



The prediction of longitudinal dispersion coefficient in natural streams using LS-SVM and ANFIS optimized by Harris hawk optimization algorithm



Naser Arya Azar^a, Sami Ghordoyee Milan^{b,*}, Zahra Kayhomayoon^c

^a Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

^b Department of Irrigation and Drainage Engineering, Aburahman Campus, University of Tehran, Tehran, Iran

^c Department of Geology, Payame Noor University, Tehran, Iran

ARTICLE INFO

Keywords:

Dispersion coefficient
Water quality modeling
Streamflow
Machine learning

ABSTRACT

Accurate calculation of the longitudinal dispersion coefficient (K_x) of pollution is essential in modeling river pollution status. Various equations are presented to calculate the K_x using experimental, analytical, and mathematical methods. Although machine learning models are more reliable than experimental equations in the presence of uncertainties missing data, they have not been widely used in predicting K_x . In this study, the K_x of the river was predicted using machine learning methods, including least square-support vector machine (LS-SVM), adaptive neuro-fuzzy inference system (ANFIS), and ANFIS optimized by Harris hawk optimization (ANFIS-HHO), and the results were compared with that of the experimental methods. Several scenarios were designed by different combinations of input variables, such as the average depth of the flow (H), average flow velocity (U), and shear velocity (u_s). The results showed that machine learning models had a more efficient performance to predict K_x than experimental equations. The ANFIS-HHO, with a scenario containing all the input variables, performed better than the other two models, with root mean square error, mean absolute percentage error, and coefficient of determination of 17.0, 0.22, and 0.97, respectively. Furthermore, the HHO algorithm slightly increased the prediction performance of the ANFIS. The discrepancy ratio (DR) evaluation criteria showed that experimental equations overestimated the values of K_x , while the machine learning models resulted in higher precision. Also, the results of Taylor's diagram showed the acceptable performance of the ANFIS-HHO model compared to other models. Given the promising results of the present study, it is expected that the proposed approach can be efficiently used for similar environmental modeling problems.

1. Introduction

In recent years, the protection of rivers has been considered by national and international organizations responsible for quality control and protection of water resources. The importance of river protection reveals in regions where large cities and industries are located near the rivers, and rivers are the main supplier of the water needs (Tayfur and Singh, 2005). Increasing the level of pollution in surface waters necessitates mixing and reduction processes in natural rivers. One of the most

important and efficient river environmental management methods is to improve the self-cleaning ability of the river (Pourabadei and Kashefpour, 2007). Currently, agricultural and industrial wastewater entrance to natural rivers has become a common management practice in environmental engineering because of its oxidization and organic matter removal capabilities. To control the quality of surface water resources, pollutants' entrance to natural rivers and open streams is conducted using costly methods. This requires accurate knowledge about the pollutant transfer to the rivers and the possibility of transmission and mixing and self-cleaning of pollutants by river flow (Pourabadei and

Abbreviations: ACOR, Ant colony optimization algorithm; ADE, Advection-dispersion equation; ANFIS, Adaptive neuro-fuzzy inference system; ANN, Artificial neural network; BA, Bee algorithm; BN, Bayesian network; CS, Cuckoo search; DE, Differential evolutionary; DR, Discrepancy ratio; EPR, Evolutionary Polynomial Regression; FM, Fuzzy Model; GA, Genetic algorithm; GEP, Gene-expression programming; GMDH, Group method of data handling; GP, Genetic programming; HHO, Harris hawks optimization; ICA, Imperialist competitive algorithm; LS-SVM, Least squares support vector machine; M5, Model tree; MAPE, mean absolute percentage error; MARS, multivariate adaptive regression spline; Max, Maximum; Min, Minimum; MT, Model Tree; NF, Neuro fuzzy; POMGGP, Pareto-Optimal-Multigene Genetic Programming; PSO, particle swarm optimization; RMSE, Root mean square error; SD, Standard deviation; SVM, Support vector machine; WOA, Whale optimization algorithm.

* Corresponding author.

E-mail address: s.milan@ut.ac.ir (S. Ghordoyee Milan).

Nomenclature	
A	Cross-sectional area of the flow
b	Regression bias
B	Water surface width
C	Cross-sectional concentration
e	Error
E	Indicates the escaping energy of the prey
E_0	Randomly changes inside the interval $(-1,1)$ in the HHO
g	Acceleration due to gravity
H	The average depth of the flow
J_{rabbit}	Represents the random jump in the HHO
K_x	Longitudinal dispersion coefficient
LB	Upper bound
N	Denotes the total number of hawks
n	Number of samples
p	The consequent model parameters (ANFIS)
q	The consequent model parameters (ANFIS)
r_1, r_2, r_3, r_4	Random numbers between 0 and 1, in the HHO
R^2	Coefficient of determination
T	Weight
T	The maximum number of iterations in the HHO
U	Average flow velocity
u^*	Shear velocity
UB	Upper bound
W	Transpose operator, respectively
x	Input variable (ANFIS)
$X(t+1)$	The position vector of hawks in the next iteration
x_o	Observed value
x_p	Predicted value (ANFIS)
$X_{rabbit}(t)$	The position of rabbit
y	Input variable
β	Equal to 1.0 in triangular channels
β	is a default constant set to 1.5 (in the HHO)
γ	The regulation term for the error
$\Delta X(t)$	Difference between the position vector of the rabbit and the current location in iteration t
σ	The channel sinuosity
$\varphi(x)$	the nonlinear mapping of inputs to a high dimensional feature space

Kashefpour, 2007). The process of pollutant transfer can be seen in three stages. In the first stage, a pollutant is diluted due to its initial movement in the channel. In the second stage, the pollutants are mixed because of the turbulent transfer processes throughout the river section. In the third stage, after completing the cross-section, longitudinal dispersion tends to clear the longitudinal changes in pollutant

concentrations (French, 1986).

The longitudinal dispersion coefficient (K_x) of pollution is an essential variable in modeling the pollution status of rivers. Accurate calculation of K_x is required in several applied hydraulic problems, such as river engineering, environmental engineering, river mouth problems, and risk assessment of injection of dangerous pollutants into rivers

Table 1
Experimental equations used in previous research for the calculation of K_x .

No	Author	Equation	Reference
1	Elder (1959)	$K_x = 5.93HU$	Tayfur and Singh (2005)
2	McQuivey and Keefer (1974)	$K_x = 0.58 \left(\frac{H}{u^*} \right)^2 UB$	Deng et al. (2001)
3	Fisher et al. (1979)	$K_x = 0.011 \frac{U^2 B^2}{Hu^*}$	Fisher et al. (1979)
4	Li et al. (1998)	$K_x = 0.55 \frac{Bu^*}{H^2}$	Seo and Bake (2004)
5	Liu (1977)	$K_x = 0.18 \left(\frac{H}{u^*} \right)^{0.5} \left(\frac{B}{H} \right)^2 Hu^*$	Seo and Bake (2004)
6	Iwasa and Aya (1991)	$K_x = 2.0 \left(\frac{B}{H} \right)^{1.5} Hu^*$	Tavakollizadeh and Kashefpur (2007)
7	Seo and Bake (2004)	$K_x = 5.92 \left(\frac{U}{u^*} \right)^{1.43} \left(\frac{B}{H} \right)^{0.62} Hu^*$	Seo and Cheong (1998)
8	Koussis and Rodriguez-Mirasol (1998)	$K_x = 0.6 \left(\frac{B}{H} \right)^2 Hu^*$	Sedighnezhad et al. (2007)
9	Li et al. (1998)	$K_x = 0.2 \left(\frac{U}{u^*} \right)^{1.2} \left(\frac{B}{H} \right)^{1.3} Hu^*$	FaghforMaghrebi and Givehchi (2007)
10	Deng et al. (2001)	$K_x = \frac{0.15}{8\epsilon_t} \left(\frac{U}{u^*} \right)^2 \left(\frac{B}{H} \right)^{1.67} Hu^*$ $\epsilon_t = 0.145 + \frac{1}{3520} \left(\frac{U}{u^*} \right) \left(\frac{B}{H} \right)^{1.38}$	Deng et al. (2001)
11	Sahay and Dutta (2009)	$K_x = 10.612 \left(\frac{U}{u^*} \right) UH$	Deng et al. (2001)
12	Tavakollizadeh and Kashefpur (2007)	$K_x = 7.428 + 1.775 \left(\frac{U}{u^*} \right)^{1.572} \left(\frac{B}{H} \right)^{0.62} HU$	Kashefpur and Falconer (2002)
13	Sahay and Dutta (2009b)	$K_x = 2 \left(\frac{B}{H} \right)^{0.96} \left(\frac{U}{u^*} \right)^{1.26} Hu^*$	Tavakollizadeh and Kashefpur (2007)
14	Zeng and Huai (2014)	$K_x = 5.4 \left(\frac{U}{u^*} \right)^{0.13} 5.4 \left(\frac{B}{H} \right)^{0.7}$	Zeng and Huai (2014)

Table 2Machine learning methods used in previous studies for the calculation of K_x .

No	Reference	Method	No	Reference	Method
1	Tayfur and Singh (2005)	ANN	15	Toprak et al. (2014)	ANN
2	Toprak and Savci (2007)	FM	16	Sahay (2013)	GA
3	Toprak and Cigizoglu (2008)	ANN	17	Parsaie and Haghabi (2015)	ANN
4	Tayfur (2009)	GA	18	Sattar and Gharabaghi (2015)	GEP
5	Riahi-Madvar et al. (2009)	ANFIS	19	Noori et al. (2016)	ANN, ANFIS, SVM
6	Sahay and Dutta (2009)	GA	20	Najafzadeh and Tafarojnoruz (2016)	NF-GMDH-PSO, GA, MT, ANN, DE
7	Riahi-Madvar et al. (2009)	ANFIS	21	Wang et al. (2017)	GP
8	Noori et al. (2009)	SVM, ANFIS	22	Alizadeh et al. (2017a, 2017b, 2017c)	GA, ICA, BA, CS, ANN
9	Adarsh (2010)	SVM, ANN	23	Alizadeh et al. (2017a, 2017b, 2017c)	PSO
10	Azamathulla and Ghani (2011)	GP	24	Alizadeh et al. (2017a, 2017b, 2017c)	BN, ANN
11	Azamathulla and Wu (2011)	SVM	25	Seifi and Riahi-Madvar (2019)	GA, ANFIS-GA
12	Etemad-Shahidi et al. (2012)	M5	26	Balf et al. (2018)	EPR
13	Azamathulla and Ahmad (2012)	GEP	27	Riahi-Madvar et al. (2019)	POMGGP
14	Toprak et al. (2014)	ANN	28	Memarzadeh et al., 2020	

(Sedighnezhad et al., 2007; Seo and Bake, 2004). Various equations have been proposed to calculate the K_x using experimental, analytical, and mathematical methods. K_x can be easily determined when actual field data are available. Experimental equations are used in rivers where the characteristics of pollution dispersion are unclear (Kashefipur and Falconer, 2002). Analytical and mathematical methods have high computational complexities and are often time-consuming. Experimental methods are of the research methods used to test hypotheses. They aim to investigate the effect of specific stimuli, methods, or environmental conditions on a group of subjects. One of the characteristics of the experimental methods is that they provide results by manipulating the variables and conditions obtained for the group selected by random assignments. In general, these methods result in significant errors in the prediction of K_x . In this regard, the researchers have presented relationships that depend on the river's hydraulic parameters, namely, flow velocity, river shear velocity, flow depth, and river width. These relationships are valid in a specific hydraulic condition and do not provide reliable results for outlier values (Azamathulla and Wu, 2011). Table 1 summarizes the available experimental equations for the calculation of K_x based on the average depth of the flow (H), average flow velocity (U), shear velocity (u_*), and water surface width (B). Elder (1959) introduced the most straightforward equation, which relates the K_x to H and U . Later, researchers have developed equations that use more parameters such as u_* and B . Deng et al. (2001) introduced a new equation in which K_x depends not only on the geometric and hydraulic characteristics of the flow but also on the transverse dispersion coefficient (ϵ_t). The latter also depends on the hydraulic and geometric parameters of the flow.

