Bold font indicates the best performance

Table 6 Evaluation performance of the applied models considering scenario (I) for the testing set (bold font implies the best value)

Model	RMSE (ppm)	R^2	MAE (ppm)	NSE	IA	SI	Rank
ANFIS[I]	151.687	0.808	92.055	0.777	0.932	0.252	6
ANFIS-BOA[I]	135.061	0.831	96.118	0.823	0.950	0.047	1
ANFIS-FA[I]	142.423	0.807	99.387	0.804	0.945	0.155	4
GMDH[I]	142.385	0.811	92.814	0.804	0.943	0.14	3
GMDH-BOA[I]	140.43	0.812	87.213	0.809	0.942	0.114	2
GMDH-FA[I]	150.243	0.803	79.324	0.781	0.943	0.168	5
MLR[I]	189.540	0.667	151.027	0.652	0.901	1.000	7

Bold font indicates the best performance

Table 7 Evaluation performance of the applied models considering scenario (II) for the training set (bold font implies the best value)

Model	RMSE (ppm)	R^2	MAE (ppm)	NSE	IA	SI	Rank
ANFIS[II]	171.976	0.753	109.34	0.749	0.922	0.553	5
ANFIS-BOA[II]	128.66	0.861	86.404	0.859	0.966	0.242	4
ANFIS-FA[II]	91.085	0.931	57.589	0.930	0.982	0.000	1
GMDH[II]	130.362	0.856	70.794	0.856	0.963	0.215	2
GMDH-BOA[II]	130.707	0.855	76.217	0.855	0.963	0.229	3
GMDH-FA[II]	130.364	0.856	70.794	0.856	0.968	0.215	2
SR[II]	218.609	0.594	147.399	0.594	0.857	1	6

Bold font indicates the best performance

Table 8 Evaluation performance of the applied models considering scenario (II) for the testing set (bold font implies the best value)

Model	RMSE (ppm)	R^2	MAE (ppm)	NSE	IA	SI	Rank
ANFIS[II]	151.304	0.779	103.807	0.778	0.935	0.39	6
ANFIS-BOA[II]	140.45	0.814	97.929	0.809	0.941	0.255	5
ANFIS-FA[II]	123.655	0.862	74.649	0.852	0.956	0.000	1
GMDH[II]	140.016	0.815	85.520	0.813	0.945	0.203	4
GMDH-BOA[II]	136.401	0.845	89.703	0.819	0.942	0.173	3
GMDH-FA[II]	132.125	0.835	81.396	0.831	0.951	0.107	2
SR[II]	188.297	0.658	143.333	0.657	0.890	1.000	7

Bold font indicates the best performance

Table 9 Evaluation performance of the applied models considering scenario (III) for the training set (bold font implies the best value)

Model	RMSE (ppm)	R^2	MAE (ppm)	NSE	IA	SI	Rank
ANFIS[III]	62.729	0.967	33.665	0.967	0.991	0.007	2
ANFIS-BOA[III]	60.354	0.969	32.656	0.969	0.992	0.000	1
ANFIS-FA[III]	63.621	0.966	36.201	0.966	0.991	0.013	3
GMDH[III]	122.026	0.874	68.467	0.874	0.965	0.248	4
GMDH-BOA[III]	125.452	0.867	70.293	0.866	0.964	0.263	5
GMDH-FA[III]	125.639	0.866	72.311	0.866	0.963	0.268	6
SR[III]	177.129	0.733	123.635	0.733	0.918	0.595	7
May et al. (1996)	268.037	0.677	183.973	0.392	0.868	1.000	9
May et al. (1989)	197.125	0.679	91.138	0.674	0.898	0.663	8

Bold font indicates the best performance

Table 10 Evaluation performance of the applied models considering scenario (III) for the testing set (bold font implies the best value)

Model	RMSE (ppm)	R^2	MAE (ppm)	NSE	IA	SI	Rank
ANFIS[III]	162.991	0.763	102.251	0.743	0.933	0.202	4
ANFIS-BOA[III]	144.881	0.811	91.037	0.797	0.946	0.009	1
ANFIS-FA[III]	153.186	0.785	93.098	0.773	0.939	0.095	3
GMDH[III]	167.152	0.768	104.859	0.730	0.933	0.206	5
GMDH-BOA[III]	176.683	0.755	116.398	0.698	0.927	0.298	6
GMDH-FA[III]	142.385	0.811	92.814	0.804	0.943	0.011	2
SR[III]	189.873	0.682	143.920	0.651	0.907	0.566	8
May et al. (1996)	257.659	0.721	181.133	0.357	0.866	0.941	9
May et al. (1989)	176.838	0.708	100.447	0.697	0.898	0.408	7

