



Application of data-driven models to predict the dimensions of flow separation zone

Amin Gharehbaghi¹ · Redvan Ghasemlounia² · Sarmad Dashti Latif³ · Amir Hamzeh Haghiabi⁴ · Abbas Parsaie⁵

Received: 3 November 2022 / Accepted: 11 April 2023 / Published online: 22 April 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

In this research, the effect of a submerged multiple-vane system on the dimensions of flow separation zone (DFSZ) is assessed via 192 measured datasets. The vanes' shape comprised two segments, curved and flat plates which are located in the connection of main channel to the lateral intake channel with an angle of 55° . In this direction, a butterfly's array for the vanes' arrangement along with different main controlling factors such as distances of vanes along the flow (δ_l), degree of curvature (β), and angles of attack to the local primary flow direction (θ) is utilized. Through capturing photos and utilizing AutoCAD and SURFER software, maximum relative length and width are calculated. Based on the experimental measurements, maximum percentage reduction of DFSZ, in comparison with the controlled test (without submerged vanes), is obtained with $\theta=30^\circ$, $\beta=34^\circ$, and $\delta_l=10$ cm with value of 78 and 76%, respectively. Moreover, several data-driven models, namely, gene expression programming (GEP), support vector regression (SVR), and a robust hybrid SVR with an ant colony optimization algorithm (ACO) (i.e., hybrid SVR-ACO model), are developed in order to predict DFSZ via the operative dimensionless variables realized by Spearman's rho and Pearson's coefficient processes. In accordance with the statistical metrics, model grading process, scatter plot, and the hybrid SVR(RBF)-ACO model are preferred as the best and most precise model to predict maximum relative length and width with a total grade (TG) of 6.75 and 5.8, respectively. The generated algebraic formula for DFSZ under the optimal scenario of GEP is equated with the corresponding measured ones and the results are within 0–10%.

Keywords Lateral intake · Submerged vanes · Dimensions of flow separation zone · Butterflies array · Data-driven models

Introduction

A lateral intake or diversion channel has been usually employed in order to divert waterway from rivers and main channels into the secondary channels, water supply

networks, and power stations. Some of the distinguishing features of a flow separation zone in a lateral intake are demonstrated in Fig. 1.

According to Fig. 1, three zones are observed: (1) a zone of separation in the vicinity of entrance (separation zone), (2) a contracted flow region in the channel (contracted flow

Responsible Editor: Marcus Schulz

✉ Sarmad Dashti Latif
Sarmad.latif@komar.edu.iq

Amin Gharehbaghi
amin.gharehbaghi@hku.edu.tr;
gharehbaghi.amin@gmail.com

Redvan Ghasemlounia
redvan.ghasemlounia@gedik.edu.tr

Amir Hamzeh Haghiabi
haghiabi@lu.ac.ir

Abbas Parsaie
Parsaie@scu.ac.ir

² Faculty of Engineering, Dept. of Civil Engineering, Istanbul Gedik Univ, Istanbul 34876, Turkey

³ Civil Engineering Department, College of Engineering, Komar University of Science and Technology, Kurdistan Region, Sulaimany 46001, Iraq

⁴ Department of Water Science and Engineering, Faculty of Agriculture and Natural Resources, Lorestan University, Lorestan, Iran

⁵ Department of Hydraulic Structures, Faculty of Water and Environmental Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran

¹ Faculty of Engineering, Dept. of Civil Engineering, Hasan Kalyoncu University, Şahinbey, Gaziantep 27110, Turkey

zone), and (3) a stagnation point next to the downstream corner of junction (stagnation zone) (Ramamurthy et al. 2007). The flow pattern in the lateral intake is extremely complicated since the flow separation results in some remarkable changes in the hydraulic conditions such as reduction of the actual width of current and further alluviation in the entrance of lateral intake. Thus, it is crucial to determine DFSZ.

In order to reduce DFSZ and manage misdirected discharge/sediments into the basin, it has generally been found effective to install submerged vanes in the entry of lateral intake. Submerged vanes frequently create downstream vortices that shift the flow direction and apply shear stress to the channel bed. It causes bed material to move in a transverse direction along the channel's cross-section, which changes the bed profile. The properties and array of vanes can be arranged logically to achieve the appropriate bed profiles (Sharma and Ahmad 2020).

Submerged vane has been employed for extensive purposes such as channel bank protections (Wang and Odgaard 1993), sediment elimination in water intakes (Wang et al. 1996), and expanding rivers for navigation (Odgaard and Wang 1991b). A summary of theoretical studies and innovative relationships for sediment control by submerged vanes is provided by Odgaard and Spoljaric (1988) and Odgaard and Wang (1991a). They reported that the small submerged vanes caused the remarkable alterations in the velocity distribution in rivers. Nakato et al. (1990) suggested a pioneering technique in the laboratory field to produce a scour-trench in front of a power-plant. They concluded that the performance of the utilized vane model can be considered an outstanding scheme. Barkdoll et al. (1999) conducted numerous laboratory tests to identify the fundamental geometrical and physical characteristics that submerged vanes can use to stop extreme bed sedimentation into channels used for diverting sediment from sedimentary rivers. His experiments proved that the recommended shape for submerged vanes caused only an insignificant rate of bed sediment pass into the diversion as the ratio of q_r was less than 0.2. Numerous instance field applications have been

presented by Odgaard (2009). Gupta et al. (2010) studied the results of a collar on the efficiency of a scour hole around the submerged vane and optimized its dimensions in the prevention of scour hole. They declared that by using a collar on the leading edge of a submerged vane, the scour depth was noticeably diminished. Behbahan (2011) investigated the effect of different shapes of submerged vanes as flat, angled, and curved types with various arrays on the scour and sediment control. He reported that based on experimental results, the curved and angled forms were more operative than the flat form in riverbank protection by 35 and 20%, respectively. Ouyang (2009), Ouyang and Lin (2016), and Karami et al. (2017) conducted in-depth analyses to evaluate the border effects of channel banks, water surface, and channel bed profile generated by various arrays of submerged vanes in various forms using novel techniques.

Characteristically, a hydrogeological process can be analyzed by three methods, including physical, numerical, and artificial intelligence (AI) methods. To date, many complicated numerical methods have been adopted to evaluate the flow characteristics and sediment transportation phenomenon in an open channel (Kaya and Gharehbaghi 2012; Gharehbaghi et al. 2016, 2017; Sui and Huang 2017; Ehteram et al. 2020; Monier et al. 2020). Nevertheless, because of complex flow interactions, these methods were mostly unable to precisely justify vane-induced flow field characteristics by submerged multiple-vane systems.

Due to the costly and time-consuming function of experimental and numerical approaches, nonlinear systems such as machine learning models (MLMs) and intelligent optimization models (IOMs), AI-based techniques, should be utilized to identify the extremely unbalanced behavior of DFSZ. These methods are well known as the top adaptive ones suitable for finding complex and nonlinear indefinite patterns in large dimensional data. As scientists' skill with these AI-based systems deepens, they are becoming more dependable, and now they are frequently utilized as robust approaches in different fields of water sciences to predict complex hydraulic and hydrological variables such as sugarcane growth based on climatological parameters (Taherei Ghazvinei et al. 2018), daily dew point temperature (Qasem et al. 2019), forecasting nitrate concentration as a water quality parameter (Latif et al. 2020), inflow forecasting (Latif et al. 2021a), phosphate forecasting in reservoir water system (Latif et al. 2021b), daily streamflow time-series prediction (Latif and Ahmed 2021; Tofiq et al. 2022), surface water quality status and prediction during movement control operation order under COVID-19 pandemic (Najah et al. 2021), groundwater level fluctuations (Ghasemlounia et al. 2021; Gharehbaghi et al. 2022), discharge coefficient of a new type of sharp-crested V-notch weirs (Gharehbaghi and Ghasemlounia 2022), and dissolved oxygen prediction (Ziyad Sami et al. 2022).

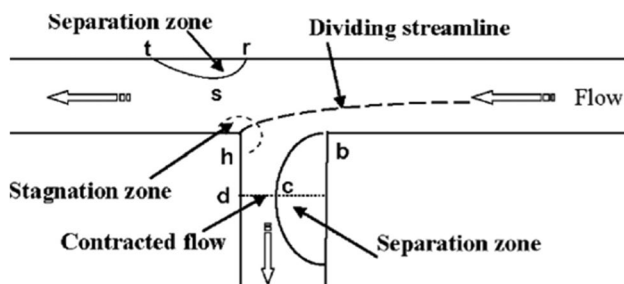


Fig. 1 Characteristics of dividing flow in open channels (Ramamurthy et al. 2007)

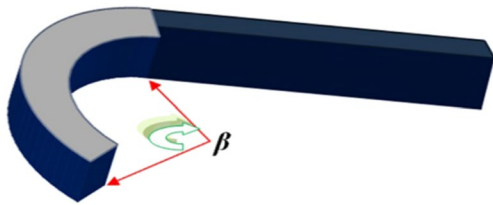


Fig. 2 The sketch view of the present submerged vane

Regarding the AI methods, GEP (gene expression program) and SVR (support vector regression) are very popular and have been also regularly utilized in various fields of engineering. GEP is contemplated as a well-recognized AI-based technique inspired by biological basics and has a satisfactory ability in estimating parameters with high nonlinearity relationships. SVR is regarded as a commendable method for classification, regression, and function estimation. In recent years, IOMs have been developed as an effective substitute for the traditional standalone data-driven models (DDMs), AI-based modeling methodologies, to improve performance and resolve issues. This study uses three alternative DDMs, including SVR, GEP, and a robust hybrid bio-inspired IOM, to both test experimentally and predict DFSZ induced by a coupled submerged vane. The hybrid model is a coupled version of SVR with an ant colony optimization (ACO) algorithm (i.e., hybrid SVR-ACO model), which is employed to progress the estimation precision of DFSZ. In the hybrid SVR-ACO model, ACO is a bio-inspired IOM that optimizes mechanical approximation parameters of the standalone SVR model in computational engineering problems. Several scenarios are defined in the current simulation by varying the hyperparameters and structure of the produced DDMs. It should be noted that finding the ideal value for a hyperparameter requires a process of trial and error in order to reach the best configuration. In this way, 192 experimental datasets are used to evaluate DFSZ while accounting for dimensionless effects.