Machine learning models have gathered considerable attention due to not requiring the physical information of the channels and their physical structure. Researchers have extensively used these models in recent years due to the lack of difficulties that occur in numerical models and not having the limitations and high costs of experimental models. Artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS), least square-support vector machine (LS-SVM), and Bayesian networks (BN) are among these models. These models have been applied for the assessment of sediment transport, demonstrating acceptable results (Azamathulla et al., 2012; Azamathulla et al., 2009; Afan et al., 2015; Kumar et al., 2016; Mount and Stott, 2008; Milan et al., 2018; Najafzadeh and Oliveto, 2020; Saberi-Movahed et al., 2020; Najafzadeh and Ghaemi, 2019; Najafzadeh and Etemad-Shahidi, 2016; Balf et al., 2018).

In recent years, researchers have used machine learning-based methods to predict K_x . In a leading study, Tayfur and Singh (2005) predicted K_x using data of 71 samples from 29 rivers in the United States using ANN and showed that ANN had an acceptable ability to estimate the amount of K_x . Riahi-Madvar et al. (2009) used the ANFIS approach to predict K_x based on 73 data samples, of which 70% of the data were

used for training and the rest for the test. Error evaluation criteria showed that ANFIS had better performance than experimental models and could be combined with mathematical models. From 2009 to 2016, besides ANN and ANFIS, other models such as SVM and GEP were used to predict K_x (Table 2). For the first time, Najafzadeh and Tafarojnoruz (2016) used the NF-GMDH-PSO hybrid model and the MT, ANN, DE, and GA models to predict K_x . They used 233 data samples, of which 75% of the data were selected randomly for training and the remaining data for the test. The models' performance was acceptable, and the NF-GMDH-PSO hybrid model was capable of predicting K_x with the highest performance.

Using the ability of evolutionary algorithms, some researchers have developed empirical relationships in this field (Tayfur, 2009; Sahay and Dutta, 2009; Sahay, 2013). In the field of evolutionary algorithms, Alizadeh et al. (2017a, 2017b, 2017c) used the PSO algorithm to derive new equations to predict K_x in natural rivers. Based on their study, PSO could improve the performance of predictive equations by finding the optimal values of the coefficients. In another study, Memarzadeh et al. (2020) used the WOA evolutionary algorithm to find the optimal coefficients and showed that the algorithm improved the performance of predictive equations.

LS-SVM models have also been used in various studies, including land temperature prediction, water quality, satellite images, and sensor development showing an excellent performance (Chauchard et al., 2004; Zheng et al., 2008; Leong et al., 2019; Najafzadeh and Ghaemi, 2019; Zhou et al., 2020). Compared with the classical SVM method, LS-SVM has less computation complexity leading to higher accuracy and lower consuming time. In a study, Kong et al. (2008) employed LS-SVM and wavelet transform to predict daily river flow. Their results indicated the appropriate accuracy of the LS-SVM model in flow prediction.

Najafzadeh and Ghaemi (2019) compared the LS-SVM and MARS models with ANN and ANFIS to predict water quality indicators. They used 200 data sets, considering nine input variables, and concluded that LS-SVM and MARS performed better than the other models.

ANFIS is another model useful for data with uncertainties. Its relatively low accuracy in the modeling and being occasionally trapped in the local minima (Asefpour Vakilian and Massah, 2016; Najafzadeh et al., 2016) necessitate the use of advanced optimization methods, such as particle swarm optimization (PSO) and differential evolution (DE), to optimize the coefficients of the ANFIS model, and therefore, increase its performance. Ebtehaj et al. (2019) used the PSO algorithm to train an ANFIS model to simulate sediment transport in open channels. The results indicated the acceptable performance of the ANFIS-PSO hybrid model compared to the ANFIS model.

Azad et al. (2018) used the ANFIS model and the GA, PSO, ACOR, and DE evolutionary algorithms to simulate the quality parameters of the Gorganrood river, Iran. Their results showed that the DE algorithm is more accurate than the other evolutionary algorithms. In another study,

Table 3

Statistical characteristics of the dataset used in this study.

Parameter	B (m)	H (m)	U (m/s)	u_* (m/s)	B/H	U/u_*	β	σ	K_x (m^2/s)
Min	11.90	0.22	0.034	0.002	13.8	1.29	2.62	1.08	1.90
Max	711.2	19.9	1.74	0.553	156.5	19.63	5.05	2.54	892
Average	82.97	1.71	0.54	0.087	51.68	7.623	3.79	1.54	107
SD	121.6	2.58	0.38	0.088	31.49	4.527	0.56	0.43	169

they used these algorithms to model rainfall-runoff in Isfahan, Iran (Azad et al., 2019). They showed that the ANFIS-ACOR method resulted in the best performance with R^2 , RMSE, and SI of 0.92, 2.73, and 0.26, respectively. Yang et al. (2019) used ANFIS-GA and ANFIS-PSO hybrid models to predict landslides. The ANFIS-GA and ANFIS-PSO models could predict ground vibration with acceptable accuracy (Yang et al., 2019).

Seifi and Riahi-Madvar (2018) used 505 data, divided into training (60%) and test (40%) samples, to predict K_x . They used the ANN and ANFIS models, along with GA evolutionary algorithm, to improve the performance of them. The results showed that GA remarkably improved the performance of ANFIS so that ANFIS-GA predicted the value of K_x more accurately than other models.

Proposed by Heidari et al. (2019), the Harris hawks optimization (HHO) is one of the most-recently algorithms available for optimization problems. In this research, HHO is used to improve the training capability of ANFIS in predicting K_x in natural streams amount.

The literature review on the application of machine learning in predicting the value of K_x shows that it is possible to increase the prediction performance with advanced machine learning models. LS-SVM is one of the models that has shown promising results in various studies, and therefore, its use in predicting K_x can be efficient along with models such as ANFIS. Also, due to the poor performance of ANFIS, especially for the test data, in some cases, evolutionary optimization algorithms can be used to improve its results. In this study, an attempt has been made to use one of the most advanced evolutionary algorithms to improve the results of machine learning algorithms. Since machine learning models are not sensitive to the nature of the input data, it has been tried in this study to create various scenarios by different combinations of input variables, i.e., B , H , U , u_* , $\frac{B}{H}$, $\frac{U}{u_*}$, β , and σ , to determine the most appropriate scenario for estimating K_x . Moreover, the most suitable machine learning model is selected based on RMSE, MAPE, R^2 ,

and Taylor's diagram. The selected models, along with the selected input scenario, are compared with the results of some experimental equations.

2. Materials and methods

2.1. Advection-dispersion equation (ADE)

K_x is an essential factor in the prediction of the concentration, which was introduced by Taylor (1954) in a one-dimensional (1D) environment (Eq. 1).

$$\frac{\partial C}{\partial t} + U \frac{\partial C}{\partial x} = K_x \frac{\partial^2 C}{\partial x^2} \quad (1)$$

where C is cross-sectional concentration, t is time, and x is the longitudinal direction throughout the flow. Fisher (1967) developed Eq. (2) to calculate the K_x for pollutant transfer (Tayfur and Singh, 2005).

$$K_x = -\frac{1}{A} \int_0^B h u' \int_0^Y \frac{1}{\varepsilon_t h} \int_0^Y h u' dy dy dy \quad (2)$$

where A is the cross-sectional area of the flow, h is the local depth of the flow, u' is the deviation of the speed at depth from the average speed, and y is the coordinates in the lateral direction. Eq. (2) is the basis of many experimental methods for determining K_x , the most important of which are presented in Table 1. Fisher (1967) proposed a simpler equation (Eq. 3) for Taylor's method.

$$K_x = 0.011 \frac{U^2 B^2}{H u_*} \quad (3)$$

Fisher et al. (1979) presented Eq. (4) to calculate ε_t for the wide and direct rivers with the uniform flow and constant depth in width (Kashefpur and Falconer, 2002).

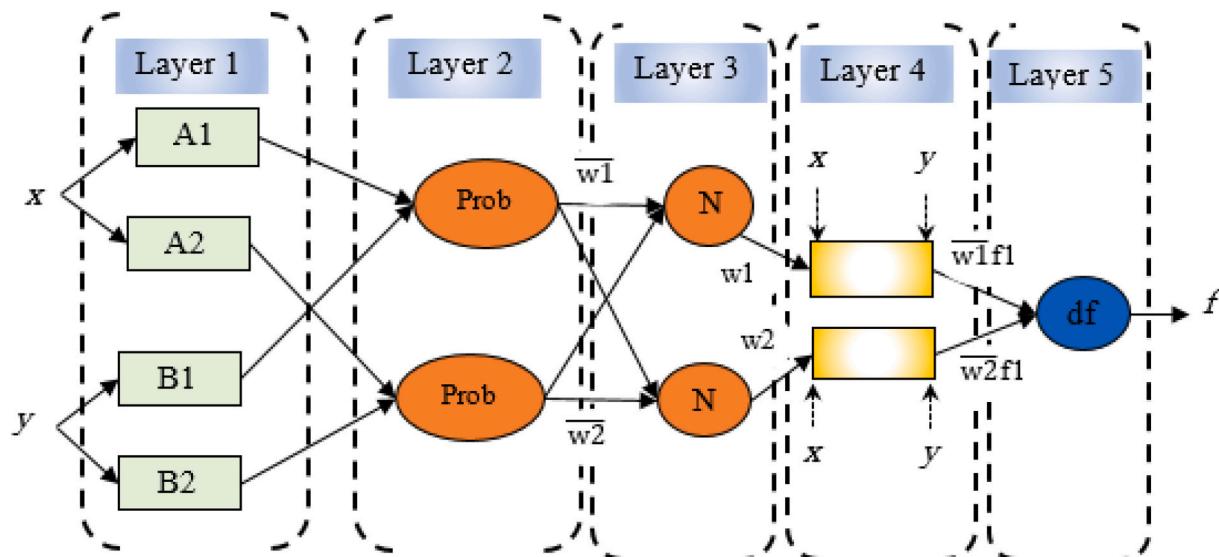


Fig. 1. Adaptive neuro-fuzzy inference system (ANFIS).