Bold font indicates the best performance

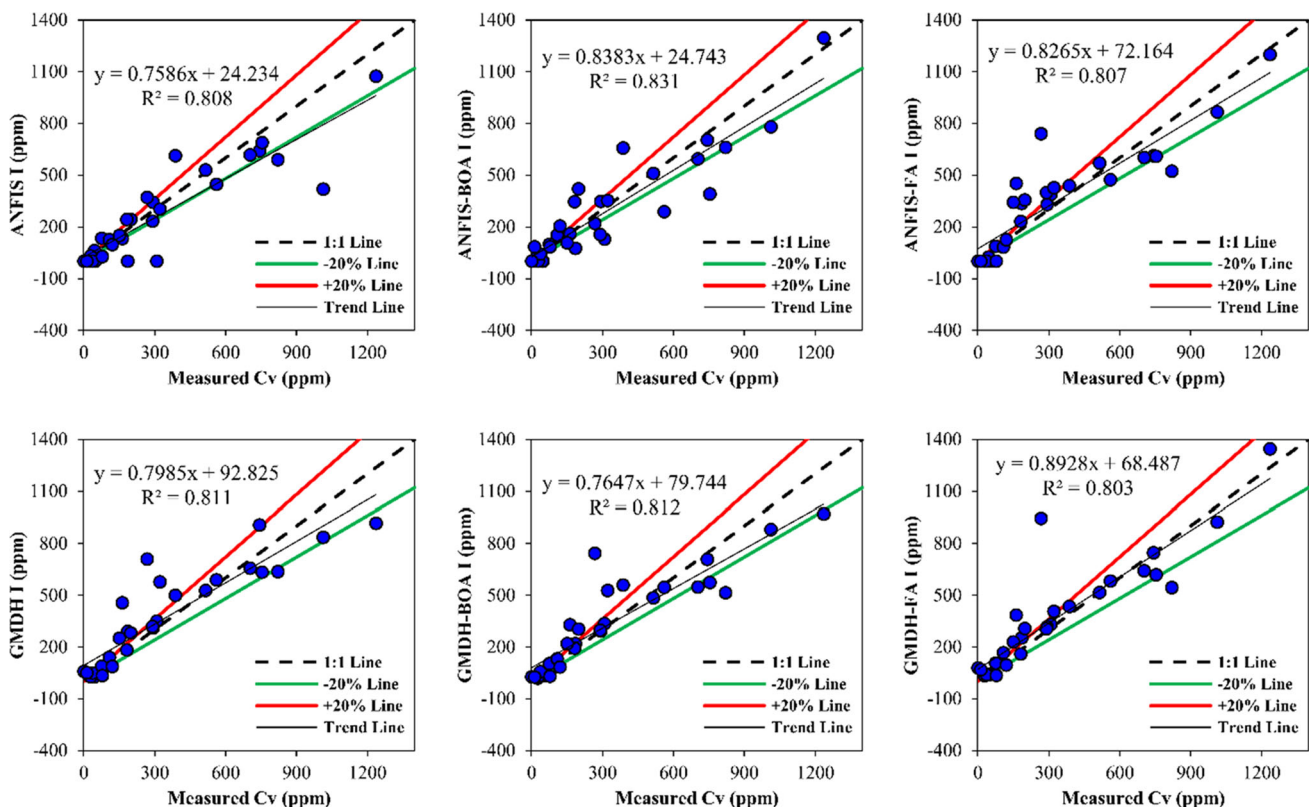


Fig. 6 Scatter plots for the performance of machine learning methods for predicting C_v in the testing set considering scenario (I)

dimensionless (scenario III) input parameters are given in Tables 7, 8, 9, and 10 for the training and testing sets. Similar to the findings of the first scenario (see Tables 5 and 6), the machine learning models surpassed the stepwise regression (SR) and empirical equations in simulating and predicting C_v values.

Note that in general, the performance of the integrated ANFIS and GMDH models with the BOA and FA is generally better than the performance of the standard

ANFIS and GMDH models. However, the superiority of the FA is more apparent for the second input scenario (II) (observing Tables 7 and 8), while considering the results of Tables 9 and 10, the BOA gave a better performance for the training and testing sets of dimensionless input variables (scenario III).

Performances of results for the testing set of the applied models are also presented in Figs. 6, 7, 8, and 9 in terms of scatter plots. From Figs. 6, 7, 8, and 9, it is evident that the