According to the literature study, earlier studies primarily focused on the most fundamental types of submerged vanes since their theories and underlying assumptions were oversimplified. The innovative aspect of the current study is the experimental assessment of the effects of coupled submerged vanes on the DFSZ and flow pattern at the entry of the lateral intake, and the use of DDMs for prediction.

The main aims and scope of this study are to (i) investigate experimentally effects of a combined submerged multiple-vane system on DFSZ and flow attributes, (ii) determine the relationships among nitty-gritty layout and physical characteristics of the present submerged multiple-vane system with DFSZ to improve vane performances, and

(iii) develop different DDM techniques to estimate DFSZ. Additionally, an evaluation is carried out to determine the best DDMs by validating applied physical and DDMs using statistical measures.

Experimental setup

The combined submerged vanes employed in this work are comprised of two segments, curved and flat plates. The curved section has different degrees of curvature (β) including 0, 17, 34, and 68°, as described in Fig. 2.

A laboratory system for experimental work is described in Fig. 3 (A and B). In this submerged multiple-vane system, the optimal angle of lateral intake channel, i.e., 55° (Keshavarzi and Habibi 2005), and the butterfly array among vanes are utilized. Many studies have been conducted in the area of submerged vanes' size and array. Two rows of vanes would theoretically be more effective than three rows for managing sediment in any vane layout (Barkdoll et al. 1999). The most successful design guidelines and dimensions are typically based on local flow-field observations and expert judgments because there are no specified all-inclusive guidelines and standards for the ideal dimensions and design essentials for submerged vanes in the literature.

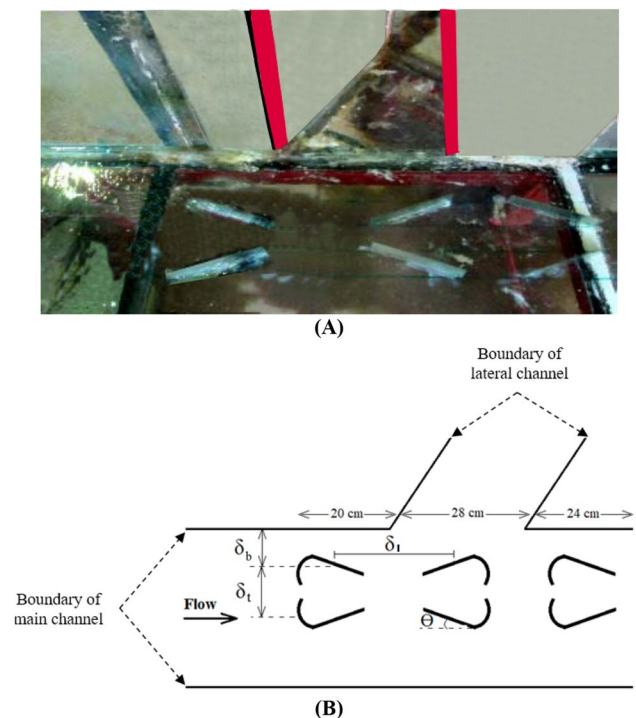


Fig. 3 Characteristics of the proposed submerged multiple-vane system in the **A** $\beta=0^\circ$ and **B** $\beta=17, 34,$ and 68°

Table 1 Specifications of present submerged multiple-vane system

Parameters	Value	Parameters	Value
θ	0,15,30, and 45°	δ_l	10, 15, and 20 (cm)
H_v	5 (cm)	δ_r	10 (cm)
L_v	18 (cm)	δ_b	10 (cm)
β	0,17, 34, and 68°		

In Fig. 3, θ is the angle of attack to the local primary flow direction, δ_l is the distances in the median of vanes along the flow (i.e., stream-wise direction); L_v and H_v are the length and height of the proposed vane, respectively; δ_r is the distance in the median of vanes across the flow (i.e., perpendicular to streamwise direction); δ_b is the distance of first row from in the median of vanes to the bank. It is necessary to mention that the dimensions of the proposed multiple-vane system are taken from the standards and recommended ranges by Odgaard and Wang (1991a, b) and Behbahan (2011), and the size of flume utilized in the hydraulic laboratory.

The hydraulic laboratory of Tabriz Azad University’s Water Engineering Department is where this study’s experimental work is done. In order to assess DFSZ in the horizontal lateral intake channel, two sets of experiments are conducted in this regard. The approved submerged multiple-vane system is employed in the second set, whereas no vane is used in the first set (control test). Table 1 provides the dimensions of the utilized system and vanes.

Based on the requirements of Froude law scaling and free steady flow conditions, all discharges are taken into account in the ranges of 20 to 35 (l/s), making the downstream flow supercritical in all tests. On the basis of flow-field analysis, potential plans to enhance vane performance are identified.

In this experiment, a 10-m long, 0.4-m wide, and 0.5-m high rectangular horizontal flume is employed as the main

channel, and a 4-m long, 0.28-m wide, and 0.5-m high rectangular horizontal flume is used as the lateral intake channel. Both channels have a longitudinal slope of 0.0001 with glassy sidewalls for better visualization of flow patterns.

In the intake of the main channel, a grid wall and a wave suppressor are employed to reduce turbulence on the water’s surface, break big eddies, and maintain flow homogeneity. To measure the depths of flow downstream of lateral intake and upstream of main channel, point gauges with Vernier scales and precision of ± 0.1 (mm) are used. Figure 4 (A and B) shows the geometric characteristics of the experimental open channel set-up and the point gauges employed in the laboratory.

A pump with a power of 200 hp discharges the water into the stilling basin in the inlet of main channel. In both channels, all discharges (Q) are gauged by using an electromagnetic flow meter in the measuring basin with a precision of $\pm 0.5\%$.

In this study, to determine the DFSZ in the entrance of lateral intake channel, numerous photos of the separation zone are captured from the top view; then, through these photos, the maximum length (L_s) and width (W_s) of separation zone are computed by AUTOCAD and SURFER software.

Dimensional analysis

In this research, experimental data is assessed to determine the potential parameters on DFSZ. In this respect, this phenomenon can be written as a function of

$$f_1(y_1, y_2, B_2, q_1, q_2, L_v, H_v, \rho, g, \mu, \delta_l, \delta_r, L_s, W_s, S_0, \alpha, \theta, \beta, n) = 0 \tag{1}$$

where y_1 and y_2 are the depths of flow in the upstream of main channel and downstream of lateral intake channel,

Fig. 4 A Layout of the experimental setup and B point gauges with Vernier scales

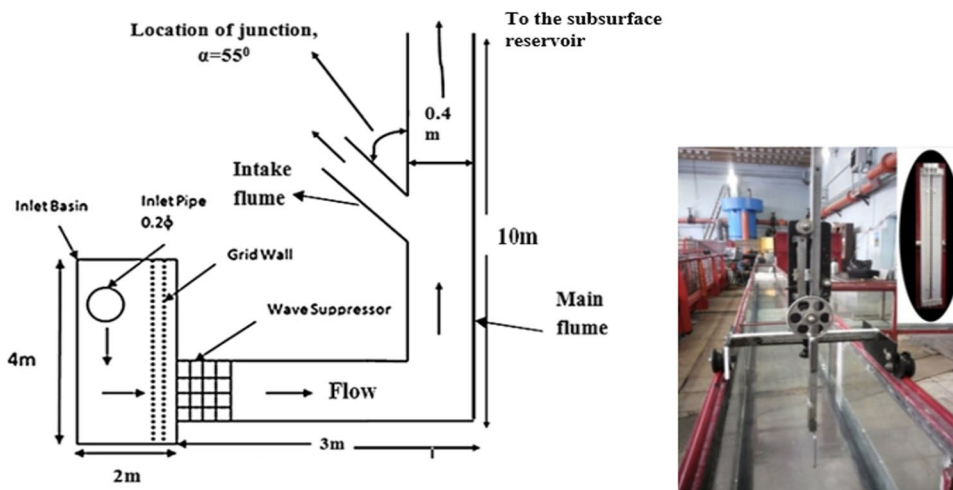


Table 2 Range of used experimental data

Variables	Range	Variables	Range
$\frac{H_V}{y_1}$	0.193–0.335	Fr_1	0.1–0.5
$\frac{L_V}{y_1}$	0.579–0.99	$\frac{\delta_l}{y_1}$	0.324–1.11
$\frac{y_2}{y_1}$	0.41–0.75	$q_r (\frac{q_2}{q_1})$	1.2–1.5
$\frac{L_s}{B_2}$	0.351–1.241	$\frac{\delta_b}{y_1}$	0.321–0.558
$\frac{W_s}{B_2}$	0.085–0.297	$\frac{\delta_b}{y_1}$	0.321–0.558

respectively, B_2 is the width of lateral intake channel, q_1 is the inflow discharge per unit width in the main channel, q_2 is the outflow discharge per unit width in the lateral intake channel, L_v and H_v are length and height of present vanes, respectively, ρ is flow density, n is the number of rows in the proposed submerged multiple-vane system, g is the gravitational acceleration, μ is the dynamic viscosity of flow, L_s and W_s are the maximum length and width of separation zone, respectively, S_o is the channel slope, α is the angle of the lateral intake, and β is the degree of curvature in the curved segment of vanes suggested.