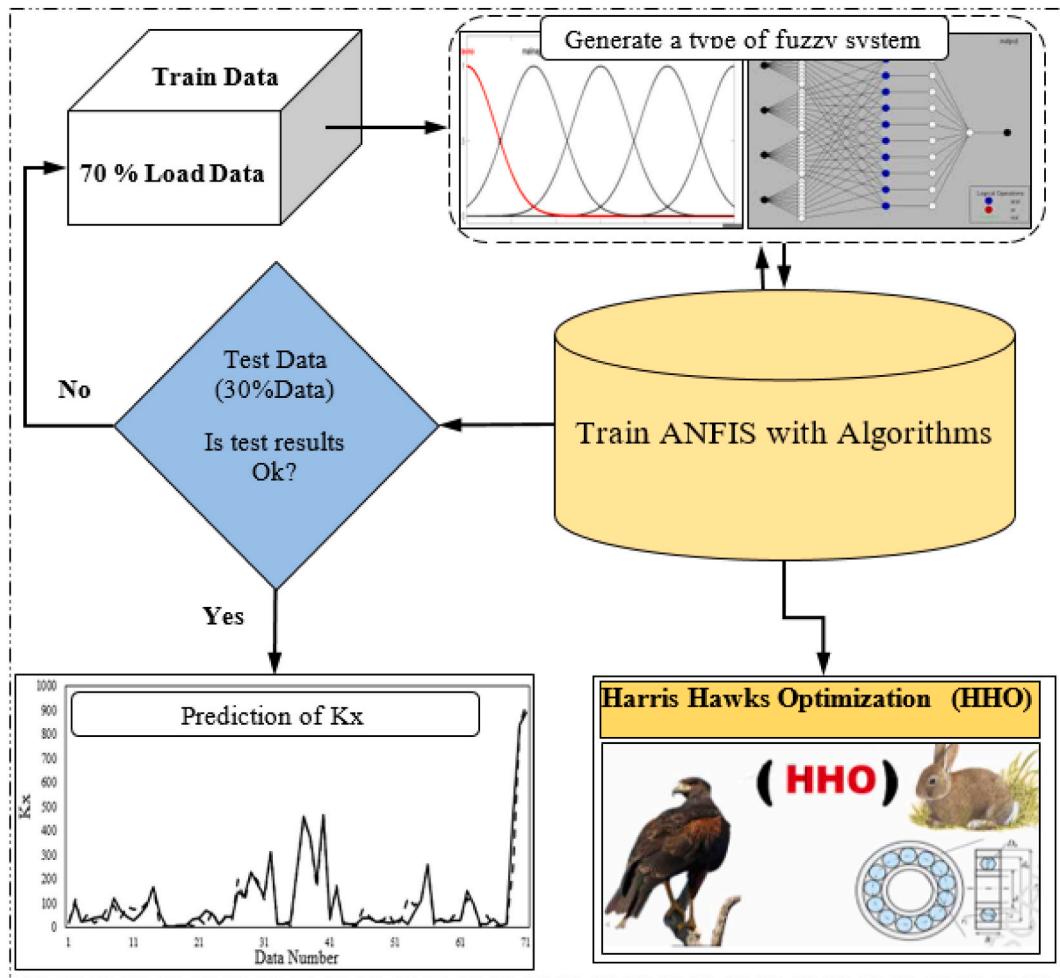


Fig. 2. The HHO – ANFIS model.

$$\epsilon_r = 0.15 Hu^*$$

2.2. Dataset

Data from 29 rivers in the United States, including 71 data samples, were extracted from Deng et al. (2001). They were used to analyze the experimental equations and construct the machine learning models. Table 3 summarizes the statistical characteristics of the dataset used in this study. It is observed that the K_x of the rivers ranged from 1.90 to 892 m^2/s , while its average value was ca. 107 m^2/s . β indicates the channel shape, introduced by Deng et al. (2001):

$$\beta = \ln\left(\frac{B}{H}\right) \quad (5)$$

β is equal to 1.0 in triangular channels. Otherwise, it has either a parabolic shape ($\beta = 2$), having a flatbed region with two curving banks ($2 < \beta < 5$), or a rectangular shape ($\beta > 5$) (Deng et al., 2001). σ is the channel sinuosity, which is defined as the ratio of the channel length to the valley length (Chang, 1988). The sinuosity of the streams and rivers varies from 1.08 to 2.54.

2.3. Least squares support vector machine (LS-SVM)

Suykens and Vandewalle (1999) introduced LS-SVM, which is very similar to the classical SVM but has lower computation and time complexity and higher performance. Given a set of training data such as

$\{(x_k, y_k)\}_{k=1}^N$, whose input and output data include $x_k \in R^N$ and $y_k \in R$, respectively, Eq. (6) shows the nonlinear regression function in the initial weighting (Suykens and Vandewalle, 1999):

$$y(x) = W^T \varphi(x) + b \quad (6)$$

where T , b , and W are weight, regression bias, and transpose operator, respectively. $\varphi(x)$ is the nonlinear mapping of inputs to a high dimensional feature space. This nonlinear regression can be solved using an optimization process (Eq. 7).

$$\min_j(w, e) = \frac{1}{2} W^2 W + \frac{1}{2} \gamma \sum_{k=1}^N e_k^2 \quad (7)$$

with the constraint

$$y_k = W^T \varphi(x) + b + e_k, k = 1, \dots, N \quad (8)$$

and γ is the regulation term for the error e . γ controls the approximation function, so the larger the γ value, the higher the error is. Solving this equation using the Lagrangian form of the main objective function, we have

$$L(w, b, e, a) = j(w, e) - \sum_{i=1}^N \alpha_i \{ W^T \varphi(x_k) + b + e_k - y_k \} \quad (9)$$

where α_i is the Lagrangian coefficient. Based on the Karush-Kuhn-Tucker

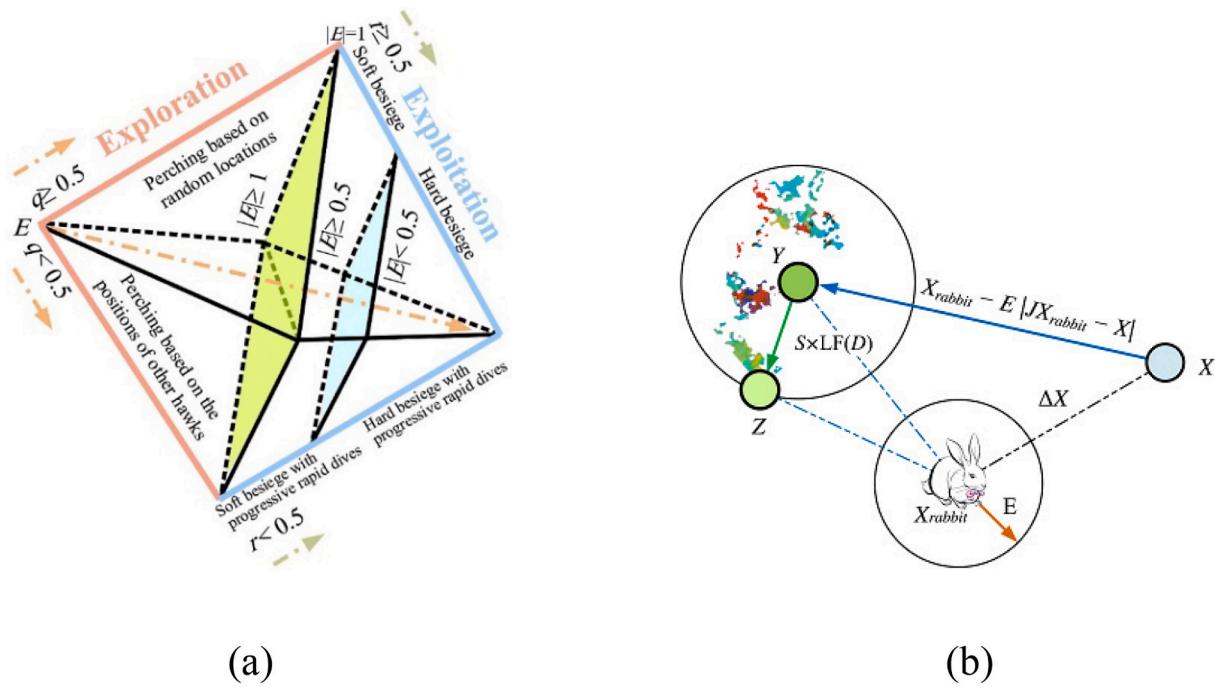


Fig. 3. (a) Procedure of Harris hawks optimization, (b) besieging the prey by Harris hawks.

condition, the LS-SVM model is written for the approximation function as Eq. (10) (Mellit et al., 2013).

$$y(x) = \sum_{k=1}^N \alpha_k K(x, x_k) + b \quad (10)$$

where $K(x, x_k)$ is called kernel function. In this study, the Gaussian kernel function is used (Eq. 6) (Mellit et al., 2013):

$$K(x, x_k) = \exp\left(-\frac{\|x - x_k\|^2}{\sigma^2}\right) \quad (11)$$

2.4. Adaptive neuro-fuzzy inference system (ANFIS)

The ANFIS model was developed by combining a fuzzy inference system (FIS) and ANN (Jang, 1993). ANFIS is a fuzzy rule-based network that uses adaptive systems to facilitate the learning and adaptation process (Ozgan et al., 2009). The main purpose of the ANFIS is to optimize the FIS parameters by using input-output data sets through a learning algorithm. In this structure, the nodes in the first and last layers represent the input and output data, while the nodes in the hidden layers are identified as membership functions and rules. The Sugeno model is widely used in this structure because of its interpretability, high computational capability, and adaptability. For two inputs (x_1, x_2) and one output (y) of the Sugeno type ANFIS structure, the rules are defined as Eqs. (12) and (13).

$$\text{Rule 1. If } (x_1 \text{ is } A_1) \text{ and } (x_2 \text{ is } B_1) \text{ then } f_1 = p_1 x_1 + q_1 x_2 + r_1 \quad (12)$$

$$\text{Rule 2. If } (x_1 \text{ is } A_2) \text{ and } (x_2 \text{ is } B_2) \text{ then } f_2 = p_2 x_1 + q_2 x_2 + r_2 \quad (13)$$

where A and B are the fuzzy sets, p, q , and r are the consequent model parameters that are determined in the training stage. ANFIS architecture includes five layers (Fig. 1). In the first layer, the input data passes through different membership functions, and the membership degree of input nodes to different fuzzy intervals is determined using membership functions. In the second layer, which contains the rule nodes, fuzzy

values are multiplied by each node, and the result is the weight of the rules. The nodes of the third layer normalize the weight of the rules. The resultant nodes create the fuzzy-based rule outputs. The fifth layer consists of a single node that calculates the total output of the system. This layer transforms the results of each fuzzy rule into a non-fuzzy output using a defuzzification process.