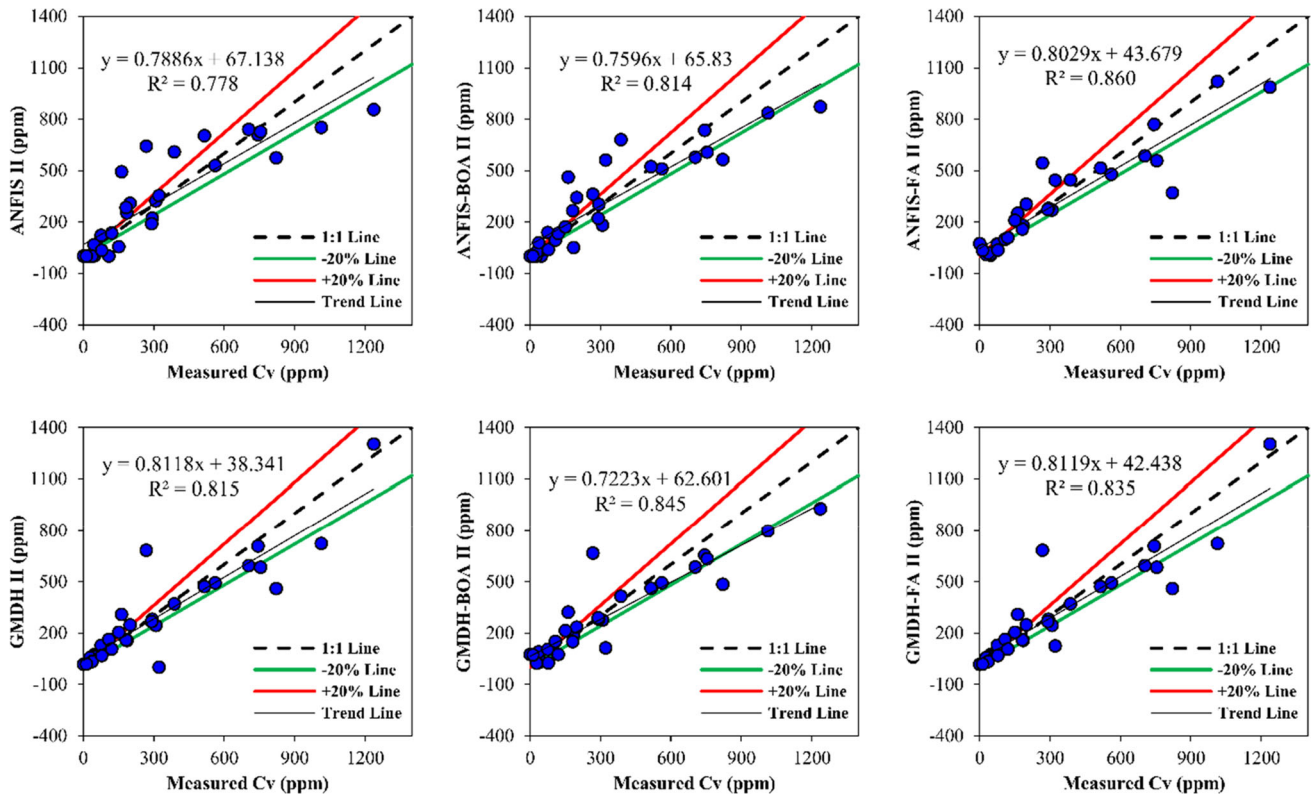


Fig. 7 Scatter plots for the performance of machine learning methods for predicting C_v in the testing set considering scenario (II)

