WILEY | Hindawi

Research Article

Assessment of Artificial Intelligence Models for Estimating Lengths of Gradually Varied Flow Profiles

Majid Niazkar (ID),[1] **Farshad Hajizadeh mishi** (ID),[1] **and Gökçen Eryılmaz Türkkan** (ID)[2]

[1]*Department of Civil and Environmental Engineering, School of Engineering, Shiraz University, Zand Blvd, Shiraz, Iran*
[2]*Department of Civil Engineering, Bayburt University, Bayburt, Turkey*

Correspondence should be addressed to Majid Niazkar; mniazkar@shirazu.ac.ir

The study of water surface profiles is beneficial to various applications in water resources management. In this study, two artificial intelligence (AI) models named the artificial neural network (ANN) and genetic programming (GP) were employed to estimate the length of six steady GVF profiles for the first time. The AI models were trained using a database consisting of 5154 dimensionless cases. A comparison was carried out to assess the performances of the AI techniques for estimating lengths of 330 GVF profiles in both mild and steep slopes in trapezoidal channels. The corresponding GVF lengths were also calculated by 1-step, 3-step, and 5-step direct step methods for comparison purposes. Based on six metrics used for the comparative analysis, GP and the ANN improve five out of six metrics computed by the 1-step direct step method for both mild and steep slopes. Moreover, GP enhanced GVF lengths estimated by the 3-step direct step method based on three out of six accuracy indices when the channel slope is higher and lower than the critical slope. Additionally, the performances of the AI techniques were also investigated depending on comparing the water depth of each case and the corresponding normal and critical grade lines. Furthermore, the results show that the more the number of subreaches considered in the direct method, the better the results will be achieved with the compensation of much more computational efforts. The achieved improvements can be used in further studies to improve modeling water surface profiles in channel networks and hydraulic structure designs.

1. Introduction

Gradually varied flow (GVF) is a nonuniform flow in natural and man-made canals. The study of GVF is crucial to water resources management as it may not only be categorized as one of the most common flow conditions in an open channel but also play a key role in various hydraulic projects. Some examples of the occurrence of GVF include flow through a change in channel bottom slope, canal constrictions and transitions, a variation of channel geometries, flow under the infection of hydraulic structures, and flow from a large reservoir to a canal. In such situations, flow variables, i.e., water depth and flow velocity, vary gradually in each cross-section along a channel.

The governing equation for computing GVF profiles in prismatic canals is shown in equation (1). It is basically a combination of energy (or momentum) equation and a resistance equation. The former presents the spatial variation of water depth in GVF profiles, while the latter relates friction slope (S_f) with flow and channel geometries of the canal under consideration:

$$\frac{dy}{dx} = S_0 \frac{1 - (y_n/y)^N}{1 - (y_c/y)^M},\tag{1}$$

where y is the water depth, x is the longitudinal distance along the channel, dy/dx is the water surface slope, y_n and y_c are the normal and critical water depths, respectively, S_0 is the channel bottom slope, and M and N are the hydraulic components for critical and uniform flows, respectively.

One of typical problem statements in GVF profiles is computation of the distance between two specific water depths. In other words, water depths at two cross-sections of the same profile are given while the distance between these

two sections (L) is unknown. According to the literature review, various attempts for solving this problem may be categorized into several groups based on their methods: (1) semianalytical methods [1, 2], (2) analytical solutions [3–7], (3) numerical schemes [8–11], (4) artificial intelligence (AI) models [12], and (5) optimization techniques [13–15].

The disadvantages of semianalytical and analytical solutions include (1) some of them are only applicable to specific conditions such as Bresse's analytical solution for wide rectangular channel and Chezy's equation and (2) the analytical solutions with a wide of range of applicability mostly have complex relations. On the other hand, the numerical schemes basically march on space between the two water depths are given. Additionally, they have been known to be susceptible to stability problems [16], while they may sometimes achieve different results [17]. Although analytical solutions are error-free, the accuracy of numerical solutions depends on several factors including the spatial interval (Δx) and the round-off characteristics of the method, which may lead to discretization and truncation errors, respectively [11]. Based on the current literature, application of AI models to GVF computation is limited. For instance, Sivapragasam et al. [12] utilized genetic programming (GP) and the artificial neural network (ANN) to predict water surface profile as a steady flow with different discharges passes over a rectangular notch. Although AI models have been successfully used for solving numerous problems in water resources management and hydraulic engineering [18–21], it has not been applied to estimate the length of GVF profiles.