2.5. Development of the ANFIS using Harris hawks optimization

Neuron-based methods suffer from several drawbacks, such as trapping in local minima and slow convergence of training, especially for wide search spaces (Asefpour Vakilian and Massah, 2018). Therefore, there is a need to use hybrid methods that use optimization techniques to optimize the parameters of the ANFIS.

In this study, the HHO evolutionary algorithm was used along with the ANFIS model to improve the ANFIS model. According to the literature, the HHO algorithm has its unique features that can significantly improve the traditional ANFIS model. The relationships of hybrid ANFIS-HHO used in this study to predict the K_x is depicted in Fig. 2. Firstly, scenarios with different variables were given to the model as input, and then the type of the fuzzy function and the ANFIS structure is determined. The ANFIS is improved with evolutionary algorithms to achieve better results. In this structure, the objective function is to minimize the error of the predicted values. Finally, the amount of K_x is predicted by the optimized model.

Like many other evolutionary algorithms, HHO is a nature-inspired and population-based algorithm, which is based on rabbit hunting by Harris hawks (Heidari et al., 2019). In the HHO method, the Harris hawks are the candidate solutions (Fig. 3a). They perch randomly on some locations and wait to detect prey based on two strategies. Their position is mathematically expressed as Eq. (14)

$$X(t+1) = \begin{cases} X_{rand}(t) - r_1 |X_{rand}(t) - 2r_2 X(t)| & q \geq 0.5 \\ (X_{rabbit}(t) - X_m(t)) - r_3(LB + r_4(UB - LB)) & q < 0.5 \end{cases} \quad (14)$$

where $X(t+1)$ is the position vector of hawks in the next iteration of t , $X_{rabbit}(t)$ is the position of rabbit, $X(t)$ is the current position vector of

Table 4
Scenarios and their input variables for the prediction of K_x .

Scenarios	Input variables							
	B	H	U	u^*	B/H	U/u^*	β	σ
Complete scenario	✓	✓	✓	✓	✓	✓	✓	✓
Scenario 1	✓	✓	✓	✓	✓	✓	✓	✗
Scenario 2	✓	✓	✓	✓	✓	✓	✗	✓
Scenario 3	✓	✓	✓	✓	✗	✗	✓	✓
Scenario 4	✓	✓	✓	✓	✓	✗	✓	✗
Scenario 5	✓	✓	✓	✓	✗	✗	✗	✓
Scenario 6	✓	✓	✓	✓	✓	✗	✗	✗
Scenario 7	✓	✓	✓	✓	✗	✗	✗	✗
Scenario 8	✓	✓	✓	✗	✗	✗	✗	✗
Scenario 9	✓	✓	✗	✗	✗	✗	✗	✗

Table 5
Optimum values of the LS-SVM parameters.

Mode	Kernel function	σ^2	γ
LS-SVM	Gaussian	5.365	136.03

hawks, r_1 , r_2 , r_3 , r_4 , and q are random numbers between 0 and 1, which are updated in each iteration, LB and UB are the lower and upper bounds of variables, $X_{rand}(t)$ is a randomly selected hawk from the current population, and X_m is the average position of the current population of hawks. The average position of hawks is obtained using Eq. (15)

$$X_m(t) = \frac{1}{N} \sum_{i=1}^N X_i(t) \quad (15)$$

where $X_i(t)$ indicates the location of each hawk in iteration t and N denotes the total number of hawks. The energy of prey decreases considerably during the escape. The energy of prey is modeled using Eq. (16)

$$E = 2E_0 \left(1 - \frac{t}{T}\right) \quad (16)$$

where E indicates the escaping energy of the prey, T is the maximum number of iterations, and E_0 is the initial energy. E_0 randomly changes inside the interval $(-1, 1)$ at each iteration.

The hawks intensify the besiege process to catch the exhausted prey effortlessly. The E parameter is utilized to model this strategy and enable the HHO to switch between soft and hard besiege processes. In this regard, when $|E| \geq 0.5$, the soft besiege happens, and when $|E| < 0.5$, the hard besiege occurs (Fig. 3b).

If $r \geq 0.5$ and $|E| \geq 0.5$, the rabbit still has enough energy and tries to escape by some random misleading jumps, but finally, it cannot. During these attempts, the Harris hawks encircle it softly to make the rabbit more exhausted and then perform the surprise pounce. This behavior is modeled by Eqs. (17) and (18)

$$X(t+1) = \Delta X(t) - E|JX_{rabbit}(t) - X(t)| \quad (17)$$

$$\Delta X(t) = X_{rabbit}(t) - X(t) \quad (18)$$

where $\Delta X(t)$ is the difference between the position vector of the rabbit

Table 6
Evaluation criteria of the LS-SVM model for the prediction of K_x .

Evaluation Criteria	Scenario										
		Complete Scenario	1	2	3	4	5	6	7	8	9
RMSE	Training	12.74	14.21	14.19	12.4	17.3	18.4	23.0	31.46	30.0	45.1
	Test	28.4	20.6	33.2	32.3	37.3	74.1	191	260	132	154
MAPE	Training	0.06	0.33	0.13	0.32	1.34	0.43	1.30	0.50	0.50	1.00
	Test	2.54	0.42	3.34	0.52	1.10	0.54	0.66	1.40	1.43	0.90
R^2	Training	0.99	0.99	0.99	0.98	0.78	0.99	0.80	0.96	0.94	0.62
	Test	0.88	0.96	0.81	0.80	0.89	0.86	0.83	0.57	0.89	0.57

and the current location in iteration t , r_5 is a random number between 0 and 1, and J_{rabbit} represents the random jump strength of the rabbit throughout the escaping procedure.

If $r \geq 0.5$ and $|E| < 0.5$, the prey is exhausted, and it has low escaping energy. Besides, Harris hawks hardly encircle the prey to perform the surprise pounce finally (Fig. 3a). In this situation, the current positions are updated using Eq. (19)

$$X(t+1) = X_{rabbit}(t) - E|\Delta X(t)| \quad (19)$$

2.6. Performance evaluation criteria

The dataset was randomly divided into two groups: 70% of the data for the training of the models and the remaining for the test. Root mean

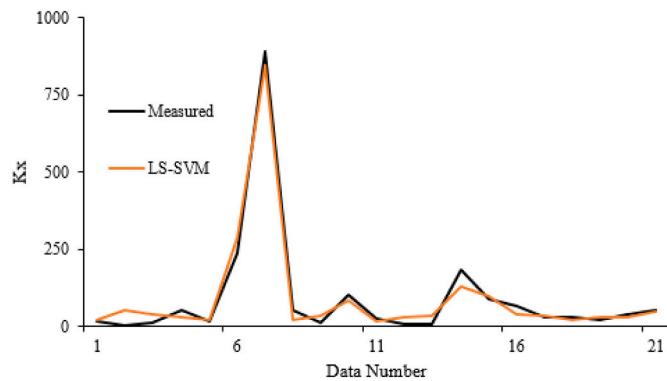


Fig. 4. The measured and predicted values of K_x for the LS-SVM model.

Table 7
Parameters of the ANFIS model.

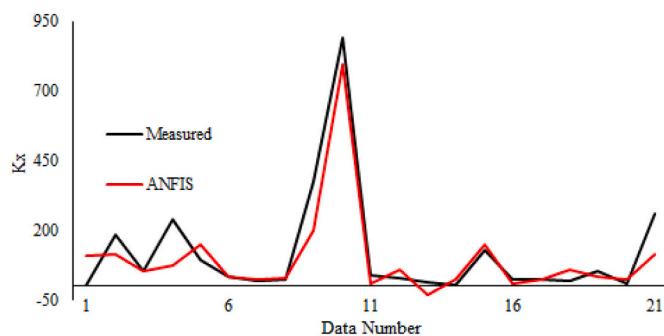
Parameter	Value
Fuzzy structure	Sugeno-type
Initial FIS for training	Genfis3
MF type	Gaussian
Output MF	Linear
Optimization method	Hybrid
Number of fuzzy rules	10
Maximum number of epochs in training	1200

Table 8
Parameters of the ANFIS-HHO model.

Parameter	Value
Number of search agent	30
Iteration number	1500
β	1.5
Range partitions (weights and biases)	[-3, +3]
Population size	30

Table 9Evaluation criteria of the ANFIS and ANFIS-HHO models for the prediction of K_x .

Evaluation Criteria		Scenario									
		Complete Scenario	1	2	3	4	5	6	7	8	9
ANFIS	Training	8.23	9.02	7.02	10.81	12.62	15.40	13.65	9.5	20.60	28.71
	Test	73.4	81.0	159	387	201	356	433	385	440	530
MAPE	Training	0.07	0.09	0.08	0.16	0.14	0.17	0.1	0.06	0.07	0.18
	Test	2.53	5.37	6.02	2.13	1.94	2.00	7.64	3.98	5.5	8.30
R^2	Training	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.98	0.98	0.70
	Test	0.87	0.50	0.07	0.2	0.30	0.55	0.26	0.18	0.12	0.03
ANFIS-HHO											
RMSE	Training	13.47	18.2	16.2	15.2	18.0	15.8	19.0	20.0	28.2	56.0
	Test	14.63	32.0	38.3	43.5	45.6	27.2	47.4	39.0	42.8	68.6
MAPE	Training	0.21	1.0	1.01	1.03	0.49	0.54	0.72	0.75	0.86	3.3.0
	Test	0.24	0.55	1.01	0.98	1.11	1.26	0.38	0.83	0.66	1.30
R^2	Training	0.98	0.92	0.90	0.93	0.87	0.89	0.68	0.89	0.70	0.52
	Test	0.97	0.89	0.85	0.86	0.79	0.92	0.65	0.71	0.65	0.48

**Fig. 5.** The measured and predicted values of K_x for the ANFIS model.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (x_o - x_p)^2}{n}} \quad (20)$$

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{x_o - x_p}{x_o} \right| \quad (21)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (x_p - x_o)^2}{\sum_{i=1}^n (x_o - \bar{x}_o)^2} \quad (22)$$

Where x_o is the observed (measured) value, x_p is the predicted value, and n is the number of samples. The lower RMSE, MAPE, and higher R^2 values present the better performance of the model.

3. Results and discussion

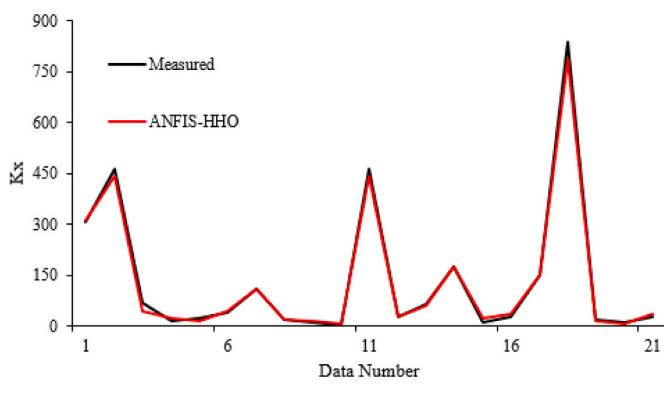
In this section, firstly, the results of the machine learning models are described and analyzed, and then, the results of the machine learning models are compared with that of the experimental equations.