Through Buckingham π -theorem and the property of dimensional analysis, non-dimensional parameters can be obtained as

$$f\left(Fr_1, Re_1, \frac{B_2}{y_1}, \theta, S_o, \alpha, n, \beta, \frac{\delta_b}{y_1}, \frac{\delta_l}{y_1}, \frac{\delta_t}{y_1}, \frac{q_2}{q_1}, \frac{L_s}{y_1}, \frac{W_s}{y_1}, \frac{y_2}{y_1}, \frac{L_V}{y_1}, \frac{H_V}{y_1}\right) = 0 \tag{2}$$

where Fr_1 and Re_1 are the approaching Froude and Reynolds numbers (dimensionless). By overlooking fixed-amounts variables including S_o , α , and n and supposing a high scale of turbulence (i.e., large Reynolds numbers), the effect of Re_1 can be disregarded (Borgheti et al. 1999). Lastly, for the sake of simplicity, Eq. (2) can be expressed in the following forms:

$$\frac{L_s}{B_2} = \left(Fr_1, \frac{\delta_b}{y_1}, \frac{\delta_l}{y_1}, \frac{\delta_t}{y_1}, \frac{q_2}{q_1}, \frac{y_2}{y_1}, \frac{L_V}{y_1}, \frac{H_V}{y_1}, \theta, \beta\right) \tag{3}$$

$$\frac{W_s}{B_2} = \left(Fr_1, \frac{\delta_b}{y_1}, \frac{\delta_l}{y_1}, \frac{\delta_t}{y_1}, \frac{q_2}{q_1}, \frac{y_2}{y_1}, \frac{L_V}{y_1}, \frac{H_V}{y_1}, \theta, \beta\right) \tag{4}$$

The $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$ as representatives of $DFSZ$ are the maximum relative length and width of flow separation zone. The range and statistical indexes of parameters used in Eqs. (3) and (4) are presented in Tables 2 and 3, respectively.

Pre-processing

Many characteristics have been utilized in many different sorts of research to explain unidentified events, albeit some of these parameters have negligible influence. Unwise choices might negatively impact on the effectiveness of approaches, because the performance of the exact estimation of the target depends on the opposite choice of forecasters. The combination of model inputs and adequate datasets for training and testing are two of the key factors limiting how accurately nonlinear techniques can model (Gharehbaghi and Ghasemlounia 2022). In this way, approaches for dimensionality reduction have been used to extract useful parameters and speed up the computation. In order to select appropriate input variables for Eqs. (3) and (4), this method is applied in this work using two special tests, Spearman’s rho and Pearson’s

correlation coefficient, which serve as examples of non-parametric and parametric statistical tests, respectively. With a 95% confidence level, Table 4 displays the correlation coefficient values of the variables utilized in Eqs. (3) and (4).

Based on Table 4, the highest correlation coefficient values in $\frac{L_s}{B_2}$ and

$\frac{W_s}{B_2}$ are observed with the $\frac{q_2}{q_1}, \frac{\delta_l}{y_1}, \frac{\delta_b}{y_1}, \frac{\delta_t}{y_1}, \frac{L_V}{y_1}, \frac{H_V}{y_1}, \beta$; and $\theta, \beta, \frac{q_2}{q_1}, \frac{\delta_l}{y_1}, Fr_1, \frac{L_V}{y_1}, \frac{H_V}{y_1}$, respectively; hence, they are considered input parameters, and $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$ are target. Finally, Eqs. (3) and (4) can be rewritten in the ultimate following forms:

Table 3 Statistical indexes of variables used in Eqs. (3) and (4)

Statistical indexes	Fr_1	$\frac{L_V}{y_1}$	$\frac{H_V}{y_1}$	$\frac{q_2}{q_1}$	$\frac{\delta_b}{y_1}$	$\frac{\delta_l}{y_1}$	$\frac{\delta_t}{y_1}$	$\frac{y_2}{y_1}$	$\frac{L_s}{B_2}$	$\frac{W_s}{B_2}$
Mean	0.31	0.71	0.23	1.35	0.39	0.59	0.39	0.57	0.62	0.18
Standard Deviation	0.11	0.11	0.039	0.07	0.06	0.19	0.06	0.11	0.22	0.049
Variance	0.01	0.01	1×10^{-4}	1×10^{-4}	0.004	0.037	0.004	0.012	0.048	0.002
Skewness	0.007	1.035	1.035	-0.08	1.035	0.579	1.03	0	1.24	1.161
Kurtosis	-1.13	0.096	0.096	-0.98	0.096	-0.11	0.09	-1.31	0.36	0.959

Table 4 Spearman’s rho and Pearson’s correlation coefficient values in Eqs. (3) and (4)

Type of correlation coefficient	Variable	θ	β	$\frac{H_V}{y_1}$	$\frac{L_V}{y_1}$	Fr_1	$\frac{\delta_b}{y_1}$	$\frac{\delta_l}{y_1}$	$\frac{q_2}{q_1}$	$\frac{y_2}{y_1}$	$\frac{\delta_t}{y_1}$
Spearman’s rho	$\frac{L_s}{B_2}$	-0.53	0.63	0.44	0.44	0.042	0.44	0.61	0.43	-0.38	0.44
Spearman’s rho	$\frac{W_s}{B_2}$	0.48	0.62	0.33	0.33	0.65	0.13	0.36	0.49	-0.61	0.13
Pearson	$\frac{L_s}{B_2}$	-0.7	0.69	0.73	0.73	0.023	0.71	0.64	0.56	-0.17	0.71
Pearson	$\frac{W_s}{B_2}$	0.54	0.66	0.43	0.43	0.55	0.11	0.46	0.51	-0.52	0.11

$$\frac{L_s}{B_2} = \left(\beta, \frac{q_2}{q_1}, \frac{\delta_l}{y_1}, \frac{\delta_b}{y_1}, \frac{\delta_t}{y_1}, \frac{L_V}{y_1}, \frac{H_V}{y_1} \right) \tag{5}$$

$$\frac{W_s}{B_2} = \left(\theta, \beta, \frac{q_2}{q_1}, \frac{\delta_l}{y_1}, Fr_1, \frac{L_V}{y_1}, \frac{H_V}{y_1} \right) \tag{6}$$

Sensitivity analysis

A sensitivity analysis for the forecasters in Eqs. (5) and (6) is implemented to depict the impact of each forecaster on the model target by altering each forecaster in a fixed amount and holding the other forecasters constant. The cosine amplitude technique is applied to adopt the sensitivity analysis (Momeni et al. 2014) as follows:

$$R_{ij} = \frac{\sum_{k=1}^N (I_{ik} O_{jk})}{\sqrt{\sum_{k=1}^N I_{ik}^2 \sum_{k=1}^N O_{jk}^2}} \tag{7}$$

where N is the total number of datasets and I_i and O_j are the input and output variables. R_{ij} value is in the range of 0 to 1 and specifies the strength of relationship between every forecaster and target in Eqs. (5) and (6), as are presented in Fig. 5 A and B. According to these figures, $\frac{H_V}{y_1}$ and $\frac{L_V}{y_1}$ are the most effective variables for estimation of $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$, respectively.

Estimation models

Based on Table 3, it is apparent that $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$ have an unstable behavior via the kurtosis and skewness, and their relationships with forecasters in Eqs. (5) and (6) are extremely intricate and nonlinear. Thus, robust accurate DDMs are essential for estimating $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$. Concerning, as above mentioned, in the present study, for predicting $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$ by SVR, GEP, and hybrid SVR-ACO methods, 192 experimental datasets are employed. In this direction, a set of 144 datasets are randomly utilized for calibration (75% of the data), and the other 48 datasets (25%) are randomly utilized for validation. Then, these sets are normalized to zero mean and unit variance as suggested by Lawrence et al. (1997).

Gene expression programming (GEP)

GEP was developed by Ferreira (2001) inspired by “Darwinian evolution theory” and is one of the circulating algorithm systems. In this method, an excellent population is elected, or else a fresh populace is regenerated to accomplish the excellent populace. The flowchart of a GEP algorithm is depicted in Fig. 6.