In this study, two AI models were employed to predict the distance between two cross-sections with known water depths in the same GVF profile. A large database was provided for different flow conditions in rectangular and trapezoidal cross-sections. The performances of these models in estimating length of GVF profiles were also compared with those of the most common numerical method available in the literature.

2. Methods and Materials

2.1. Problem Statement of GVF Profile Length.
Steady GVF is one of the most common time-independent flow conditions occurred in open channels, while length of GVF profiles is necessary to channel design, design of hydraulic structures, and budget estimation of open-channel water conveyance projects [1, 22]. In the current literature, Swamee [1] presented empirical relations between the control section and the section with $0.99 y_n$ or $1.01 y_n$ for triangular, wide, and narrow rectangular. However, in this study, the distance between two arbitraries but known water depths (y_1 and y_2) within the same profile in trapezoidal sections is of interest.

In addition to equation (1), Manning's equation, which is the most widely used resistance equation in open-channel hydraulics [23], governs the flow filed in GVF profiles:

$$Q = \frac{1}{n} A R^{2/3} S^{1/2}, \qquad (2)$$

where Q is the discharge, n is Manning's coefficient, A is the flow area, $R = A/P$ is the hydraulic radius, P is the wetter

perimeter, and S is the channel slope. When $S = S_0$, the water depth in equation (2) exclusively corresponds to y_n, while it can be any other water depth for $S = S_f$.

In the problem statement of computing the length of GVF profiles, Q channel geometries including canal bottom width (b), channel side slope (z), S_0, and n are the given information, while L is meant to be estimated.

The distance of varied flows is a determinant parameter in various water engineering problems, such as determining the location of hydraulic jump [13], predicting budget of channel design [24, 25], and estimating backwater impacts on hydraulic structures [26]. For instance, the influence length of GVF profiles propagating from a uniform or critical depth has been investigated in the literature [11, 27]. In modeling water surface of GVF profiles in real-life projects, the routine procedure in professional hydraulic software, such as HEC-RAS, is to divide the canal under consideration into several parts so called reaches so that each reach as similar flow conditions and channel geometries [28]. Due to spatial variation of channel geometries along natural streams, a reach-average value has been frequently designated to canal characteristics such as b, z, S_0, and n [28, 29]. As n is known to be a flow-dependent parameter and is not a measurable parameter [26, 30], a flow-independent bed roughness predictor may be utilized to estimate a reach-average value for steady GVF.

Steady GVF profiles is generally categorized based on the comparison between y_n and y_c: (1) when $y_n = y_c$, the channel slope is called critical slope ($S_0 = S_c$), (2) if $y_n > y_c$, the channel slope is mild slope ($S_0 < S_c$), and (3) steep channel occurs ($S_0 > S_c$) when $y_n < y_c$. In this classification, when water depth is higher than both y_n and y_c, it is located in M1 or S1 zone if the channel slope is mild or steep, respectively. Furthermore, when water depth is located between y_n and y_c, it may be called either M2 or S2 for mild or steep slope, respectively. Finally, M3 or S3 profile happens when the water depth is lower than both y_n and y_c in mild or steep slope, respectively.

2.2. Varied Flow Function.
In this method, it is commonly assumed that M and N that appeared in equation (1) are the flow invariants for simplicity. In other words, the variation of M and N with water depth can be neglected in engineering application [31]. Using the substitution of $u = y/y_n$, equation (1) is rewritten:

$$x = \frac{y_n}{S_0} \left[u - \int_0^u \frac{\mathrm{d}u}{1 - u^N} + \left(\frac{y_c}{y_n} \right)^M \int_0^u \frac{u^{N-M} \mathrm{d}u}{1 - u^N} \right] + c, \qquad (3)$$

where x is the location of a specific cross-section along the channel, $M = 3[1 + 2z(y/b)]^2 - 2z(y/b)[1 + z(y/b)]/[1 + 2z(y/b)][1 + z(y/b)]$, $N = (10/3)(1 + 2z(y/b)/1 + z(y/b)) - (8/3)(\sqrt{1 + z^2}(y/b)/1 + 2\sqrt{1 + z^2}(y/b))$ for trapezoidal cross-sections, and c is a constant.