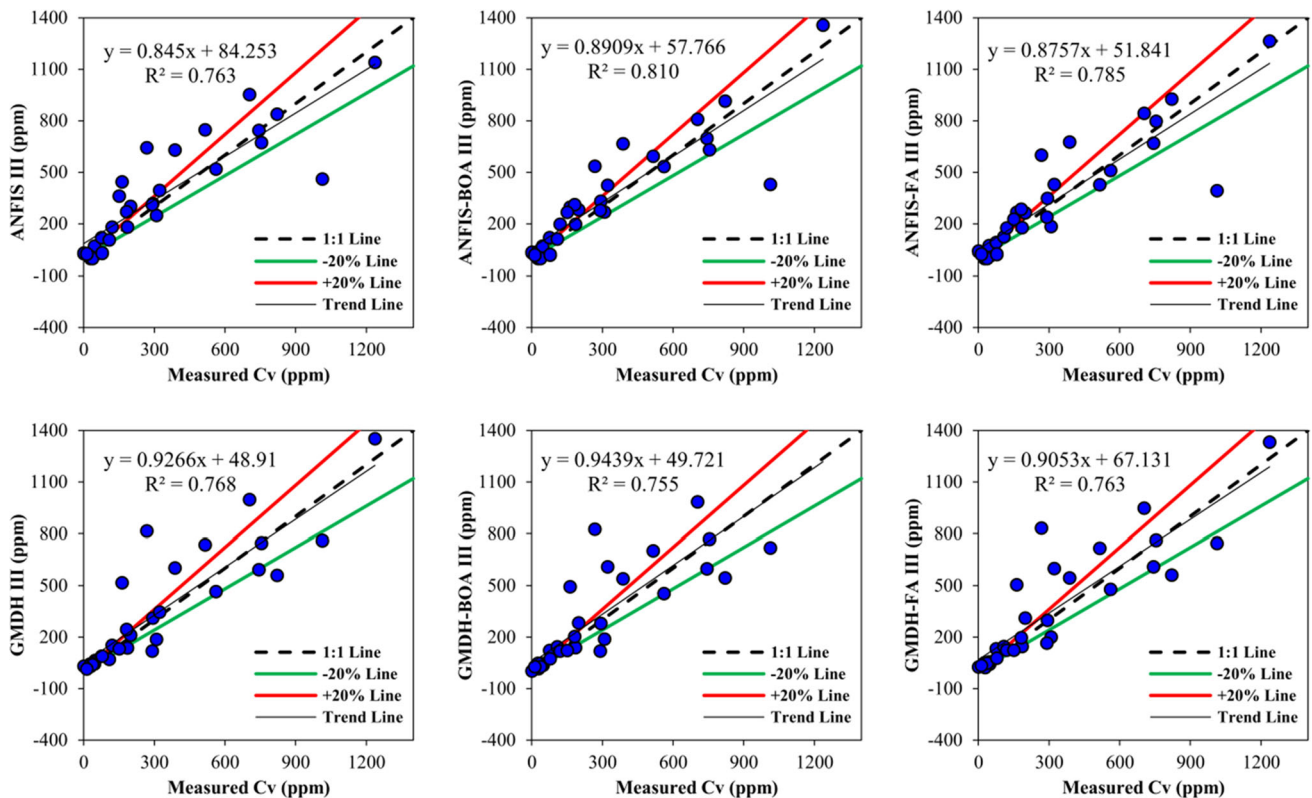


Fig. 8 Scatter plots for the performance of machine learning methods for predicting C_v in the testing set considering scenario (III)

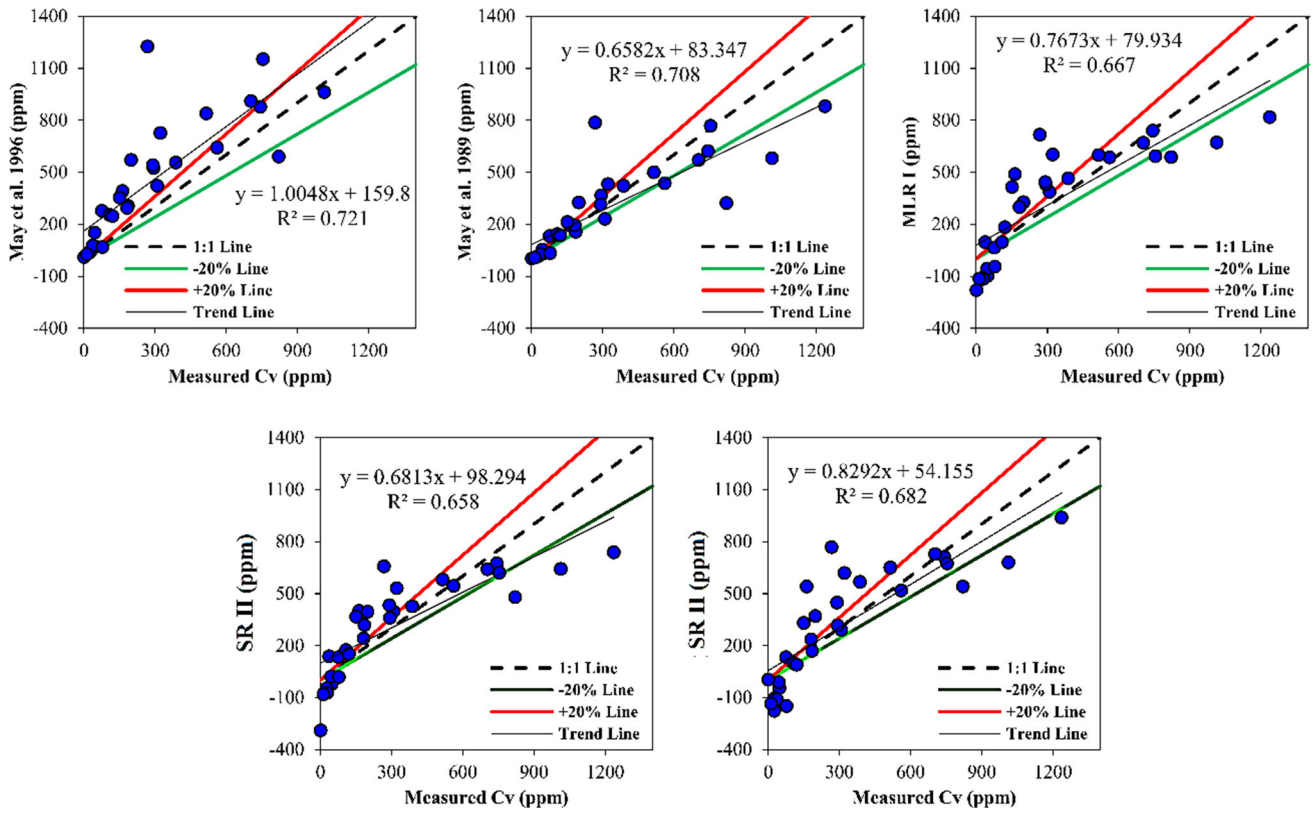


Fig. 9 Scatter plots for the MLR[I], SR[II], SR[III], and empirical equations for predicting Cv in the testing set