Since several variables are effective in the prediction of K_x , 10 scenarios were defined with different combinations of input variables, according to Table 4. The complete scenario included all the input variables, i.e., B , H , U , u^* , $\frac{B}{H}$, $\frac{U}{u^*}$, β , and σ . Each scenario was used separately to train the machines, and the performances of the machines were compared to obtain the most efficient scenario.

3.1. Results of the LS-SVM model

Table 5 shows the optimum values of the LS-SVM parameters. The Gaussian kernel in the LS-SVM method has two parameters, namely, σ^2 and γ , the optimum value of which was obtained equal to 5.365 and 136.03, respectively.

According to Table 6, the scenarios that included more input variables had better accuracy than the scenarios with three or four input variables, except Scenario 8, which included only three input variables

**Fig. 6.** The measured and predicted values of K_x for the ANFIS-HHO model.

square error (RMSE), mean absolute percentage error (MAPE), and coefficient of determination (R^2) were used to evaluate the scenarios and machine learning methods.

Table 10Evaluation criteria of the machine learning models and experimental equations for the prediction of K_x .

Evaluation Criteria	Equation						ML Model		
	Fisher et al. (1979)	Seo and Bae (2004)	Tavakollizadeh and Kashefipur (2007)	Sahay and Dutta (2009)	Zeng and Huai (2014)	Sahay and Dutta (2009b)	LS-SVM	ANFIS	ANFIS-HHO
RMSE	1708	439.80	3734	283	163	369	18.3	49.1	17.0
MAPE	2.67	1.28	6.48	0.85	3.45	1.07	0.26	0.86	0.22
R^2	0.37	0.45	0.47	0.43	0.40	0.45	0.98	0.91	0.97

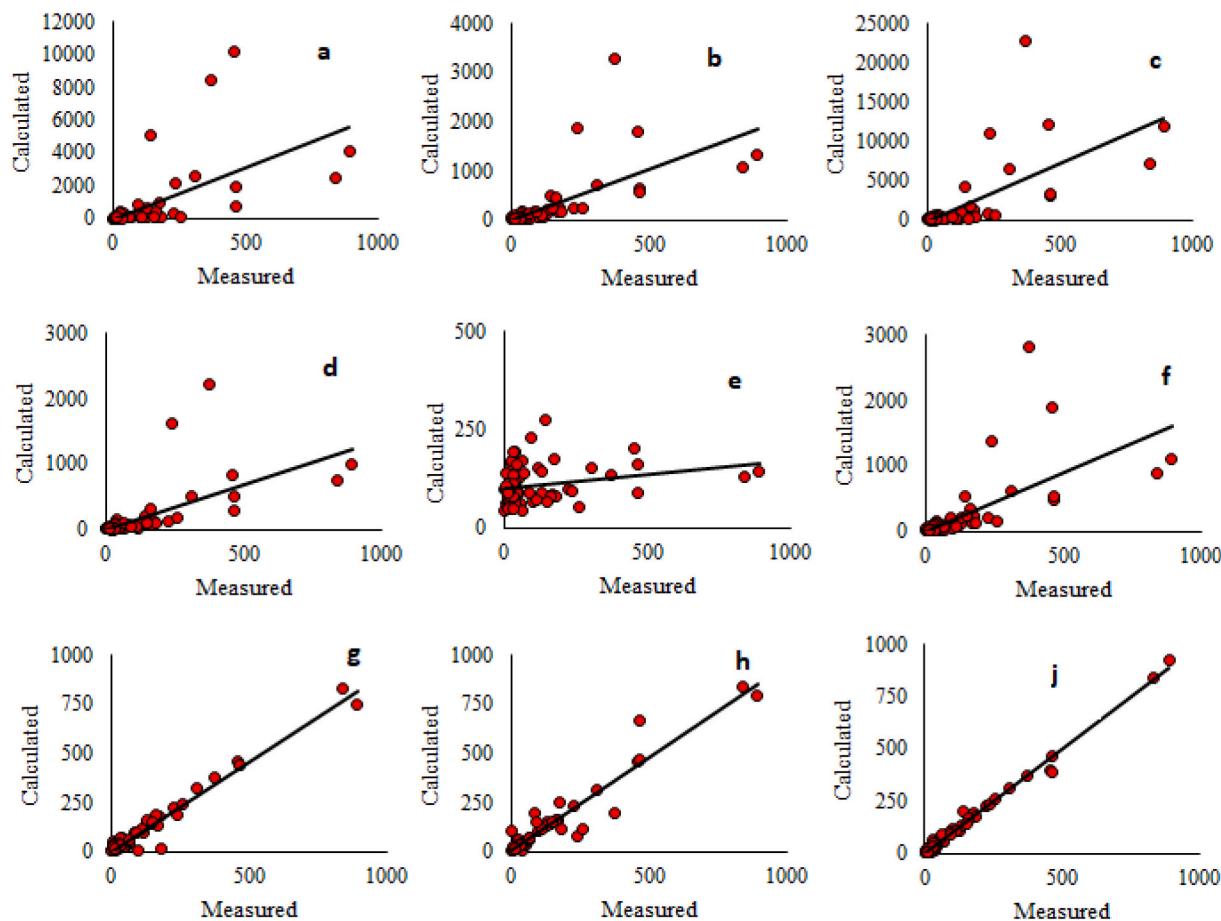


Fig. 7. Comparison of measured and predicted values of K_x , (a) Fisher et al. (1979), (b) Seo and Bae (2004), (c) Tavakkollizadeh and Kashefipur (2007), (d) Sahay and Dutta (2009), (e) Zeng and Huai (2014), (f) Sahay and Dutta (2009), (g) LS-SVM, (h) ANFIS, (i) ANFIS-HHO.

of B , H , and U . Therefore, in some cases where only the values of B , H and U are available, K_x can be predicted with appropriate performance. Scenario 1, which includes all the input parameters except σ , is the best scenario to predict the value of K_x . Using this scenario, the values of RMSE, MAPE, and R^2 for the test data were obtained equal to 24.6, 0.42, and 0.96, respectively. The complete scenario also resulted in an acceptable performance. However, in this scenario, despite its good accuracy on the training data, the prediction of the test data was conducted with lower performance compared to Scenario 1.

The trend of K_x changes for the measured and predicted values of the test data shows that using the LS-SVM model, Scenario 1 could efficiently predict the K_x behavior (Fig. 4). As can be seen in Fig. 4, the measured values are well predicted by the LS-SVM model. Furthermore, the measured and predicted values are very close to each other based on the line fitted to the data.

3.2. Results of the ANFIS and ANFIS-HHO models

The appropriate structure of the ANFIS model and the optimal values for its parameters are brought in Table 7. According to the table, the Gaussian membership function was selected as the most appropriate fuzzy membership function. Since the ANFIS uses the Sugeno-type method, the linear function was selected as the best output function of the model. The optimal structure and parameter values of the HHO algorithm can be seen in Table 8. Achieving the appropriate values for each parameter requires implementing the model with a wide range of values, which ends with determining the appropriate values of the

parameter. The population for HHO was 30, while the maximum number of iterations was obtained to be 1500.

Table 9 shows the values of the performance evaluation criteria for the ANFIS and ANFIS-HHO models. The ANFIS model predicted the training data with acceptable performance while exerted poor results for the test data. It can be said that the training algorithms of the ANFIS are trapped in local optimal minima, and as a result, irrational predictive errors are observed for the test data. For this purpose, the HHO algorithm was used to optimize ANFIS parameters to solve this problem and improve the results. For most of the scenarios, a large difference was observed between the measured and predicted data obtained by ANFIS. However, in the ANFIS-HHO model, the prediction performance of the training and test data are relatively similar. One of the essential points of using machine learning algorithms that should always be assessed is the slight difference between the prediction accuracy of training and test data, which was not observed in the ANFIS model.

The complete scenario, which included all the input parameters, resulted in the best results in both models. Using this scenario, the values of RMSE, MAPE, and R^2 of the ANFIS model for the test data were 73.4, 2.53, and 0.87, respectively. These values for the ANFIS-HHO model were 23.63, 0.38, and 0.94, respectively. The results indicate that the use of the HHO algorithm has significantly improved the predictive performance of the ANFIS model.

The trend of K_x changes for the measured and predicted values of the test data is depicted in Figs. 5 and 6 for the ANFIS and ANFIS-HHO models, respectively. The figures show that the complete scenario was capable of efficient prediction of K_x behavior. Fig. 5 shows that although

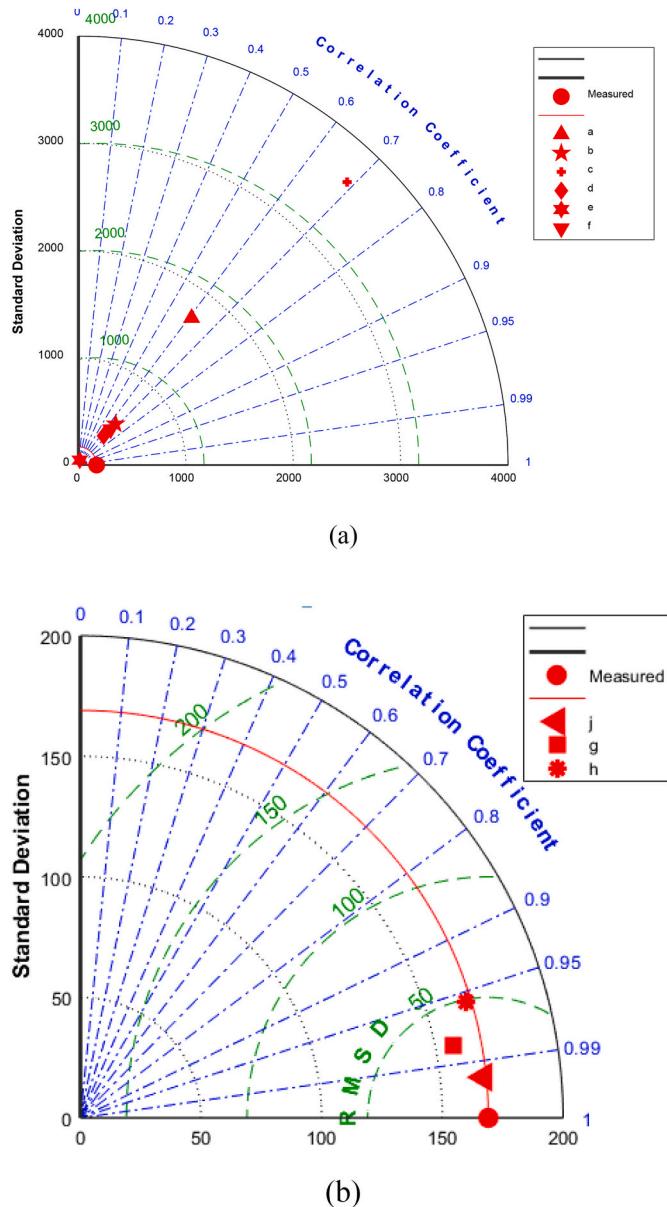


Fig. 8. Taylor's diagrams for the prediction of K_x , (a) Fisher et al. (1979), (b) Seo and Bake (2004), (c) Tavakollizadeh and Kashefipur (2007), (d) Sahay and Dutta (2009), (e) Zeng and Huai (2014), (f) Sahay and Dutta (2009), (g) LS-SVM, (h) ANFIS, (j) ANFIS-HHO.