The first integral term in equation (3) is called varied flow function ($F(u, N) = \int_0^u \mathrm{d}u/1 - u^N$), which depends on two parameters u and N. The second integral in equation (3) can be turned into a varied flow function by setting $v = u^{N/J}$,

where $J = (N/N - M + 1)$. By further manipulating equation (3), the distance between two cross-sections 1 and 2

with known water depths is achieved by the following equation:

$$L = \frac{y_n}{S_0} \left\{ u_2 - u_1 + F(u_1, N) - F(u_2, N) + \frac{J}{N} \left(\frac{y_c}{y_n} \right)^M [F(v_2, J) - F(v_1, J)] \right\}, \tag{4}$$

where 1 and 2 subscripts correspond to the first and second cross-sections in the channel reach.

Traditionally, magnitudes of the varied flow function are provided in tables covering numerous values of state variables [6]. Although this method has been introduced as a standard approach in hydraulic engineering texts [31, 32], its major drawback is determination of the varied flow function [6, 33]. The reasons for which this method is not suitable in practice may be as follows: (1) the varied flow function is relatively complicated to be solved by the hand, (2) according to equation (4), four integrals $(F(u_1, N), F(u_2, N), F(v_2, J),$ and $F(v_1, J))$ need to be computed to determine L, and (3) the tables provided for varied flow function can only be exploited for a set of discrete values of state variables. Consequently, an interpolation may be required for intermittent values, which produces inevitable errors. On the other hand, when varied flow function is calculated without interpolation, the obtained results may be used as benchmark since it considers no approximate assumption.

2.3. Simple Direct Method. In this method, the distance between two cross-sections 1 and 2 is computed by using finite difference:

$$L = \frac{1}{S_0 - S_f} \left[\left(y_2 + \frac{Q^2}{2gA_2^2} \right) - \left(y_1 + \frac{Q^2}{2gA_1^2} \right) \right], \tag{5}$$

where g is the gravitational acceleration.

As shown equation (5), L can be computed directly when Q and channel properties are known, while S_f in equation (5) is substituted with the reach-average friction slopes using the first and second cross-sections. In this study, the direct method is utilized for calculating GVF profile length using (1) one-spatial step, (2) three-spatial steps, and (3) five-spatial steps. In the one-step version, equation (5) is exploited only one time to obtain L between the two water depths given in the problem statement, while the whole distance between sections 1 and 2 is divided into three and five subreaches in the three-step and five-step versions, respectively. For the better clarification, Figure 1 depicts schematic division of a channel reach into five subreaches in the five-step scenario. As shown, y_n and y_c are the same for all five subreaches in the five-step direct method. According to Figure 1, four additional water depths are required to be used between the first and second sections in the five-step direct method. For $y_1 > y_2$, the four additional water depths are $y_3 = y_1 - (y_1 - y_2/5)$, $y_4 = y_3 - (y_1 - y_2/5)$, $y_5 = y_4 - (y_1 - y_2/5)$, and $y_6 = y_5 - (y_1 - y_2/5)$. Therefore, when y_1

and y_2 are given, the additional water depths can be calculated one after another from y_3 to y_6. In the five-step direct method, the distance between two consecutive sections, such as L_{13} between sections 1 and 3, is computed using equation (5). Finally, the algebraic summation of the distances between the successive cross-sections is computed, which is equal to L as shown in Figure 1.

2.4. Artificial Neural Network. Artificial neural network (ANN) is a well-documented AI model and has been successfully applied to various problems in water resources and hydraulic engineering [34, 35]. Basically, it consists of three layers, named as input, hidden, and output layers, while each layer includes some components called neurons. The number and objective of neurons are defined based on the layer to which they belong. For instance, the neurons of the input layer take the vector of input data. The structure of the ANN provides connections between neurons of two successive layers, while there is no connection between the neurons within a layer. Using these connections, the data flow through the network until an adequate relation between the input and output data is achieved [30].