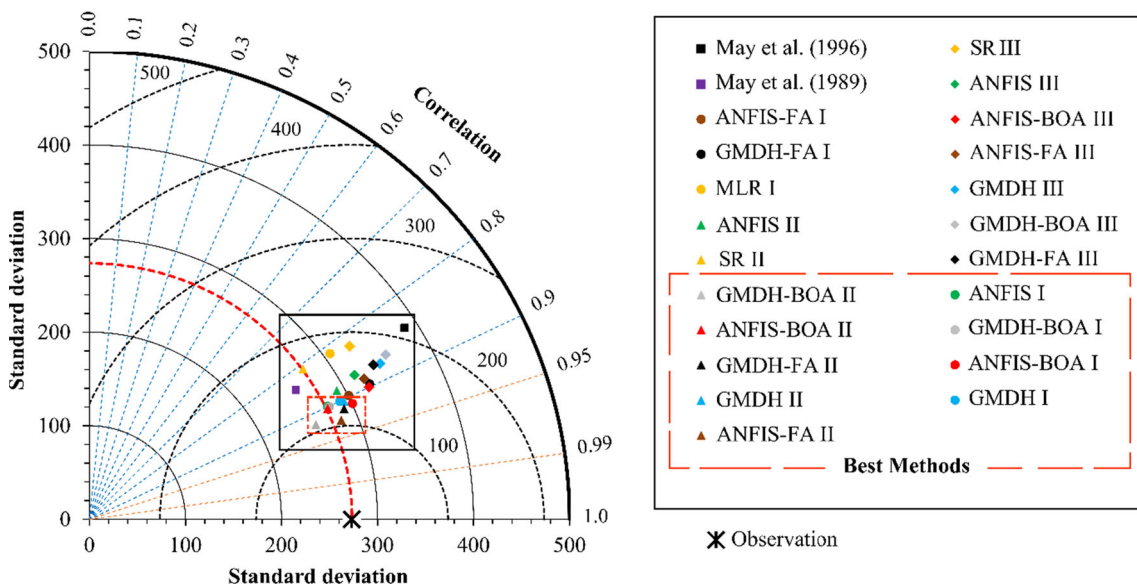


Fig. 10 Taylor diagram for displaying the goodness of the predictive models' performance using three statistic measure (correlation coefficient, RMSE, and standard deviation)

Table 11 Percentage of the improved efficiency of BOA and FA on GMDH and ANFIS models

Statistic	Algorithm	Training		Testing		Total average (%)
		Standard GMDH (%)	Standard ANFIS (%)	Standard GMDH (%)	Standard ANFIS (%)	
RMSE _{Ave}	FA	1.58*	4.19	0.06	10.13	3.99**
	BOA	0.24	11.79	0.02	9.75	5.45
R ² _{Ave}	FA	0.51	5.05	0.27	4.39	2.55
	BOA	0.04	5.53	0.70	4.50	2.69
Average		0.59	6.64	0.26	7.19	3.67

*Integrating the FA with the standard GMDH (GMDH-FA) improved the averaged RMSE of the training set equal to 1.58%; ** integrating the FA to the machine learning models (GMDH-FA and ANFIS-FA) improved the total averaged RMSE of the standard GMDH and ANFIS equal to 3.99%

empirical equation by May et al. (1996) mostly over-predicted the *C_v* values, whereas ANFIS[I], AFIS-BOA[1], GMDH[II], GMDH-BOA[II], and GMDH-FA[II] under-predicted the *C_v* values. All the regression models (MLR[I], SR[II], and SR[III]) over-predicted lower *C_v* values (*C_v* < 600 ppm) and under-predicted higher *C_v* values (*C_v* > 600 ppm). Furthermore, the regression models provided some unfavourable negative predicted results for *C_v* which is totally unacceptable.

For having a better judgement and visualization, the final efficiency of all the applied models in the testing set is depicted in terms of a polar plot called the Taylor diagram in Fig. 10. Taylor diagram provides a statistical summary of how well simulated and predicted values match observed values in terms of three statistics of correlation coefficient (*r*) as the azimuth angle, standard deviation as the radial distance from the origin, and RMSE as the distance from the reference observation point (Taylor 2001). In the diagram shown in Fig. 8, all models have been plotted as points, as if their positions are precise indicators of their true predictive performance. Through visual diagnosis, the best predictive models are distinguished as the “Best Methods” from the other models in Fig. 10.