K_x always has a positive value, negative values are predicted for K_x using the ANFIS model, which is one of the limitations of the ANFIS model. However, in Fig. 6, despite the presence of multiple peak data, the ANFIS-HHO model has detected these values correctly with relatively acceptable performance. Furthermore, the measured and predicted values are very close to each other based on the line fitted to the data.

3.3. The comparison of the machine learning models and experimental equations

Some experimental models for the calculation of K_x that were proposed by previous studies are compared with machine learning algorithms in this section. Table 10 shows the evaluation criteria of the models, including RMSE, MAPE, and R^2 . According to the table, the machine learning models resulted in a better performance than the experimental equations. Among the experimental models, the equations introduced by Sahay and Dutta (2009) and Zeng and Huai (2014) had a

relatively better performance compared to other experimental models. RMSE, MAPE, and R^2 for Sahay and Dutta (2009) equation were equal to 283, 0.85, and 0.43, respectively, while these values were equal to 163, 3.45, and 0.4, respectively for the equation of Zeng and Huai (2014). In contrast, all the machine learning models used in this study were very accurate; among them, the ANFIS-HHO model resulted in the highest performance. Using this model, the values of the RMSE, MAPE, and R^2 were equal to 17.0, 0.22, and 0.97, respectively. It can be seen that there is a significant difference between the performance of the experimental and machine learning models.

Fig. 7 shows the measured and predicted data obtained by the experimental equations and machine learning models. As can be seen, unacceptable prediction results are obtained using the experimental equations. In these equations, the difference between the measured and predicted values and compared to the regression fitted line is sometimes large. According to the figure, the most suitable match is achieved by the machine learning models, and among them, the ANFIS-HHO model. It can be concluded that the experimental equations are not capable of predicting the K_x , and the relationships between the input variables and the output are more complex than a simple linear model. Therefore, machine learning models should be used to predict the K_x values of the rivers.

One of the most suitable methods to compare the performance of several predictive models is to use Taylor's diagram. Taylor's diagrams of the experimental and machine learning models are depicted in Fig. 8 (a) and (b), respectively. The x and y axes indicate the standard deviation of the data. The quarter-circle arc shows the value of the correlation coefficient of arbitrary data and the observation data, which varies from 0 to 1. The measured data lies on the x-axis, and a predicted data close to the x-axis indicates a strong correlation with the measured data. The green arcs indicate the root mean square deviation (RMSD). According to Fig. 8(a), the results of all experimental equations had a correlation coefficient between 0.6 and 0.7, while according to Fig. 8(b), the correlation coefficient is more than 0.95 for the machine learning models.

The highest correlation coefficient was obtained for the ANFIS-HHO model, which was equal to 0.99. Among the experimental models, the highest standard deviation of data was observed for Tavakollizadeh and Kashefipur (2007) equation, which was ca. 3500, while for other experimental equations, it was less than 1800. This coefficient was much lower in machine learning models (ca. 170). The lowest value of this coefficient was obtained for the ANFIS-HHO model. Furthermore, the RMSD values for the experimental equations were very large, in the range of 500–3500, indicating a significant difference between the data predicted by the machine learning models and the experimental equations. In total, the results show that the ANFIS-HHO method had better efficiency compared to the other two methods.

Furthermore, the discrepancy ratio (DR) (Kashefipur and Falconer, 2002) was used in this study to evaluate the performance of the models (Eq. (23))

$$DR = \log\left(\frac{K_{xp}}{K_{xm}}\right) \quad (23)$$

where K_{xm} and K_{xp} indicate the measured and predicted values for the K_x of the river, respectively. If the value of the ratio is zero ($K_{xp} = K_{xm}$; $DR = 0$), there is no difference between the measured and predicted values. If the ratio is positive ($K_{xp} > K_{xm}$; $DR > 0$), the model has overestimated the values, and if the ratio is negative ($K_{xp} < K_{xm}$; $DR < 0$), the values are underestimated by the model. According to Fig. 9, the DR of the LS-SVM model was in a range between -0.3 to 0.3. In other words, 45% of the data was slightly overestimated, while 45% of the remaining was underestimated. The ANFIS and ANFIS-HHO models had similar behavior to the LS-SVM model. The DR values higher than zero for most of the experimental equations show that these models have been accompanied by an overestimation. For example, 50% of the data predicted by the equation of Zeng and Huai (2014) had a DR value between

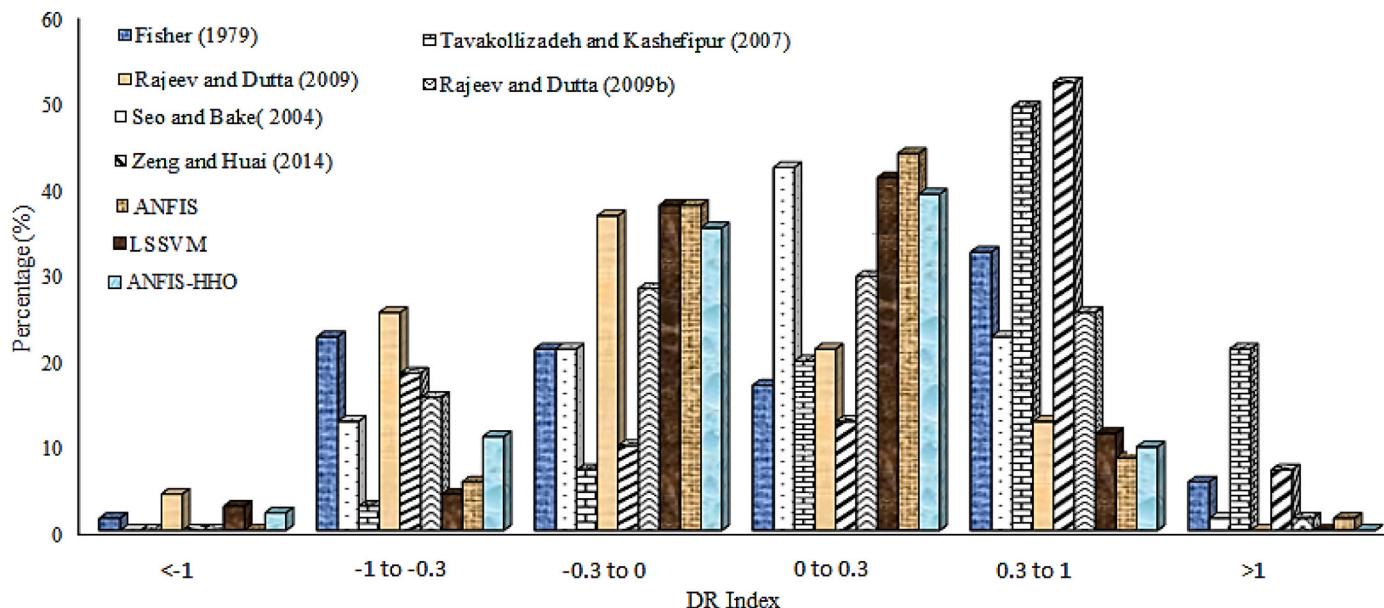


Fig. 9. Comparison of discrepancy ratio.

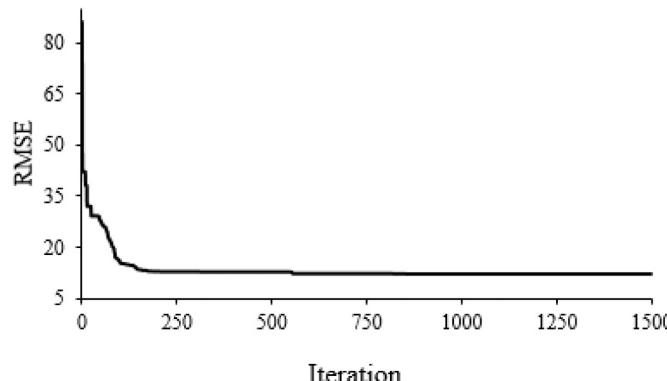


Fig. 10. Converging trend of the models to the minimum RMSE during the training.

0.3 and 1, which indicates that in most cases, it resulted in an overestimation. No *DR* values higher than 1 or lower than -1 was observed using this equation. Finally, according to Fig. 9, it can be said that all three models of LS-SVM, ANFIS, and ANFIS-HHO have resulted in predicted data in a balanced range.

The converging trend of a model to the minimum RMSE is essential for model training. According to Fig. 10, this trend was different for the studied models. The HHO algorithms had an initial RMSE of ca. 85 m and then converged rapidly to the optimum point after a small number of iterations. Finally, after 600 to 1500 iterations, all the algorithms have reached a particular error, and maximum convergence has been obtained. It is observed that after convergence, the RMSE of the HHO algorithm was 13.8, which was higher than that of the ANFIS-HHO model after convergence.

The measured and predicted values of the experimental equations and machine learning models used in the present study are brought in Fig. 11(a) and (b), respectively. According to Fig. 11(a), in the steps where the river's K_x values have a short peak, the experimental equations resulted in a significant overestimation. The equation of Tavakollizadeh and Kashefipur (2007) resulted in an irrational prediction in step 38, which has decreased its performance significantly. It can be said

that the experimental equations are not able to correctly calculate the K_x values of the outlier samples since they usually face remarkable errors for these samples. However, as shown in Fig. 11(b), the machine learning models have been able to correctly identify the values of K_x , while the ANFIS-HHO model has had the most compatibility among them. Therefore, it can be claimed that the ANFIS-HHO model is the best model for the prediction of K_x values.