Predicting the length of GVF profiles using the ANN is conducted for the first time in this study, based on authors' knowledge. The input data include the dimensionless u and N, while the output data are dimensionless $F(u, N)$. Furthermore, there is a trade-off between the number of neurons in the hidden layer in the ANN and the computational efforts. To be more specific, the more the number of neurons in the hidden layer is used, the more accurate results may be obtained. In this study, several hidden layers with four to ten neurons were tested, while a seven-neuron hidden layer was selected. After the ANN completed the prediction process, the estimated varied flow functions were substituted into equation (4) to compute the corresponding length of GVF profiles. As shown in equation (4), four $F(u, N)$ are required to determine the length between two specified water depths.

2.5. Genetic Programming. Genetic programming (GP) is one of AI models that employs the genetic algorithm to create a powerful prediction tool. In essence, GP is an improved version of the genetic algorithm which is capable of finding a relation between two vectors of variables regardless of the physical background of those data. GP begins with creation of a random population comprising of random functions and coefficients [36]. It also uses the genetic algorithm features such as crossover and mutation to improve the fitness of new generations in light of minimizing an objective function. The objective function basically reflects
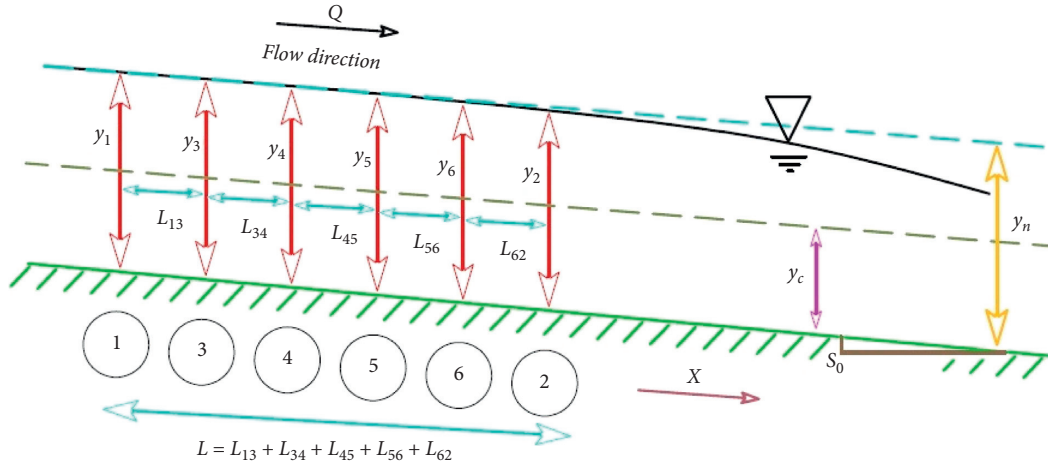
Figure 1: Schematic division of a channel reach into several subreaches.

the errors between the input and output data. This process continues until an expression with a desirable error is achieved. Such correlations may further be used for estimation purposes [25].

GP has a tree-like structure in which a variety of mathematical functions and variables may be adopted to seek for an appropriate relationship between the input and output data. As a result of these characteristics, Discipulus [37] software, which has been used for applying GP in the literature [35], was exploited to many problems in the fields' water resources and hydraulic engineering. The input data given to this program include u and N, while the output data were $F(u, N)$. The latter values were exploited to estimate L for each data point.

2.6. The Database. The data considered in this study consist of two parts: (1) train data and (2) test data. The former includes 5154 rows of u, N (input data), and $F(u, N)$ (output data). They were basically gathered from the tables presented in engineering text books [31, 38]. These data were used for training the ANN and GP. On the other hand, the second part of data consists of 165 data for mild slope and 165 data for steep slope. To be more specific, the test data contain 70 data of M1, 54 data of M2, 41 data of M3, 48 data of S1, 51 data of S2, and 66 data of S3 profile. Furthermore, the values of Q, b, z, S_0, n, y_1, and y_2 of the test data were generated by the random function embedded in Excel [39]. Finally, y_n and y_c can be computed when Q and the reach-average values of n and S_0 are given.

Since the test data were developed randomly, they need to be checked. In this regard, three requirements were considered: (1) channel geometries and Q of each row should be practically feasible, (2) each row of data should only have one type of channel slope (i.e., mild or steep slope), and (3) each row of data should belong to one specific type of flow profile. For instance, the GVF profile is M1 or M2 or M3 when the channel slope is mild. In case a row of data did not satisfy the mentioned requirements, it was replaced with another randomly generated row of data to keep 330 rows of data, which correspond to 1320 pairs of u and N.