Table 12 Comparison of the results employing different input vectors on the performance of machine learning models

Input scenario	Training		Testing	
	RMSE _{Ave}	R ² _{Ave}	RMSE _{Ave}	R ² _{Ave}
I	121.89*	0.87	143.71	0.81
II	130.47	0.85	137.32	0.83
III	93.30	0.92	162.33	0.77

*Averaged RMSE values for the ANFIS[II], ANFIS-FA[II], and ANFIS-BOA[II] in the training set. Bold font indicates the best performance

However, it can be observed that the ANFIS-FA[II] gave a better performance than the other models.

5 Discussion

Since the major objective of this study is to challenge the capability of FA and BOA in optimization problems, in Table 11 a statistical analysis for the magnitude of improvement brought by these algorithms to the standard ANFIS and GMDH models is given. In Table 11, the averaged values of the coefficient of determinations (R²_{Ave}) and the RMSE_{Ave} between models’ predicted outputs and observations (Tables 5, 6, 7, 8, 9, and 10) are used as the indicators for the percentage of the improved efficiency.

Both the FA and BOA improved the performance of the standard ANFIS and GMDH models. However, these algorithms presented considerable improvement for the ANFIS model. In general, the BOA slightly outperformed the FA for improving the performance of machine learning models regarding the RMSE_{Ave} and R²_{Ave} criteria. In Table 12, the general performance of the machine learning models in terms of the input vector scenarios (I, II, and III) using the averaged values of the coefficient of determinations (R²_{Ave}) and the RMSE_{Ave} between models’ predicted outputs and observations (Tables 5, 6, 7, 8, 9, and 10) is shown.

The summary results of Table 12 imply that utilizing the forward selection procedure in the second scenario has boosted the effectiveness of machine learning models in predicting *C_v* values. Although employing a dimensionless input vector (scenario III) has improved the performance of the machine learning models in the training phase, it failed to elevate the efficiency of these models in the testing phase. This conclusion is also verified by the outcome of the Taylor diagram in Fig. 10 so that none of the third scenario models located in the “Best Methods” box.

Table 13 The Mann–Whitney test results between measurement and predicted values (scenario III)

Method	<i>p</i> value	Significantly different (90%)	Significantly different (95%)
Measurement versus May et al. (1996)	0.096	Yes	No
Measurement versus May et al. (1998)	0.959	No	No
Measurement versus MLR	0.959	No	No
Measurement versus ANFIS	0.739	No	No
Measurement versus ANFIS-BOA	0.842	No	No
Measurement versus ANFIS-FA	0.982	No	No
Measurement versus GMDH	0.877	No	No
Measurement versus GMDH-BOA	0.947	No	No
Measurement versus GMDH-FA	0.684	No	No

The findings of this study also revealed that the empirical equations and regression models could not surpass any of the machine learning models. Surprisingly, May et al.'s (1996) equation was the weakest predictive model and acted worse than even the former equation by May et al. (1989). In order to evaluate the models' performances statistically, the results of the Mann–Whitney test are considered. The Mann–Whitney is a statistical test that can be used to figure out if there is a significant difference between the measured and predicted data. In Table 13, the results of the Mann–Whitney test for the third scenario are provided. The results of Table 13 indicate that in a 95% confidence level there is no significant difference between measured and predicted values of all models. On the other hand, in a 90% confidence level, just the predicted results of May et al.'s (1996) equation show a significant difference with the measured values.

6 Conclusions

Novel integrated approaches were proposed in this paper for sewers' volumetric concentration of sediments (C_v) predicting. The proposed approaches are based on the combination of two nature-inspired algorithms (firefly algorithm (FA) and butterfly optimization algorithm (BOA)) and two machine learning approaches (ANFIS and GMDH). The selection of the best input features for C_v prediction was accomplished by forward selection procedure and dimensional analysis using the Buckingham Π theorem so that three scenarios were employed for constructing the applied methods. Accordant with the obtained

results, the following conclusions can be drawn from the research:

- Due to the complexity of the sediment transport process in sewers and the wide ranges of input and output data used for the training and testing sets in predicting sediment concentration, the regression (MLR and SR) and empirical equations failed to yield promising results. However, machine learning approaches (ANFIS and GMDH) acted far better than those traditional methods.
- Forward selection method for selecting input parameters improved the prediction capability of both ANFIS and GMDH models. It not only reduced the output predictive error but also simplified the machine learning models structure due to having fewer input variables.
- Considering several statistical measures (e.g. RMSE, MAE, R^2 , and NSE), both ANFIS and GMDH models performed satisfactorily in predicting C_v values. It was not possible to represent a dominant superior model between both of them.
- The proposed application of FA and BOA integration with the ANFIS model (ANFIS-FA, ANFIS-BOA) noticeably improved the performance of standard ANFIS. Nevertheless, engaging FA and BOA with the standard GMDH model (GMDH-FA, GMD-FOA) did not remarkably enhance the efficiency of the GMDH model.