The first step in the GEP method, which can run randomly or with some information, is the development of the primary populace. The structure of an expression tree is then given, along with the chromosomes. To determine whether a resolution is appropriate, the effects are evaluated using a fitness function. If a respectable resolution value is attained, the progression

Fig. 5 Values of R_{ij} for the variables used in Eqs. (5) and (6): **A** $\frac{L_s}{B_2}$ and **B** $\frac{W_s}{B_2}$

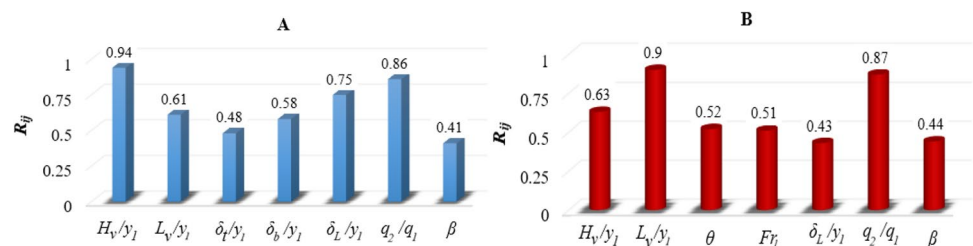
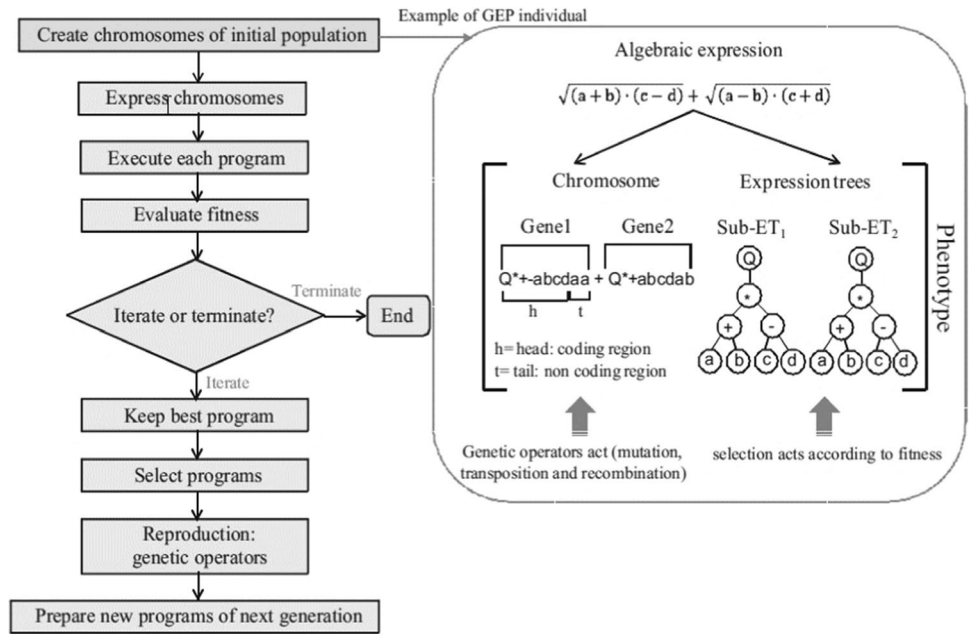


Fig. 6 Flowchart of a GEP algorithm (Ferreira 2001)



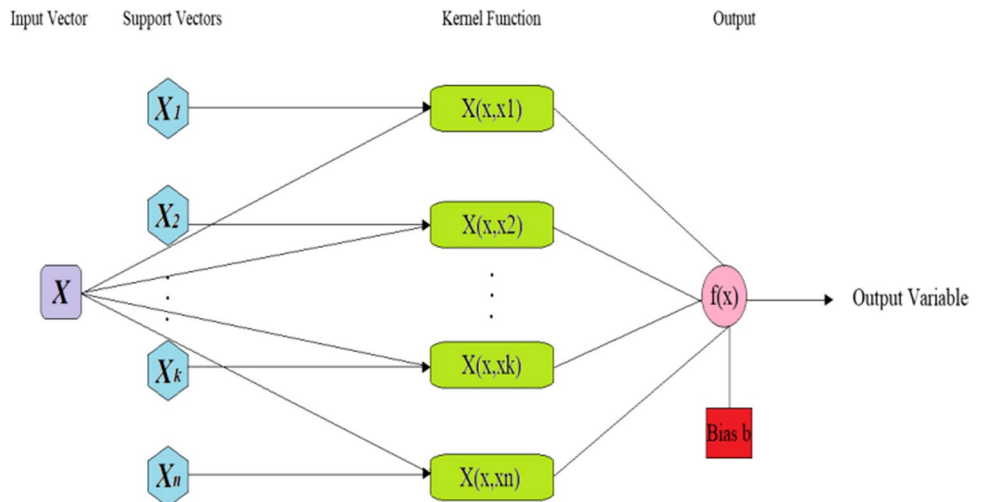
mechanism is stopped, and the superb resolution for this phase is then shown. The best option for the current group is held in reserve in the absence of stop conditions. This process is performed for a predetermined number of groups in order to achieve the desired outcome (Ferreira 2001).

GEP model development

In the current research, GeneXpro Tools 4.0 program is employed to adopt GEP algorithm for estimating the values of $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$ by following steps (Mehdzadeh et al. 2017). In the first step, *RMSE* has been opted as the suitable fitness function. The second step is to define variables and function sets to generate the chromosomes. In

this direction, forecasters and target variables ($\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$) in Eqs. (5) and (6) have been, respectively, opted. Function sets are chosen for four dominant mathematics operators including $\{+ - / \times\}$ and several arithmetical functions including x^2, x^3, e^x , etc. The third step is identifying the basic components of chromosomes, such as head size, number of genes, and chromosomes. The fourth phase uses additional functions like an expression tree connecting function to connect subsections. Lastly, the stop condition is a generation number of 10,000. Several situations are adopted in this work by adjusting important hyperparameters including the number of chromosomes and genes, head size, and kind of linking function. It

Fig. 7 Structure of an SVR algorithm



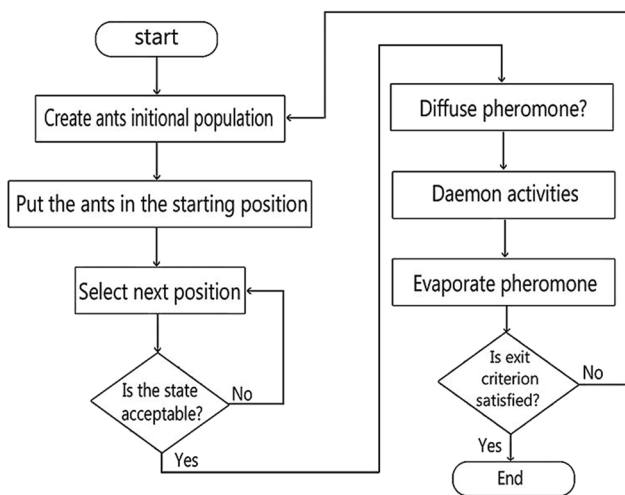


Fig. 8 Flowchart of hybrid SVR-ACO model

should be noted that these hyperparameters are chosen through a process of trial and error to achieve the optimal model configuration.

Support vector regression (SVR)

SVM was created and descended from statistical learning concepts. It functions as a subclass of supervised MLMs and employs an operational risk reduction procedure to reduce high oversimplification error. It may also be used to realize patterns or arrange goods in certain sectors, which is known as SVR. SVR is viewed as an effective method for forecasting, function estimation, and regression. Additionally, by effectively minimizing the observed training and distribution error, it can learn and estimate nonlinear relationships among the predictor-target datasets in vast dimensions (Yu et al. 2018). The construction of an SVR algorithm is shown in Fig. 7.

In keeping with Fig. 7, the *n*-dimensional observed datasets are categorized as the predictor vectors and one-dimensional target vectors are categorized as the output. SVR distinguishes a fitting function *f(x)*, by dropping the distribution error limitations to accomplish a general regression efficacy as

$$f(x) = wT \times \varphi(x) + b \tag{8}$$

where $\varphi(x)$ denotes a nonlinear function, *b* is the bias statement, and *w* is the weight vectors specified by decreasing the risk function. Therefore, it is vital to decrease the error function (Eq. 9) by applying some conditions presented in Eq. (10) (Pai and Hong 2006) as follows:

$$Minimize_{w,b} \frac{1}{2} w^T \times w + C \sum_{i=1}^l (\xi_i^* + \xi_i) \tag{9}$$

$$yi - (wT \times \varphi(x) + b) \leq \epsilon + \xi_i; (wT \times \varphi(x) + b) - yi \leq \epsilon + \xi_i^*; \xi_i^*, \xi_i \geq 0 \tag{10}$$

where *C* is a legalization argument with a positive integer amount. Bigger values of *C* imply a more complicated and precise learning machine, yet a lesser *C* value implies large error toleration and a fairly weaker one (Smola and Schölkopf 2004; Yu et al. 2006). ξ_i and ξ_i^* are the slack variables (Olatomiwa et al. 2015). As SVR network is calibrated, it approximates outputs via the next equation as follows:

$$f(x) = \sum_{i=1}^l \theta_i \times \varphi(x, x_i) + b \tag{11}$$

where θ_i exemplifies the mean Lagrangian coefficient. The calculation of $\varphi(x)$ in futures space is very complicated.

The SVR approach can make use of a variety of kernel functions, including linear, RBF, and polynomial. The type and parameters of a kernel function are often chosen based on the data, which has a substantial impact on prediction accuracy. There are no standard guidelines that specify which kernel function to use with a given set of data. Given that the choice of kernel parameters directly organizes the complexity of computations, the optimal rates of *C* and error for the sensitive area should be specified (Zaji et al. 2016; Mohammadi and Mehdizadeh 2020).

SVR model development

The SVR model’s calibration and validation accuracy are used in this work, and the MATLAB 2019b program is used. In each case, the objective is DFSZ, and the predictor sets are the dimensionless parameters from Eqs. (5) and (6).