After checking each row of the developed data, u and N were determined for the specified water depths (y_1 and y_2). They were used to compute the corresponding $F(u, N)$ without interpolation. Table 1 presents the ranges of different parameters in the test data. As shown, the train data have a wider range of values than that of the test data. According to Table 1, the test data include a wide range of values for each parameter involved.

As previously mentioned, four $F(u, N)$ are required to calculate L for each row of data. Thus, the calculated values of $F(u, N)$ were substituted into equation (4) to give the benchmark distance between y_1 and y_2 for each row of the test data. Finally, the test data were also solved by other methods considered in this study for comparison purposes.

2.7. Performance Evaluation Criteria. Six evaluation criteria were selected from the literature to compare the performances of different methods in estimation of the length of GVF profiles between two specified water depths [40, 41]. These indices are (1) root mean square error (RMSE), (2) mean absolute relative error (MARE), (3) mean absolute error (MAE), (4) relative absolute error (RAE), (5) relative squared error (RSE), and (6) coefficient of determination (R^2). These criteria are presented in equations (6)–(11), respectively:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left[L_{\text{estimated}} - L\right]^2}, \tag{6}$$

$$\text{MARE} = \frac{1}{N} \sum_{i=1}^{N} \left|\frac{L_{\text{estimated}} - L}{L}\right|, \tag{7}$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} \left|L_{\text{estimated}} - L\right|, \tag{8}$$

$$\text{RAE} = \frac{\sum_{i=1}^{N} \left|L_{\text{estimated}} - L\right|}{\sum_{i=1}^{N} \left|L - \left(\sum_{i=1}^{N} L/N\right)\right|}, \tag{9}$$

TABLE 1: Range of different parameters in the database.

Parameter	Mean	Maximum	Minimum	Standard deviation
Train data				
u	1.48	20.00	0.00	2.19
N	5.25	9.80	2.20	2.37
$F(u, N)$	0.49	3.52	0.00	0.50
Test data				
u	2.10	19.75	0.00	3.24
N	3.04	5.88	2.41	0.42
$F(u, N)$	0.42	2.98	0.00	0.42
Q	104.57	494.00	1.00	70.68
b	30.40	95.00	3.00	14.41
z	1.23	3.50	0.00	0.95
S	0.01	0.10	0.00	0.01
n	0.02	0.05	0.01	0.01
y_n	1.61	10.13	0.05	1.31
y_1	1.95	20.00	0.02	2.42
y_2	2.11	17.00	0.01	2.35
y_c	1.13	6.41	0.06	0.76

$$\text{RSE} = \frac{\sum_{i=1}^{N} \left(L_{\text{estimated}} - L\right)^2}{\sum_{i=1}^{N} \left(\left(\sum_{i=1}^{N} L/N\right) - L\right)^2}, \tag{10}$$

$$R^2 = \left(\frac{\sum_{i=1}^{N}\left[\left(L - \left(\sum_{i=1}^{N} L/N\right)\right)\left(L_{\text{estimated}} - \left(\sum_{i=1}^{N} L_{\text{estimated}}/N\right)\right)\right]}{\sqrt{\sum_{i=1}^{N}\left[\left(L - \left(\sum_{i=1}^{N} L/N\right)\right)^2\left(L_{\text{estimated}} - \left(\sum_{i=1}^{N} L_{\text{estimated}}/N\right)\right)^2\right]}}\right)^2. \tag{11}$$

where $L_{\text{estimated}}$ is the estimated length, i is a counter, and N is the number of data.

3. Results and Discussion

The ANN, GP, varied flow function, and 1-step direct method compute the length of GVF profile between y_1 and y_2 by considering one channel reach. However, 3-step and 5-step direct methods divided the channel reach into 3 and 5 subreaches, respectively. These methods were used for estimation of GVF profile length between two specified water depths. For comparison purposes, the test data consist of different GVF profiles to investigate performances of different methods in prediction of L between y_1 and y_2 of each row of the test data.