3.4. Discussion

The results showed acceptable performance of LS-SVM and ANFIS-HHO models. According to the results obtained from the error evaluation criteria (RMSE, MAPE, R^2 , and Taylor's diagram), the ANFIS-HHO model was proposed. Although both models had similar results, the number of optimized parameters in the LS-SVM model is less than that of ANFIS-HHO, and therefore, the LS-SVM model is easier to use than the ANFIS-HHO hybrid model. DR values also showed the similar performance of both prediction models because both models overestimated or underestimated the output with 0.40 to 0.45% errors, which was very small compared to the ANFIS model and experimental equations. In general, it can be said that evolutionary algorithms had a high ability to

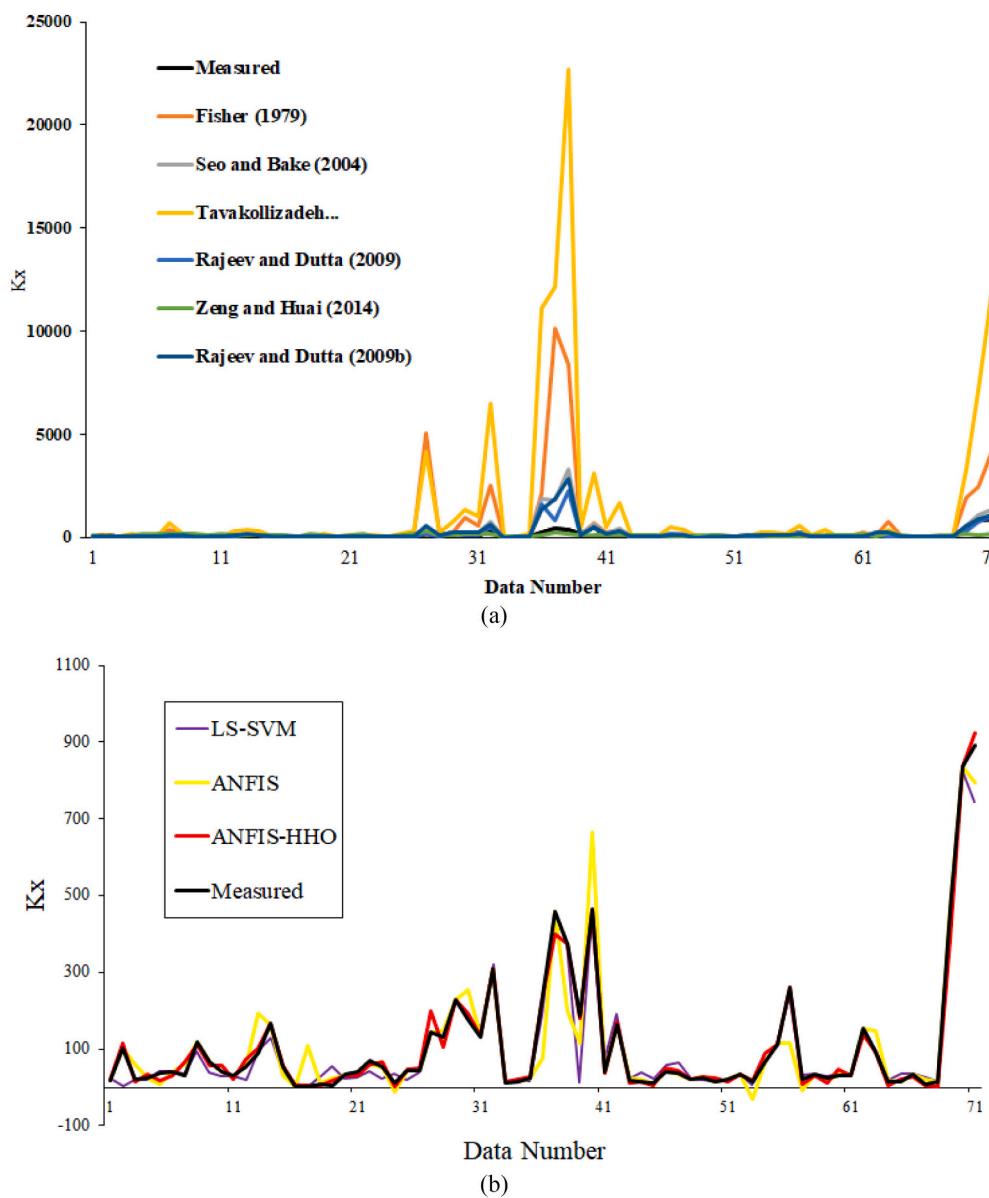


Fig. 11. Measured and predicted values for all the data used in this study, (a) measured and predicted data from experimental equations, (b) measured and predicted data from the machine learning models.

improve machine learning models for better prediction results. The results of this study were in along with Seifi and Riahi-Madvar (2019), which used algorithms for optimizing ANN, ANFIS, and GMDH to provide better results than classical models.

Another approach considered in this study and was much less considered in previous studies was the number of input variables to predict K_x and their combination to achieve the most appropriate accuracy. For this purpose, several scenarios were developed in this study, which were dimensional and non-dimensional combinations of variables. The results showed that all input variables were required to have proper accuracy in the ANFIS model and its hybrid model, but when using LS-SM, the presence of all input parameters except σ was sufficient to achieve the appropriate accuracy. Furthermore, when the non-dimensional parameters were not considered as input parameters, the accuracy of the results was somewhat reduced. Therefore, combining both dimensional and non-dimensional input variables is recommended to have proper prediction results. In predicting K_x values at peaks, the

hybrid model was able to estimate the values correctly in almost all data, while in the other two models, especially ANFIS, these values were associated with significant errors (Fig. 6). Nevertheless, in this study, for the first time, a combination of various dimensional and non-dimensional input variables was considered to evaluate their effects.

4. Conclusion

Before the emergence of machine learning models, experimental equations were used to predict the K_x of the river. However, recently, various researchers have been able to increase the accuracy of K_x predictions by using machine learning models. In this study, several advanced machine learning models were developed to predict the K_x of the river. Three models, i.e., LS-SVM, ANFIS, and ANFIS-HHO, were used to predict the value of K_x . Furthermore, the results of some of the well-known experimental equations proposed by the previous studies were compared with the results of machine learning models. The river

data based on Deng et al. (2002) were used to evaluate the performance of models and experimental equations. The results of this study can be summarized as follows:

1. Machine learning models require more input variables for efficient prediction of K_x .
2. Although ANFIS exerted an acceptable performance in the prediction of the training data, its results were poor in predicting the test data, which indicates this model is trapped in the local optimal minima during the training.
3. The performance of the machine learning models was much better than that of the experimental equations, which proves the necessity of using machine learning models in this study.
4. Performance evaluation criteria and Taylor's diagram showed that the best model for the prediction of the K_x values is ANFIS-HHO. Furthermore, the ANFIS-HHO hybrid model performed better than the ANFIS model, which shows that using the evolutionary optimization algorithms can improve the ANFIS model's performance in predictions.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Adarsh, S., 2010. Prediction of longitudinal dispersion coefficient in natural channels using soft computing techniques. *Sci. Iran.* 17 (5), 363–371.
- Afan, H.A., El-Shafie, A., Yaseen, Z.M., Hameed, M.M., Mohhtar, W.H.M.W., Hussain, A., 2015. ANN based sediment prediction model utilizing different input scenarios. *Water Resour. Manag.* 29 (4), 1231–1245.
- Alizadeh, M.J., Ahmadyar, D., Afghantlooei, A., 2017a. Improvement on the existing equations for predicting longitudinal dispersion coefficient. *Water Resour. Manag.* 31 (6), 1777–1794.
- Alizadeh, M.J., Shabani, A., Kavianpour, M.R., 2017b. Predicting longitudinal dispersion coefficient using ANN with metaheuristic training algorithms. *Int. J. Environ. Sci. Technol.* 14 (11), 2399–2410.
- Alizadeh, M.J., Shahheydari, H., Kavianpour, M.R., Shamloo, H., Barati, R., 2017c. Prediction of longitudinal dispersion coefficient in natural rivers using a cluster-based Bayesian network. *Environ. Earth Sci.* 76 (2), 86.
- Asefpour Vakilian, K., Massah, J., 2016. An apple grading system according to European fruit quality standards using Gabor filter and artificial neural networks. *Sci. Stud. Res. Chem. Eng. Biotechnol. Food Ind.* 17 (1), 75–85.
- Asefpour Vakilian, K., Massah, J., 2018. A portable nitrate biosensing device using electrochemistry and spectroscopy. *IEEE Sensors J.* 18 (8), 3080–3089.
- Azad, A., Karami, H., Farzin, S., Saeedian, A., Kashi, H., Sayyahi, F., 2018. Prediction of water quality para meters using ANFIS optimized by intelligence algorithms (case study: Gorganrood River). *KSCE J. Civ. Eng.* 22 (7), 2206–2213.
- Azad, A., Manochehri, M., Kashi, H., Farzin, S., Karami, H., Nourani, V., Shiri, J., 2019. Comparative evaluation of intelligent algorithms to improve adaptive neuro-fuzzy inference system performance in precipitation modelling. *J. Hydrol.* 571, 214–224.
- Azamathulla, H.M., Ahmad, Z., 2012. Gene-expression programming for transverse mixing coefficient. *J. Hydrol.* 434, 142–148.
- Azamathulla, H.M., Ghani, A.A., 2011. Genetic programming for predicting longitudinal dispersion coefficients in streams. *Water Resour. Manag.* 25 (6), 1537–1544.
- Azamathulla, H.M., Wu, F.C., 2011. Support vector machine approach for longitudinal dispersion coefficients in natural streams. *Appl. Soft Comput.* 11 (2), 2902–2905.
- Azamathulla, H.M., Chang, C.K., Ghani, A.A., Ariffin, J., Zakaria, N.A., Hasan, Z.A., 2009. An ANFIS-based approach for predicting the bed load for moderately sized rivers. *J. Hydro Environ. Res.* 3 (1), 35–44.
- Azamathulla, H.M., Ghani, A.A., Fei, S.Y., 2012. ANFIS-based approach for predicting sediment transport in clean sewer. *Appl. Soft Comput.* 12 (3), 1227–1230.
- Balf, M.R., Noori, R., Berndtsson, R., Ghaemi, A., Ghiasi, B., 2018. Evolutionary polynomial regression approach to predict longitudinal dispersion coefficient in rivers. *J. Water Supply Res. Technol. AQUA* 67 (5), 447–457.
- Chaudhary, F., Cogdill, R., Roussel, S., Roger, J.M., Bellon-Maurel, V., 2004. Application of LS-SVM to nonlinear phenomena in NIR spectroscopy: development of a robust and portable sensor for acidity prediction in grapes. *Chemom. Intell. Lab. Syst.* 71 (2), 141–150.
- Chang, H.H., 1988. Fluvial processes in river engineering: Malabar. Fla. Krieger.
- Deng, Z.Q., Singh, V.P., Bengtsson, L., 2001. Longitudinal dispersion coefficient in single channel streams. *J. Hydraul. Eng.* 128 (10), 901–916.
- Deng, Z.Q., Bengtsson, L., Singh, V.P., Adrian, D.D., 2002. Longitudinal dispersion coefficient in single-channel streams. *Journal of Hydraulic Engineering* 128 (10), 901–916.
- Ebtehaj, I., Bonakdari, H., Es-haghi, M.S., 2019. Design of a hybrid ANFIS-PSO model to estimate sediment transport in open channels. *Iran. J. Sci. Technol. Trans. Civil Eng.* 43 (4), 851–857.
- Elder, J.W., 1959. The dispersion of a marked fluid in turbulent shear flow. *J. Fluid Mech.* 5 (4), 544–560.
- Etemad-Shahidi, A., Taghipour, M., 2012. Predicting longitudinal dispersion coefficient in natural streams using M5' model tree. *Journal of hydraulic engineering*. 138 (6), 542–554.
- FaghforMaghrebi, M., Givehchi, M., 2007. Using non-dimensional velocity curves for estimation of longitudinal dispersion coefficient. In: Proceedings of the Seventh International Symposium River Engineering, 16–18 October, Ahwaz, Iran, pp. 87–96.
- Fisher, B.H., 1967. The mechanics of dispersion in natural streams. *J. Hydraul. Div. Am. Soc. Civil Eng.* 93 (6), 187–216.
- Fisher, B.H., List, E.J., Koh, R.C.Y., Imberger, J., Brooks, N.H., 1979. Mixing in Inland and Costal Waters. Academic Press Inc., San Diego, pp. 104–138.
- French, R.H., 1986. Open-Channel Hydraulics. McGraw-Hill Book Company, New York.
- Heidari, A.A., Mirjalili, S., Faris, H., Aljarah, I., Mafarja, M., Chen, H., 2019. Harris hawk optimization: algorithm and applications. *Futur. Gener. Comput. Syst.* 97, 849–872.
- Iwasa, Y., Aya, S., 1991. Predicting longitudinal dispersion coefficient in open-channel flows. In: International Symposium on Environmental Hydrology, Hong Kong, pp. 505–510.
- Jang, J.S., 1993. ANFIS: adaptive-network-based fuzzy inference system. *IEEE Trans. Syst. Man Cybern.* 23 (3), 665–685.
- Kashefipour, S.M., Falconer, A., 2002. Longitudinal dispersion coefficients in natural channels. In: Proceedings of the Fifth International Hydro Informatics Conference, 1–5 July, Cardiff University, pp. 95–102.
- Kong, Y., Shi, Y., Yuan, J., 2008. Prediction method of time series data stream based on wavelet transform and least squares support vector machine. In: 2008 Fourth International Conference on Natural Computation, Vol. 2. IEEE, pp. 120–124.
- Koussis, A.D., Rodriguez-Mirasol, J., 1998. Hydraulic estimation of dispersion coefficient for streams. *J. Hydraul. Eng. ASCE* 124, 317–320.
- Kumar, D., Pandey, A., Sharma, N., Flügel, W.A., 2016. Daily suspended sediment simulation using machine learning approach. *Catena* 138, 77–90.
- Leong, W.C., Bahadori, A., Zhang, J., Ahmad, Z., 2019. Prediction of water quality index (WQI) using support vector machine (SVM) and least square-support vector machine (LS-SVM). *Int. J. River Basin Manag.* 1–8.
- Li, Z.H., Huang, J., Li, J., 1998. Preliminary study on longitudinal dispersion coefficient for the gorge reservoir. In: Proceedings of the Seventh International Symposium Environmental Hydraulics, 16–18 December, Hong Kong.
- Liu, H., 1977. Predicting dispersion coefficient of streams. *J. Environ. Eng. Div.* 103 (1), 59–69.
- McQuivey, R.S., Keefer, T.N., 1974. Simple method for predicting dispersion in streams. *J. Environ. Eng. Div. Am. Soc. Civil Eng.* 100 (4), 997–1011.
- Mellit, A., Pavan, A.M., Benghanem, M., 2013. Least squares support vector machine for short-term prediction of meteorological time series. *Theor. Appl. Climatol.* 111 (1–2), 297–307.
- Memarzadeh, R., Zadeh, H.G., Dehghani, M., Riahi-Madvar, H., Seifi, A., Mortazavi, S.M., 2020. A novel equation for longitudinal dispersion coefficient prediction based on the hybrid of SSMD and whale optimization algorithm. *Sci. Total Environ.* 716, 137007.
- Milan, S.G., Roozbahani, A., Banihabib, M.E., 2018. Fuzzy optimization model and fuzzy inference system for conjunctive use of surface and groundwater resources. *J. Hydrol.* 566, 421–434.
- Mount, N., Stott, T., 2008. A discrete Bayesian network to investigate suspended sediment concentrations in an Alpine proglacial zone. *Hydrol. Process. Int.* 22 (18), 3772–3784.
- Najafzadeh, M., Ghaemi, A., 2019. Prediction of the five-day biochemical oxygen demand and chemical oxygen demand in natural streams using machine learning methods. *Environ. Monit. Assess.* 191 (6), 1–21.
- Najafzadeh, M., Oliveto, G., 2020. Riprap incipient motion for overtopping flows with machine learning models. *J. Hydroinf.* 22 (4), 749–767.
- Najafzadeh, M., Tafarjomrutz, A., 2016. Evaluation of neuro-fuzzy GMDH-based particle swarm optimization to predict longitudinal dispersion coefficient in rivers. *Environ. Earth Sci.* 75 (2), 157.
- Najafzadeh, M., Etemad-Shahidi, A., Lim, S.Y., 2016. Scour prediction in long contractions using ANFIS and SVM. *Ocean Eng.* 111, 128–135.
- Noori, R., Karbassi, A., Farokhnia, A., Dehghani, M., 2009. Predicting the longitudinal dispersion coefficient using support vector machine and adaptive neuro-fuzzy inference system techniques. *Environ. Eng. Sci.* 26 (10), 1503–1510.
- Noori, R., Deng, Z., Kiaghadi, A., Kachoo sangi, F.T., 2016. How reliable are ANN, ANFIS, and SVM techniques for predicting longitudinal dispersion coefficient in natural rivers? *J. Hydraul. Eng.* 142 (1), 04015039.
- Ozgan, E., Kap, T., Beycioglu, A., Emiroglu, M., 2009. The prediction of marshall stability of asphalt concrete by using adaptive neuro fuzzy inference system. In: International Advanced Technologies Symposium.
- Parsaie, A., Haghabi, A.H., 2015. Predicting the longitudinal dispersion coefficient by radial basis function neural network. *Model. Earth Syst. Environ.* 1 (4), 34.
- Pourabadei, M., Kashefipour, S.M., 2007. Investigation of flow parameters on dispersion coefficient of pollutants in canal. In: Proceedings of the Sixth International Symposium River Engineering, 16–18 October, Ahwaz, Iran, pp. 29–38.
- Riahi-Madvar, H., Ayyoubzadeh, S.A., Khadangi, E., Ebardzadeh, M.M., 2009. An expert system for predicting longitudinal dispersion coefficient in natural streams by using ANFIS. *Expert Syst. Appl.* 36 (4), 8589–8596.

- Riahi-Madvar, H., Dehghani, M., Seifi, A., Singh, V.P., 2019. Pareto optimal multigene genetic programming for prediction of longitudinal dispersion coefficient. *Water Resour. Manag.* 33 (3), 905–921.
- Saberi-Movahed, F., Najafzadeh, M., Mehrpooya, A., 2020. Receiving more accurate predictions for longitudinal dispersion coefficients in water pipelines: training group method of data handling using extreme learning machine conceptions. *Water Resour. Manag.* 34 (2), 529–561.
- Sahay, R.R., 2013. Predicting longitudinal dispersion coefficients in sinuous rivers by genetic algorithm. *J. Hydrol. Hydromech.* 61 (3), 214–221.
- Sahay, R.R., Dutta, S., 2009. Prediction of longitudinal dispersion coefficients in natural rivers using genetic algorithm. *Hydrol. Res.* 40 (6), 544–552.
- Sattar, A.M., Gharabaghi, B., 2015. Gene expression models for prediction of longitudinal dispersion coefficient in streams. *J. Hydrol.* 524, 587–596.
- Sedighnezhad, H., Salehi, H., Mohein, D., 2007. Comparison of different transport and dispersion of sediments in mard intake by FASTER model. In: Proceedings of the Seventh International Symposium River Engineering, 16–18 October, Ahwaz, Iran, pp. 45–54.
- Seifi, A., Riahi-Madvar, H., 2019. Improving one-dimensional pollution dispersion modeling in rivers using ANFIS and ANN-based GA optimized models. *Environ. Sci. Pollut. Res.* 26 (1), 867–885.
- Seo, I.W., Bake, K.O., 2004. Estimation of the longitudinal dispersion coefficient using the velocity profile in natural streams. *J. Hydraul. Eng.* 130 (3), 227–236.
- Seo, I.W., Cheong, T.S., 1998. Predicting longitudinal dispersion coefficient in natural stream. *J. Hydraul. Eng.* 124 (1), 25–32.
- Suykens, J.A., Vandewalle, J., 1999. Least squares support vector machine classifiers. *Neural. Process. Lett.* 9 (3), 293–300.
- Tavakollizadeh, A., Kashefipur, S.M., 2007. Effects of dispersion coefficient on quality modeling of surface waters. In: Proceedings of the Sixth International Symposium River Engineering, 16–18 October, Ahwaz, Iran, pp. 67–78.
- Tayfur, G., 2009. Optimized model predicts longitudinal dispersion coefficient in natural channels. *Hydrol. Res.* 40 (1), 60–78.
- Tayfur, G., Singh, V.P., 2005. Predicting longitudinal dispersion coefficient in natural streams by artificial neural network. *J. Hydraul. Eng.* 131 (11), 991–1000.
- Taylor, G.I., 1954. Dispersion of matter in turbulent flow through a pipe. *Proc. R. Soc. Lond. A* 223, 446–468.
- Toprak, Z.F., Cigizoglu, H.K., 2008. Predicting longitudinal dispersion coefficient in natural streams by artificial intelligence methods. *Hydrol. Process.* 22, 4106–4129.
- Toprak, Z.F., Savci, M.E., 2007. Longitudinal dispersion coefficient modeling in natural channels using fuzzy logic. *Clean* 35 (6), 626–637.
- Toprak, Z.F., Hamidi, N., Kisi, O., Gerger, R., 2014. Modeling dimensionless longitudinal dispersion coefficient in natural streams using artificial intelligence methods. *KSCE J. Civ. Eng.* 18 (2), 718–730.
- Wang, Y.P., Huai, W.X., Wang, W.J., 2017. Physically sound formula for longitudinal dispersion coefficients of natural rivers. *J. Hydrol.* 544, 511–523.
- Yang, H., Hasanipanah, M., Tahir, M.M., Bui, D.T., 2019. Intelligent prediction of blasting-induced ground vibration using ANFIS optimized by GA and PSO. *Nat. Resour. Res.* 1–12.
- Zeng, Y., Huai, W., 2014. Estimation of longitudinal dispersion coefficient in rivers. *J. Hydro Environ. Res.* 8 (1), 2–8.
- Zheng, S., Shi, W.Z., Liu, J., Tian, J., 2008. Remote sensing image fusion using multiscale mapped LS-SVM. *IEEE Trans. Geosci. Remote Sens.* 46 (5), 1313–1322.
- Zhou, S., Chu, X., Cao, S., Liu, X., Zhou, Y., 2020. Prediction of the ground temperature with ANN, LS-SVM and fuzzy LS-SVM for GSHP application. *Geothermics* 84, 101757.