Table 5 Maximum percentage reduction of $\frac{L_2}{B_2}$ and $\frac{W_2}{B_2}$ in the applied $\delta_r, \beta^\circ,$ and θ

β°	θ°	$\delta_{(cm)}$	$\frac{L_2}{B_2}(\%)$	$\frac{W_2}{B_2}(\%)$
0	0	10	28	21
0	15	10	34	29
0	30	10	41	37
0	45	10	36	35
17	0	10	39	32
17	15	10	47	44
17	30	10	61	49
17	45	10	56	45
34	0	10	52	41
34	15	10	76	72
34	30	10	78	76
34	45	10	75	69
68	0	10	45	35
68	15	10	58	59
68	30	10	69	65
68	45	10	63	58

Table 6 MPR_{WD} in $\beta=34^\circ$ and applied δ_i and θ

θ°	$\delta_i(\text{cm})$	$MPR_{WD}(\%)$
30	15	43
45	10	39
15	20	36
0	20	18

Additionally, the strong nonlinearity of the relationships in Eqs. (5) and (6) is investigated using RBF and polynomial kernel functions. Then, their key parameters i.e., d, ϵ, C in SVR-poly and γ, ϵ, C in SVR-RBF models are determined by developing several scenarios and using a trial-and-error process to acquire the ideal structure. In this sense, different ranges and step size are randomly taken into account for kernel amounts. $RMSE$ and R^2 values of each scenario are regarded as the evaluation metrics; consequently, a kernel with the maximum R^2 and minimum $RMSE$ values is contemplated as the ideal one.

Hybrid SVR-ACO model

ACO is a successful IOM inspired by the actual and group actions of ants in their realm (Dorigo et al. 1999), which is suitable for dealing with issues in the direct paths. The next equations are used to adopt the math behind this algorithm:

$$c_i^t = Ant_j^t + c^t \tag{12}$$

$$d_i^t = Ant_j^t + d^t \tag{13}$$

where c_i^t is the minimum of all variables for i^{th} ant, c^t is the minimum amongst variables in t^{th} iteration, d_i^t is the maximum of all variables for i^{th} ant, and d^t is the vector including

the maximum of all variables in the t^{th} repetition. Ant_j^t determines the location of the picked j^{th} ant in the t^{th} repetition. The next equations are also applied to transfer the ants at random as

$$c^t = \frac{c^t}{I} \tag{14}$$

$$d^t = \frac{d^t}{I} \tag{15}$$

where I is a ratio (Dorigo et al. 1996, 1999; Gharehbaghi and Ghasemlounia 2022).

Hybrid SVR-ACO development

The hybrid SVR-ACO model is developed to boost the capability of conventional standalone SVR model for estimating $\frac{L_\epsilon}{B_2}$ and $\frac{W_\epsilon}{B_2}$. Since RBF appraises the scattered data with high accuracy, it is employed as the kernel function. In this method, ACO algorithm is coupled with SVR to optimize RBF functional parameters. The differences in kernel amount of RBF function affect meaningfully the precision of results.

The layers of the conventional SVR-RBF approach’s membership functions, weights, and biases are trained using ACO. The optimization procedure in the new d-dimensional space needs to be constructed in order to apply ACO to update weights. This methodology modifies each ACO circumstance resident and, as a consequence, regenerates the best possible ACO. In this regard, the SVR-RBF weights are updated based on data obtained from the superior ACO. The original weights are then replicated using the calibration set, and ACO optimized the membership functions of the original weights. When the variance between the SVR output and the original output is less

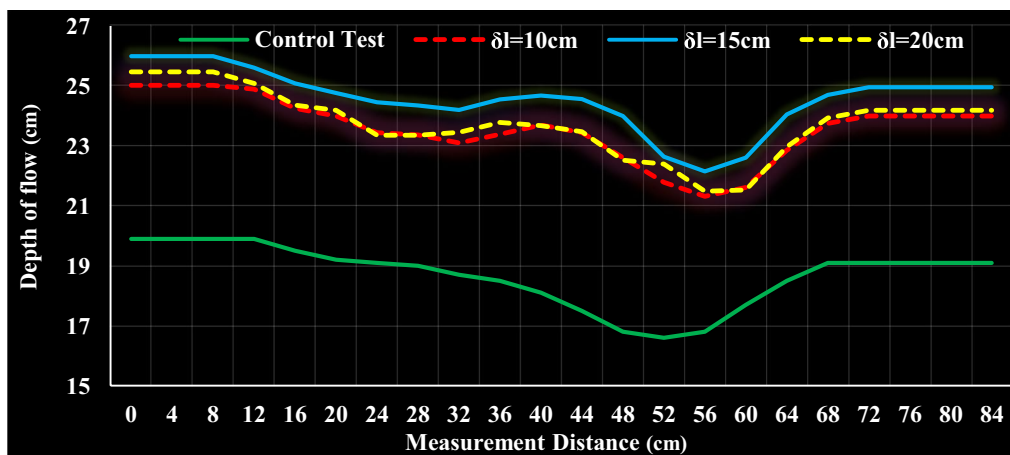


Fig. 9 Depths of flow in the main channel in $Q=25$ (l/s), $\theta=30^\circ$, $\beta=34^\circ$, and the applied δ_i

Table 7 Rate of genetic operators under the ideal scenario in GEP for both $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$

Inversion	0.1	Gene recombination	0.1	One-point recombination	0.3
Transportation	0.1				
Mutation	0.044	Insertion sequence transposition	0.1	Two-point recombination	0.3
Type of linking function	+				
Number of chromosomes	30	Root insertion sequence transposition	0.1	Head size	8
		Number of genes	3		

than a specified range, the original weights are then modified and improved with each iteration. The hybrid SVR-ACO model’s simulation process is visually shown in Fig. 8.

Performance evaluation metrics

In this investigation, the next statistical metrics are applied to assess the capability and accuracy of models used as follows (Kışı 2007):

$$\text{Determination coefficient } (R^2); R^2 = \left(\frac{\sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y)}{N\sigma_x\sigma_y} \right)^2 \tag{16}$$

$$\text{Root mean square error (RMSE); } RMSE = \sqrt{\frac{\sum_{i=1}^N (x_i - y_i)^2}{N}} \tag{17}$$

$$\text{Mean absolute error (MAE); } MAE = \frac{\sum_{i=1}^N |y_i - x_i|}{N} \tag{18}$$

$$\text{Mean bias error (MBE); } MBE = \frac{\sum_{i=1}^N (y_i - x_i)}{N} \tag{19}$$

where N is the number of datasets and x_i and y_i are the observed and forecasted values of target variables ($\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$), respectively. σ_x and σ_y are the standard deviations of observed and forecasted values of target variables, respectively. μ_x and μ_y are the average of observed and forecasted values of target variables, respectively. Lesser amounts for

Table 8 Statistical metrics in the testing stage under the ideal scenario of GEP for $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$

Statistical metrics (dimensionless)	$\frac{L_s}{B_2}$	$\frac{W_s}{B_2}$
R^2	0.93	0.92
RMSE	0.053	0.07
MAE	0.046	0.05
MBE	0.046	-0.043

$RMSE$, MAE , and MBE accompanied by higher amounts for R^2 denote a superior modeling performance.

Results and discussion

Variation of DFSZ

According to the experimental measurements, in the same value of δ_l , DFSZ increased by rising Q . By increasing Q , flow velocity increased and flow passed faster in the downstream of main channel rather than intake channel. Because in the same value of Q by increasing the value of δ_l from 10 to 20 (cm) the DFSZ increased, $\delta_l=10$ ponders as the optimal one.

Table 5 illustrates the maximum percentage reduction of $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$ as the representatives of DFSZ in the applied δ_l , β , and θ in comparison with the control test. Consistent with Table 5, the maximum percentage reduction occurs at $\theta=30^\circ$, $\beta=34^\circ$, and $\delta_l=10$ (cm) with the value of 78 and 76%, respectively.

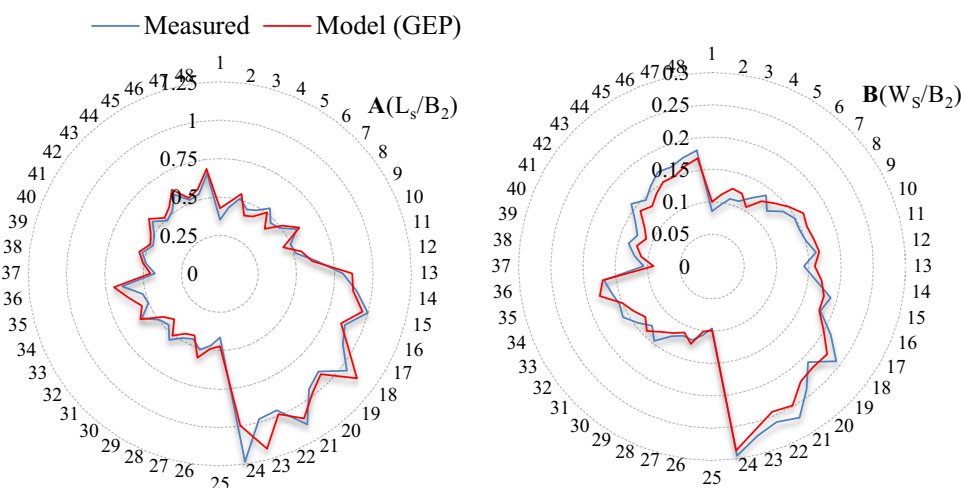
Variation of water depth in the main channel

Table 6 shows the maximum percentage of rising water depth (MPR_{WD}) by using the present submerged multiple-vanes system in the main channel compared to the control test at $\beta=34^\circ$ and the applied δ_l and θ . According to the experimental observations and Table 6, in comparison with the control test, water depth in the main channel by using the current submerged multiple-vane system noticeably increased. Moreover, based on Table 6, $\delta_l=15$ (cm) and $\theta=30^\circ$ are confirmed as the optimum ones. Figure 9 demonstrates water depth in the main channel at $Q=25$ (l/s) in the $\theta=30^\circ$, $\beta=34^\circ$, and applied δ_l .