The performances of the methods described for calculating GVF profile lengths were compared for the test data with mild slopes shown in Table 2. As shown, the 5-step direct method achieved the closet results to the benchmark solutions in comparison with 1-step and 3-step direct methods. Thus, the accuracy of the direct step method enhances with the increase of the number of spatial intervals considered. Additionally, Table 2 depicts that the ANN obtained better RMSE, RAE, MAE, RSE, and R^2 than the 1-step direct method, while the latter achieved a better MARE than the former. Based on Table 2, GP performed better than 1-step and 3-step direct methods in terms of RMSE and RSE, while it yields to the same R^2 as the 3-step direct method. According to Table 2, the 5-step direct method outperformed others for estimating the length of GVF profiles

TABLE 2: Comparison of different methods for calculating distance of GVFs with mild slopes.

Method	RMSE	MARE	R^2	RAE	MAE	RSE
Direct method (1-step)	2068.69	0.25	0.89	0.18	735.29	0.12
Direct method (3-step)	799.92	0.07	0.99	0.06	242.64	0.02
Direct method (5-step)	429.3	0.04	1	0.03	132.61	0.01
ANN (this study)	1139.68	4.36	0.97	0.15	627.09	0.04
GP (this study)	710.16	1.21	0.99	0.08	321.89	0.01

having mild slopes, while GP performs similar to the 3-step direct method.

The improvement obtained by the AI techniques may be interpreted in the light of the connectivity topic. In a holistic point of view, a better understanding of connectivity may be beneficial to develop better schemes for modeling water resources. To be more specific, estimating the distance between two known water depths in a GVF profile may improve modeling of water movement through a man-made canal or a natural stream. This may help river engineering, which includes dam construction/removal, river restoration, and channel regulations. Since, river engineering is connected with channel processes and geomorphic channel response in small- to meso-scale fluvial systems [42]. Therefore, the improvement provided by the AI techniques in this specific application does not confine in design of open channels, particularly when catchment connectivity is assessed [43].

Table 3 compares different methods for calculating L for trapezoidal sections with steep slopes. As shown, the best RMSE, RSE, and R^2 were achieved by GP. Moreover, RMSE obtained by the ANN was lower than that of all direct step methods, while R^2 calculated by the ANN was equal to the best R^2 computed by the latter. According to Table 3, the 5-step direct method achieved better metrics than the 3-step direct method, while the latter reached better results than the 1-step direct method. Thus, Tables 2 and 3 imply that the fewer the spatial interval (Δx) is considered in a typical numerical scheme, the more accurate the results achieve and the more the computational efforts are required. Furthermore, the comparisons carried out in Tables 2 and 3 indicate that GP is an accurate method for computing the length of GVF profiles, particularly when channel has a mild slope.

Although Tables 2 and 3 compared the performances of different methods in estimation of L, they do not provide clearly how each considered method performs in each of GVF profiles. In this regard, six indices were separately calculated by each method for each GVF profiles. The obtained results are depicted in Figure 2, which provide a better detailed perspective required for comparison purposes. Figure 2 shows that RMSE values achieved for mild slopes, particularly M2 profile, are larger than that of steep slopes. This clearly addresses why RMSE values of mild slope shown in Tables 2 and 3 are relatively larger than that of steep slopes. According to Figures 2(a) and 2(b), RMSE values of M2 profile are relatively higher than RMSE of other mild and

TABLE 3: Comparison of different methods for calculating distance of GVFs with steep slopes.

Method	RMSE	MARE	R^2	RAE	MAE	RSE
Direct method (1-step)	191.75	0.26	0.96	0.18	735.29	0.12
Direct method (3-step)	142.51	0.1	0.97	0.06	242.64	0.02
Direct method (5-step)	135.56	0.07	0.98	0.03	132.61	0.01
ANN (this study)	110.17	0.73	0.98	0.15	627.09	0.04
GP (this study)	54.05	0.27	1	0.08	321.89	0.01