List of symbols ε : Random variable in FA and BOA; γ : Absorption coefficient in FA; ϑ : Number of input variables in GMDH model; ϕ : Power exponent in BOA; α : Randomization parameter in FA and BOA; \sum : Summation operator; A : Cross-sectional area of flow; a_i : Partial regression coefficients in MLR and SR models; ANFIS: Adaptive neuro-fuzzy inference system; B : Water surface width; BOA: Butterfly optimization algorithm; c : Sensory modality in BOA; C_v : Volumetric concentration of sediment; D : Internal diameter of pipe channel; d : Median diameter of particles in a mixture; FA: Firefly optimization algorithm; f_i : The perceived magnitude of fragrance in BOA; Fr_d : Densimetric Froude number; g : Gravitational constant; GMDH: Group method of data handling; G_s : Specific gravity of sediment; I : Stimulus intensity in BOA; IA : Index of agreement; K_s : Overall equivalent sand roughness with sediment; M : Number of samples in training and testing sets; MAE : Mean absolute error; MLR: Multiple linear regression model; n : Overall Manning roughness coefficient with sediment; NSE : Nash–Sutcliffe model efficiency coefficient; O_i^j : Functional node of ANFIS network; p : Environmental fraction in BOA; P : Wetted parameter of flow; p_i, q_i, r_i : Parameter set of ANFIS model; Q : Flow discharge in sewer system; r : Distance between any two fireflies in FA; R : Hydraulic radius; R^2 : Coefficient of determination; RMSE: Root-mean-square error; S_0 : Longitudinal slope of sewer; SI : Synthesis index for models' evaluation; SR: Stepwise regression model; T : Temperature; V : Mean velocity of flow; V_c : Critical incipient motion velocity of sediment; W_s : Width of sediment spread; x_i : Input variables for quadratic polynomial in GMDH model; y_0 : Depth of uniform flow; z_i : Position of the i th agents in FA and BOA; β_0 : Initial attractiveness of a firefly in FA; λ_s : Overall friction factor in sewers; ν : Kinematic viscosity

Acknowledgements The first author wishes to extend his special thanks to the Alexander von Humboldt Foundation for providing financial support for this research project within the framework of the Return Fellowship program.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

- Apostolopoulos T, Vlachos A (2010) Application of the firefly algorithm for solving the economic emissions load dispatch problem. Hindawi publishing corporation international. J Comb 2011:1–23
- Arora S, Singh S (2019) Butterfly optimization algorithm: a novel approach for global optimization. Soft Comput 23(3):715–734
- Azamathulla HM, Ghani AA, Fei SY (2012) ANFIS-based approach for predicting sediment transport in clean sewer. Appl Soft Comput 12(3):1227–1230
- Chou JS, Tsai CF, Pham AD, Lu YH (2014) Machine learning in concrete strength simulations: multi-nation data analytics. Constr Build Mater 73:771–780
- De Giorgi MG, Malvoni M, Congedo PM (2016) Comparison of strategies for multi-step ahead photovoltaic power forecasting models based on hybrid group method of data handling networks and least square support vector machine. Energy 107:360–373
- DeSilva D, Marlow D, Beale D, Marney D (2011) Sewer blockage management: Australian perspective. J Pipeline Syst Eng Pract 2(4):139–145
- Ebtehaj I, Bonakdari H (2013) Evaluation of sediment transport in sewer using artificial neural network. Eng Appl Comput Fluid Mech 7(3):382–392
- Ebtehaj I, Bonakdari H, Shamshirband S, Mohammadi K (2016) A combined support vector machine-wavelet transform model for prediction of sediment transport in sewer. Flow Meas Instrum 47:19–27
- Farlow SJ (1984) Self-organizing methods in modelling: GMDH type algorithms, vol 54. CRC Press, Boca Raton
- Ghani AA (1993) Sediment transport in sewers. Civil Engineering Department, Newcastle University, PhD thesis
- Ghani AA, Md. Azamathulla H (2010) Gene-expression programming for sediment transport in sewer pipe systems. J Pipeline Syst Eng Pract 2(3):102–106
- Hofer T, Montserrat A, Gruber G, Gamerith V, Corominas L, Muschalla D (2018) A robust and accurate surrogate method for monitoring the frequency and duration of combined sewer overflows. Environ Monit Assess 190(4):209
- Ivakhnenko AG, Ivakhnenko GA (1995) The review of problems solvable by algorithms of the group method of data handling (GMDH). Pattern Recognit Image Anal C/C Of Raspoznavaniye Obrazov I Analiz Izobrazhenii 5:527–535
- Jang JS (1993) ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 23(3):665–685
- Keshtegar B, Kisi O, Arab HG, Zounemat-Kermani M (2018) Subset modeling basis ANFIS for prediction of the reference evapotranspiration. Water Resour Manag 32(3):1101–1116
- Kisi O, Zounemat-Kermani M (2014) Comparison of two different adaptive neuro-fuzzy inference systems in modelling daily reference evapotranspiration. Water Resour Manag 28(9):2655–2675
- Mahdavi-Meymand A, Zounemat-Kermani M (2020) A new integrated model of the group method of data handling and the firefly algorithm (GMDH-FA): application to aeration modelling on spillways. Artif Intell Rev 53(4):2549–2569
- Mahdavi-Meymand A, Scholz M, Zounemat-Kermani M (2019) Challenging soft computing optimization approaches in modeling complex hydraulic phenomenon of aeration process. ISH J Hydraul Eng. <https://doi.org/10.1080/09715010.2019.1574619>
- May RWP, Brown PM, Hare GR, Jones KD (1989) Self-cleansing conditions for sewers carrying sediment. Hydraulic Research Ltd (Wallingford), Report SR 221
- May RW, Ackers JC, Butler D, John S (1996) Development of design methodology for self-cleansing sewers. Water Sci Technol 33(9):195–205
- Mo L, Xie L, Jiang X, Teng G, Xu L, Xiao J (2018) GMDH-based hybrid model for container throughput forecasting: selective combination forecasting in nonlinear subseries. Appl Soft Comput 62:478–490
- Mrugalski M, Witczak M (2002) Parameter estimation of dynamic GMDH neural networks with the bounded-error technique. J Appl Comput Sci 10(1):77–90
- Najafzadeh M, Laucelli DB, Zahiri A (2017) Application of model tree and evolutionary polynomial regression for evaluation of sediment transport in pipes. KSCE J Civ Eng 21(5):1956–1963
- Rajaei T, Nourani V, Zounemat-Kermani M, Kisi O (2010) River suspended sediment load prediction: Application of ANN and wavelet conjunction model. J Hydrol Eng 16(8):613–627