Table 9 Kernel factors in SVR-RBF and SVR-polynomial models under the optimal scenario

Model C	γ	d	ϵ
SVR-RBF 100	0.01	-	0.001
SVR-Polynomial100	-	1	0.001

Fig. 10 Comparison plots between the observed and predicted, **A** $\frac{L_s}{B_2}$ and **B** $\frac{W_s}{B_2}$, in the testing stage under the ideal scenario of GEP method



Validation of GEP method

The value of genetic operators of GEP to estimate $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$ under the optimal scenario are presented in Table 7. The results of GEP model in the validation stage under the optimal scenario for $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$ are presented in Table 8.

According to Table 8, the value of *MBE* indicates that the method for the.

$\frac{L_s}{B_2}$ overestimated and for $\frac{W_s}{B_2}$ underestimated corresponding observed ones. Also, it can be concluded that the used structure in the optimum scenario can be employed with sufficient precision. Mathematical equations generated in the ideal scenario are presented in Eqs. (20) and (21) as follows:

$$\frac{L_s}{B_2} = (\log_{10}(\frac{L_V}{y_1} \times \sqrt[3]{4.73 \times \beta + \frac{q_2}{q_1} - \frac{\delta_l}{y_1}})^{0.25} + \text{Ln}(\frac{H_V}{y_1}) + \frac{q_2}{q_1}^{1.25} + 0.7 \frac{\delta_l}{y_1} \times \frac{\delta_l}{y_1} \tag{20}$$

$$\frac{W_s}{B_2} = (\text{Sin}\theta \times \frac{H_V}{y_1})^2 + (\text{Fr}_1 \times \sqrt[3]{\text{Atan}(\beta)}) + (\frac{L_V}{y_1} \times \frac{q_2}{q_1} \times \frac{\delta_l}{y_1}) \tag{21}$$

The estimated $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$ by Eqs. (20) and (21) are compared with the corresponding observed ones, and results are within 0–10%. Hence, it can be concluded that these equations can be applied as multivariate mathematical relationships with high precision. *R*² of these equations is in the range of 0.91 to 0.93, which statistically can be considered apposite equations only for the initial approximation of $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$ in hydraulic studies.

The exact prediction of $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$ mainly in the peak values shows the ability of model developed. Figure 10 A and B shows comparison plots between the observed and predicted $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$ by the ideal scenario in the testing stage, respectively.

Based on Fig. 10 A and B, it can be inferred that GEP method has partially failed to fit data, particularly in the peak

values that indicate some deviations in comparison with the measured ones.

Validation of SVR model

The values of *C*, ϵ , *d*, and γ in RBF and polynomial kernel functions in the optimal scenario are given in Table 9. Also, the number of input and output neurons were achieved 3 and 1, respectively. The values of statistical metrics in the testing stages under the optimal scenario are presented in Table 10.

According to Table 10, the value of *MBE* indicates that the models in the testing stage have overestimated corresponding observed ones. Also, it can be concluded that SVR-RBF has outperformed SVR-polynomial model.

Validation of SVR-ACO model

In this hybrid model, RBF is used as the appropriate kernel function based on the findings of the SVR model. The value of its optimized kernel parameters is obtained as *C* = 250, γ = 0.08, and ϵ = 0.009. Also, the optimal number of input and output neurons is obtained as 3 and 1, respectively.

Every inhabitant of the ACO scenario is adjusted by the hybrid SVR-ACO, which then outputs the best ACO. With 100

Table 10 Statistical metrics in the testing stage under the ideal scenario of SVR method for $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$

Variable	Model	Statistical metrics (dimensionless)			
		<i>R</i> ²	<i>RMSE</i>	<i>MAE</i>	<i>MBE</i>
$\frac{L_s}{B_2}$	SVR-RBF	0.92	0.062	0.05	0.05
$\frac{L_s}{B_2}$	SVR-polynomial	0.9	0.072	0.053	0.056
$\frac{W_s}{B_2}$	SVR-RBF	0.91	0.071	0.064	0.06
$\frac{W_s}{B_2}$	SVR-polynomial	0.9	0.082	0.072	0.068

Table 11 Statistical metrics in the validation stage under the ideal scenario of hybrid SVR-ACO model for $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$.

Variable	Statistical indices (dimensionless)			
	R^2	$RMSE$	MAE	MBE
$\frac{L_s}{B_2}$	0.99	0.029	0.022	0.012
$\frac{W_s}{B_2}$	0.98	0.036	0.027	0.019

as the population size, 1000 as the maximum number of repetitions, 20 as the initial pheromone, 0.02 as the heuristic exponential weight, 0.5 as the pheromone exponential weight, and 0.1 as the pheromone’s evaporation rate, the optimal scenario results in the optimal values of ACO parameters.

The values of statistical metrics in the validation stage for developed hybrid SVR-ACO in the optimal scenario are given in Table 11.

The value of MBE signifies that the model in the validation stage for the $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$ has overestimated corresponding observed ones. Consistent with Table 11, it can be inferred that the model can be utilized to predict with high accuracy. However, the performance of model in predicting $\frac{L_s}{B_2}$ is fairly higher than $\frac{W_s}{B_2}$.

Figure. 11 A and B demonstrates a comparison plot of the observed and forecasted in the validation stage under the optimal scenario

As seen in Fig. 11 A and B, because of the high ability of model, it can appositely capture the variation trend of the measured $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$, particularly in the peak and lowermost values. Figure 12 A and B demonstrates a scatter plot comparing the observed and forecasted $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$ in the validation stage by the optimal scenario, respectively.

It can be seen from the trend line equation (presume that the equation is $y = a_0x + a_1$) in the scatter plot that a_0 and a_1 coefficients for the model are, respectively, closer to 1 and 0 with a high value of R^2 .

Performance comparison of models used

Accompanied by Tables 8, 10, and 11 and Figs. 10, 11, and 12, it can be deduced that developed DDMs have almost acceptable precision with slight differences in the prediction of $DFSZ$. In this study, the best method is picked via the model grading process suggested by Vaheddoost et al. (2016). In this method, SG (success grade) and FG (failure grade) take into account performance pivotal metrics expressed as follows:

$$\text{Spearmanrhocorrelationcoefficient}(\rho); \rho = 1 - \frac{6 \sum d_i^2}{N^3 - N} \tag{22}$$

$$\text{Performanceindex}(PI); PI = \frac{1}{N} \sum_{i=1}^N \frac{y_i}{x_i} \tag{23}$$

$$SG(R^2)_i = \frac{R_i^2}{R_{max}^2} \times 10 \tag{24}$$

$$SG(CC)_i = \left| \frac{CC_i}{CC_{max}} \right| \times 10 \tag{25}$$

$$FG(PI)_i = \left(\frac{|1 - PI|_{min}}{|1 - PI|_i} \times 10 \right) - 10 \tag{26}$$

$$FG(RMSE)_i = \left(\frac{RMSE_{min}}{RMSE_i} \times 10 \right) - 10 \tag{27}$$

where N is the number of datasets and d_i is the difference in paired ranks; x_i and y_i are measured and forecasted values of $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$. Total grade (TG) of each model is gained by adding up FG and SG of each model distinctly that can differ between -20 and $+20$ (Table 12) and is expressed as

Fig. 11 Comparison plots between the observed and forecasted, **A** $\frac{L_s}{B_2}$ and **B** $\frac{W_s}{B_2}$, in the validation stage by the ideal scenario of hybrid SVR-ACO model

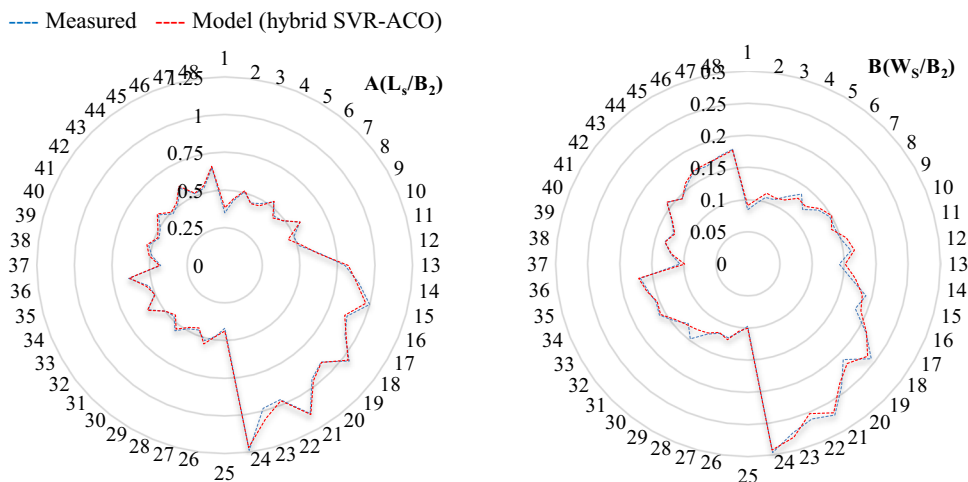
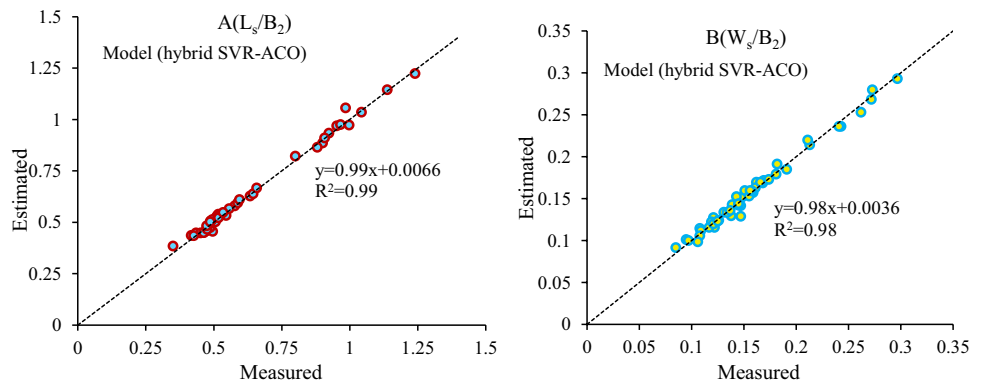


Fig. 12 Scatter plots between the observed and forecasted: **A** $\frac{L_s}{B_2}$ and **B** the $\frac{W_s}{B_2}$ in the validation stage by the optimal scenario of hybrid SVR-ACO model



$$TG = SG(CC) + SG(R^2) - FG(RMSE) - FG(PI) \quad (28)$$

Consistent with *TG* values in Table 12, the hybrid SVR-ACO model is considered the best method for the prediction of $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$.

Conclusion

In this research, the dimensions of flow separation zone (DFSZ) in a combined submerged multiple-vane system in diverse θ , δ_l , and Q on the steady and free overflow conditions were investigated. The technique used a butterfly array of applied vanes, which were composed of two portions, curved and flat plates. In order to evaluate the impact of the current system on DFSZ and flow pattern, extensive lab work was done over the scaled-down model in this regard. All experiments were carried out in a range of $0.1 < Fr_1 < 0.5$. Certain features about interaction effect were discussed, which can be supportive in design of submerged multiple-vane systems at initial phase. A summary of the experimental results is presented as follows:

1. Pre-processing by Spearman’s rho and Pearson’s correlation coefficients showed that the $\frac{q_2}{q_1}, \frac{\delta_b}{y_1}, \frac{\delta_l}{y_1}, \frac{\delta_r}{y_1}, \frac{L_v}{y_1}, \frac{H_v}{y_1}, \beta$ and $\theta, \beta, \frac{q_2}{q_1}, \frac{\delta_l}{y_1}, Fr_1, \frac{L_v}{y_1}, \frac{H_v}{y_1}$ were potential input variables to predict $\frac{L_s}{y_1}$ and $\frac{W_s}{y_1}$, respectively. Moreover,

sensitivity analysis by the cosine amplitude technique verified the $\frac{H_v}{y_1}$ and $\frac{L_v}{y_1}$ as the most effective variables in estimating the $\frac{L_s}{B_2}$ and $\frac{W_s}{B_2}$ as the target parameters, respectively.

2. In the same value of δ_l , DFSZ increased by increasing Q . Also, in the same value of Q , by increasing δ_l , DFSZ increased and, consequently, $\delta_l = 10$ (cm) was selected as the optimum option. trial
3. In comparison with the control test, the maximum percentage of reduction in DFSZ occurred at $\theta = 30^\circ$, $\beta = 34^\circ$, and $\delta_l = 10$ (cm) with the value of 78 and 76%, respectively. It caused a flow suitable length profile in the main and lateral intake channels and diminished flow turbulence in the intake position. Besides, the maximum percentage of rising water depth at $\beta = 34^\circ$ occurred at $\delta_l = 15$ (cm) and $\theta = 30^\circ$ with a value of 43%.
4. Based on the observations of flow-field, the specific nature of the present submerged multiple-vane system induced to generate more severe turbulence in the shear layer at the linking boundaries of the main channel and lateral intake channel compared to the control test.

Besides, in this study, several DDMs were characterized through a trial-and-error process to estimate DFSZ. Assessment of DDMs was conducted in the

Table 12 Results of all models in A and B, the test data set, and related grades

Target	Model	R^2	ρ	PI	$RMSE$	TG
$\frac{L_s}{B_2}$	GEP	0.93	0.96	1.058	0.053	3.94
	SVR-RBF	0.92	0.94	1.029	0.062	3.36
	SVR-Poly	0.9	0.91	1.017	0.072	2.19
	SVR-ACO	0.99	0.98	1.068	0.029	6.75
$\frac{W_s}{B_2}$	GEP	0.92	0.94	1.038	0.07	3.59
	SVR-RBF	0.91	0.9	1.022	0.071	3.31
	SVR-Poly	0.9	0.9	1.018	0.082	2.92
	SVR-ACO	0.98	0.98	1.061	0.036	5.8

testing stages. For this aim, several performance criteria, namely, R^2 , CC , PI , $RMSE$, MAE , MBE , and total grade (TG) were applied. The outcomes confirmed that the hybrid SVR-ACO model was the most effective and precise DDMs were created for the estimation of DFSZ. It should be emphasized that even though all possible DFSZ variables were evaluated in this work, the conclusions cannot be applied to other variations of the submerged multiple-vane system. The results of this study will help with the construction of hydraulic structures, diversion channels, sediment movement in rivers, and other things. The authors strongly encourage conducting experiments using various array types in open channels.

Acknowledgements The authors would like to thank Shahid Chamran University of Ahvaz for their support.

Author contribution Writing (original draft): Amin Gharehbaghi and Redvan Ghasemlounia; methodology: Amir Hamzeh Haghiabi; analysis: Abbas Parsaie; writing (review and editing): Sarmad Dashti Latif.

Declarations

Conflict of interest The authors declare no competing interests.