- Rajae T, Khani S, Ravansalar M (2020) Artificial intelligence-based single and hybrid models for prediction of water quality in rivers: a review. *Chemom Intell Lab Syst* 200:103978
- Roy DK, Barzegar R, Quilty J, Adamowski J (2020) Using ensembles of adaptive neuro-fuzzy inference system and optimization algorithms to predict reference evapotranspiration in subtropical climatic zones. *J Hydrol* 591:125509
- Safari MJS, Mohammadi M, Ab Ghani A (2018) Experimental studies of self-cleansing drainage system design: a review. *J Pipeline Syst Eng Pract* 9(4):04018017
- Sihag P, Esmailbeiki F, Singh B, Ebtehaj I, Bonakdari H (2019) Modeling unsaturated hydraulic conductivity by hybrid soft computing techniques. *Soft Comput* 23:12897–12910
- Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. *J Geophys Res* 106:7183–7192
- Vongvisessomjai N, Tingsanchali T, Babel MS (2010) Non-deposition design criteria for sewers with part-full flow. *Urban Water J* 7(1):61–77
- Wang H, Wang W, Cui L, Sun H, Zhao J, Wang Y, Xue Y (2018) A hybrid multi-objective firefly algorithm for big data optimization. *Appl Soft Comput* 69:806–815
- Yang XS (2009) Firefly algorithms for multimodal optimization. In: *International symposium on stochastic algorithms*. Springer, Berlin, pp 169–178
- Yaseen ZM, Ebtehaj I, Bonakdari H, Deo RC, Danandeh Mehr A, Mohtar WHMW, Diop L, El-shafie A, Singh VP (2017) Novel approach for streamflow forecasting using a hybrid ANFIS-FFA model. *J Hydrol* 554:263–276
- Zounemat-Kermani M (2012) Hourly predictive Levenberg–Marquardt ANN and multi linear regression models for predicting of dew point temperature. *Meteorol Atmos Phys* 117(3–4):181–192
- Zounemat-Kermani M (2017) Assessment of several nonlinear methods in forecasting suspended sediment concentration in streams. *Hydrol Res* 48(5):1240–1252
- Zounemat-Kermani M, Mahdavi-Meymand A (2019) Hybrid meta-heuristics artificial intelligence models in simulating discharge passing the piano key weirs. *J Hydrol* 569:12–21
- Zounemat-Kermani M, Matta E, Cominola A, Xia X, Zhang Q, Liang Q, Hinkelmann R (2020) Neurocomputing in surface water hydrology and hydraulics: a review of two decades retrospective, current status and future prospects. *J Hydrol* 588:125085

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.