References

- Barkdoll B, Ettema R, Odgaard A (1999) Sediment control at lateral diversions: limits and enhancements to vane use. *J Hydraul Eng*. [https://doi.org/10.1061/\(ASCE\)0733-9429\(1999\)125:8\(862\)](https://doi.org/10.1061/(ASCE)0733-9429(1999)125:8(862))
- Behbahan TS (2011) Laboratory investigation of submerged vane shapes effect on river banks protection. *Aust J Basic Appl Sci* 5(12):1402–1407
- Borghei SM, Jalili MR, Ghodsian M (1999) Discharge coefficient for sharp-crested side weir in subcritical flow. *J Hydraul Eng* 125(10):1051–1056
- Dorigo M, Maniezzo V, Colorni A (1996) Ant system: optimization by a colony of cooperating agents. *IEEE Trans Syst Man, Cybern Part B Cybern* 26:29–41. <https://doi.org/10.1109/3477.484436>
- Dorigo M, Di Caro G, Gambardella LM (1999) Ant algorithms for discrete optimization. *Artif Life* 5:137–172. <https://doi.org/10.1162/106454699568728>
- Ehteram M, Ahmed AN, Latif SD et al (2020) Design of a hybrid ANN multi-objective whale algorithm for suspended sediment load prediction. *Environ Sci Pollut Res*. <https://doi.org/10.1007/s11356-020-10421-y>
- Ferreira C (2001) Gene expression programming: a new adaptive algorithm for solving problems. *Complex Syst*. <https://doi.org/10.48550/arXiv.cs/0102027>
- Gharehbaghi A, Ghasemlounia R (2022) Application of AI approaches to estimate discharge coefficient of novel kind of sharp-crested V-notch Weirs. *J Irrig Drain Eng* 148:. [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0001646](https://doi.org/10.1061/(ASCE)IR.1943-4774.0001646)
- Gharehbaghi A, Kaya B, Saadatnejadgharahassanlou H (2016) Numerical simulation of two dimensional unsteady flow by total variation diminishing scheme. *Int J Eng Appl Sci* 8:1–1. <https://doi.org/10.24107/ijeas.255030>
- Gharehbaghi A, Kaya B, Saadatnejadgharahassanlou H (2017) Two-dimensional bed variation models under non-equilibrium conditions in turbulent streams. *Arab J Sci Eng* 42:999–1011. <https://doi.org/10.1007/s13369-016-2258-4>
- Gharehbaghi A, Ghasemlounia R, Ahmadi F, Albaji M (2022) Groundwater level prediction with meteorologically sensitive gated recurrent unit (GRU) neural networks. *J Hydrol* 612:128262. <https://doi.org/10.1016/j.jhydrol.2022.128262>
- Ghasemlounia R, Gharehbaghi A, Ahmadi F, Saadatnejadgharahassanlou H (2021) Developing a novel framework for forecasting groundwater level fluctuations using bi-directional long short-term memory (BiLSTM) deep neural network. *Comput Electron Agric* 191:106568. <https://doi.org/10.1016/j.compag.2021.106568>
- Gupta UP, Ojha CSP, Sharma N (2010) Enhancing utility of submerged vanes with collar. *J Hydraul Eng* 136:651–655. [https://doi.org/10.1061/\(asce\)hy.1943-7900.0000212](https://doi.org/10.1061/(asce)hy.1943-7900.0000212)
- Karami H, Farzin S, Sadrabadi MT, Moazeni H (2017) Simulation of flow pattern at rectangular lateral intake with different dike and submerged vane scenarios. *Water Sci Eng* 10:246–255. <https://doi.org/10.1016/j.wse.2017.10.001>
- Kaya B, Gharehbaghi A (2012) Modelling of sediment transport with finite volumes method under unsteady conditions. *J Fac Eng Archit Gazi Univ* 27:26–27
- Keshavarzi A, Habibi L (2005) Optimizing water intake angle by flow separation analysis. *Irrig Drain* 54:543–552. <https://doi.org/10.1002/ird.207>
- Kişçi Ö (2007) Streamflow forecasting using different artificial neural network algorithms. *J Hydrol Eng* 12:532–539. [https://doi.org/10.1061/\(asce\)1084-0699\(2007\)12:5\(532\)](https://doi.org/10.1061/(asce)1084-0699(2007)12:5(532))
- Latif SD, Ahmed AN (2021) Application of deep learning method for daily streamflow time-series prediction : a case study of the Kowmung River at Cedar Ford, Australia. *Int J Sustain Dev Plan* 16:497–501. <https://doi.org/10.18280/ijstdp.160310>
- Latif SD, Azmi MSBN, Ahmed AN, et al (2020) Application of artificial neural network for forecasting nitrate concentration as a water quality parameter: a case study of Feitsui Reservoir, Taiwan. *Int J Des Nat Ecodynamics*. <https://doi.org/10.18280/ijdne.150505>
- Latif SD, Ahmed AN, Sathiamurthy E, et al (2021a) Evaluation of deep learning algorithm for inflow forecasting : a case study of Durian Tunggal Reservoir, Peninsular Malaysia. *Nat Hazards*. <https://doi.org/10.1007/s11069-021-04839-x>
- Latif SD, Birima AH, Najah A et al (2021b) Development of prediction model for phosphate in reservoir water system based machine learning algorithms. *Ain Shams Eng J*. <https://doi.org/10.1016/j.asej.2021.06.009>
- Lawrence S, Back AD, Tsoi AC, Giles CL (1997) On the distribution of performance from multiple neural-network trials. *IEEE Trans Neural Networks* 8:1507–1517. <https://doi.org/10.1109/72.641472>
- Mehdizadeh S, Behmanesh J, Khalili K (2017) Application of gene expression programming to predict daily dew point temperature. *Appl Therm Eng* 112:1097–1107. <https://doi.org/10.1016/j.applthermaleng.2016.10.181>
- Mohammadi B, Mehdizadeh S (2020) Modeling daily reference evapotranspiration via a novel approach based on support vector regression coupled with whale optimization algorithm. *Agric Water Manag* 237:106145. <https://doi.org/10.1016/j.agwat.2020.106145>
- Momeni E, Nazir R, JahedArmaghani D, Maizir H (2014) Prediction of pile bearing capacity using a hybrid genetic algorithm-based ANN. *Meas J Int Meas Confed* 57:122–131. <https://doi.org/10.1016/j.measurement.2014.08.007>
- Monier JF, Gao F, Boudet J, Shao L (2020) Turbulence modelling analysis in a corner separation flow. *Comput Fluids* 213: <https://doi.org/10.1016/j.compfluid.2020.104745>
- Najah A, Teo FY, Chow MF, et al (2021) Surface water quality status and prediction during movement control operation order under

- COVID-19 pandemic: case studies in Malaysia. *Int J Environ Sci Technol* <https://doi.org/10.1007/s13762-021-03139-y>
- Nakato BT, Kennedy JF, Bauerly D (1990) Pump-station intake-shoaling control with submerged vanes. *J Hydraul Eng* 116:119–128
- Odgaard AJ (2009) River training and sediment management with submerged vanes. ASCE Press. <https://doi.org/10.1061/9780784409817>
- Odgaard AJ, Spoljaric A (1988) Sediment control by submerged vanes. *J Hydraul Eng* 112:1164–1180. [https://doi.org/10.1061/\(asce\)0733-9429\(1986\)112:12\(1164\)](https://doi.org/10.1061/(asce)0733-9429(1986)112:12(1164))
- Odgaard AJ, Wang Y (1991a) Sediment management with submerged vanes. I: Theory. *J Hydraul Eng* 117:267–267. [https://doi.org/10.1061/\(asce\)0733-9429\(1991\)117:3\(267\)](https://doi.org/10.1061/(asce)0733-9429(1991)117:3(267))
- Odgaard AJ, Wang Y (1991b) Sediment management with submerged vanes. II: APPLICATIONS. *J Hydraul Eng* 117:284–302
- Olatomiwa L, Mekhilef S, Shamshirband S et al (2015) A support vector machine-firefly algorithm-based model for global solar radiation prediction. *Sol Energy* 115:632–644. <https://doi.org/10.1016/j.solener.2015.03.015>
- Ouyang H-T (2009) Investigation on the dimensions and shape of a submerged vane for sediment management in alluvial channels. *J Hydraul Eng* 135:209–217. [https://doi.org/10.1061/\(asce\)0733-9429\(2009\)135:3\(209\)](https://doi.org/10.1061/(asce)0733-9429(2009)135:3(209))
- Ouyang HT, Lin CP (2016) Characteristics of interactions among a row of submerged vanes in various shapes. *J Hydro-Environment Res* 13:14–25. <https://doi.org/10.1016/j.jher.2016.05.003>
- Pai P-F, Hong W-C (2006) A recurrent support vector regression model in rainfall forecasting. *Hydrol Process* <https://doi.org/10.1002/hyp.6323>
- Qasem SN, Samadianfard S, Nahand HS, et al (2019) Estimating daily dew point temperature using machine learning algorithms. *Water (Switzerland)* 11: <https://doi.org/10.3390/w11030582>
- Ramamurthy AS, Qu J, Vo D (2007) Numerical and experimental study of dividing open-channel flows. *J Hydraul Eng* 133:1135–1144. [https://doi.org/10.1061/\(asce\)0733-9429\(2007\)133:10\(1135\)](https://doi.org/10.1061/(asce)0733-9429(2007)133:10(1135))
- Sharma H, Ahmad Z (2020) Turbulence characteristics of flow past submerged vanes. *Int J Sediment Res* 35:42–56. <https://doi.org/10.1016/j.ijsrc.2019.07.002>
- Smola AJ, Schölkopf B (2004) A tutorial on support vector regression. *Stat, Comput*
- Sui B, Huang SH (2017) Numerical analysis of flow separation zone in a confluent meander bend channel. *J Hydrodyn* 29:716–723. [https://doi.org/10.1016/S1001-6058\(16\)60783-7](https://doi.org/10.1016/S1001-6058(16)60783-7)
- Taherei Ghazvinei P, Darvishi HH, Mosavi A et al (2018) Sugarcane growth prediction based on meteorological parameters using extreme learning machine and artificial neural network. *Eng Appl Comput Fluid Mech* 12:738–749. <https://doi.org/10.1080/19942060.2018.1526119>
- Tofiq YM, Latif SD, Ahmed AN et al (2022) Optimized model inputs selections for enhancing river streamflow forecasting accuracy using different artificial intelligence techniques. *Water Resour Manag*. <https://doi.org/10.1007/s11269-022-03339-2>
- Vaheddoost B, Aksoy H, Abghari H (2016) Prediction of water level using monthly lagged data in Lake Urmia. *Iran Water Resour Manag* 30:4951–4967. <https://doi.org/10.1007/s11269-016-1463-y>
- Wang Y, Odgaard AJ (1993) Flow control with vorticity: Modification d'un écoulement au moyen de tourbillons. *J Hydraul Res* 31:549–562. <https://doi.org/10.1080/00221689309498877>
- Wang Y, Odgaard A, Melville BW, Jain SC (1996) Sediment control at water-intakes. *J Hydraul Eng*. [https://doi.org/10.1061/\(ASCE\)0733-9429\(1996\)122:6\(353\)](https://doi.org/10.1061/(ASCE)0733-9429(1996)122:6(353))
- Yu PS, Chen ST, Chang IF (2006) Support vector regression for real-time flood stage forecasting. *J Hydrol* 328:704–716. <https://doi.org/10.1016/j.jhydrol.2006.01.021>
- Yu X, Zhang X, Qin H (2018) A data-driven model based on Fourier transform and support vector regression for monthly reservoir inflow forecasting. *J Hydro-Environment Res*. <https://doi.org/10.1016/j.jher.2017.10.005>
- Zaji AH, Bonakdari H, Khodashenas SR, Shamshirband S (2016) Firefly optimization algorithm effect on support vector regression prediction improvement of a modified labyrinth side weir's discharge coefficient. *Appl Math Comput* 274:14–19. <https://doi.org/10.1016/j.amc.2015.10.070>
- Ziyad Sami BF, Latif SD, Ahmed AN et al (2022) Machine learning algorithm as a sustainable tool for dissolved oxygen prediction: a case study of Feitsui Reservoir. *Taiwan Sci Rep* 12:1–12. <https://doi.org/10.1038/s41598-022-06969-z>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.