steep profiles. Also, RMSE values decrease with the increase of number of intervals in the step direct method for all GVF profiles, while this reduction is significant for M1, M2, S1, and S3 profiles. Based on Figures 2(a) and 2(b), the lowest RMSE was achieved by GP for S1 and S3 profiles and by the 5-step direct method for M1, M2, and M3 profiles. On the other hand, RMSE values computed by all methods for S2 profiles are close to one another. MARE values depicted in Figures 2(c) and 2(d) indicate that the ANN and GP achieved close values to different versions of direct step methods, whereas they obtained higher MARE values for M3 and S2 profiles. Moreover, MARE decreases with the increase of number of subreach considered in the direct method. According to Figures 2(c) and 2(d), the ANN and GP yield to better MARE than the 1-step direct method for M1 and S3 profiles, while they perform better than the 5-step direct method for S1 profile based on MARE criterion. The comparison of R^2 values shown in Figures 2(e) and 2(f) implies that the accuracy of estimation of L increases considering more intervals in direct step methods. Moreover, the 1-step direct method did not achieve acceptable R^2 for M2 and S1 profiles. Based on Figures 2(e) and 2(f), the ANN and GP resulted in promising R^2 values for M1, M2, S1, and S3 profiles, while their R^2 for M3 profile is lower than that of direct step methods. Finally, MARE values computed by all methods for S2 profile are comparable and relatively low.

Figure 3 depicts percentages of estimated L in error ranges for different GVF profiles. This figure provides a suitable opportunity for detailed accurate comparison so that a swift glance reveals which method performs well for each specific GVF profile. According to Figure 3, the percentages may increase with the increase of error ranges; while the more the percentage in one specific error range, the more precise the results are. For M1 profile, Figure 3(a) shows that all methods reach closer results to the benchmark solutions in comparison with the 1-step direct method. Based on Figure 3(b), estimated L by the 1-step method and the ANN for M2 profile contains significant errors, while GP results are relatively better. In addition, 3-step and 5-step direct methods reach more accurate solutions than the rest for M2 profile. Figures 3(c) and 3(e) depict the poor performance of the ANN and GP in predicting L of M3 and S2 profiles, while it also manifests inadequate performance of the 1-step direct method in comparison with 3-step and 5-step direct methods for M3 profile. On the contrary, Figure 3(d) clearly demonstrates the improvement achieved by both the ANN and GP in computation of L in S1 profile compared to all direct step methods. Figure 2(e) demonstrates the impact on considering more subreaches in direct

step methods since it shows that the 5-step direct method has better results than the 3-step direct method and 3-step better than the 1-step direct method, as well. Moreover, the error ranges shown in Figure 2(e) imply that ANN estimations are much closer to the benchmark solutions in each error range considered. Furthermore, Figure 2(e) indicates that GP is capable of accurate prediction of L in S3 profile, even better than direct step methods. In summary, GP was found to compute lengths of M1, S1, and S3 profiles with high accuracy, while the ANN performs well in prediction of L in M1 and S1 profiles.

Confidence limits of the lengths of GVF profiles using all considered methods are shown in Figure 4. It clearly shows that considering more intervals in the direct step method makes the estimated confidence limit closer to that of the benchmark solution for each and every GVF profile. According to Figure 4, all methods exploited for estimation of L yielded to close confidence limits as that of the benchmark solutions for M1 and S1 profiles. However, the confidence limits achieved for M2 and S3 profiles are different, while GP and the 5-step direct method reach to the closest confidence limits to that of the benchmark solutions. Figure 4(c) manifests the poor performance of the ANN in estimating L of M3 profile, while GP replicates the confidence limit of the benchmark solution with high accuracy for this GVF profile. Moreover, the comparison of the confidence limits shown in Figure 4(e) reveals that all methods failed to achieve the whole range of the confidence limit of the benchmark solution, while all confidence limits are within the minimum and maximum points of the benchmark confidence limit. In summary, Figures 2–4 indicate that GP computed the lengths of M1, S1, and S3 profiles with high accuracy compared, while the ANN was successful in estimation of M1 and S1 profile lengths.

According to the comparative analysis conducted in this study, the obtained results show that AI models can improve the prediction of the distance between two water depths specified when they both belong to M1, S1, and S3 profiles. Therefore, one of the main advantages of the AI techniques is that they can improve the accuracy of estimating lengths of gradually varied flow profiles, particularly when the common numerical modeling produces errors, e.g., when the water depth approaches either the normal or critical depth. Additionally, the AI methods, when they are trained, can give estimations with a less computational effort than the benchmark solution and numerical schemes. The two latter methods required integration calculations and marching on space to compute a length of a gradually varied flow profile. These benefits may attract attention toward using the AI techniques in this specific application. In conclusion, the AI
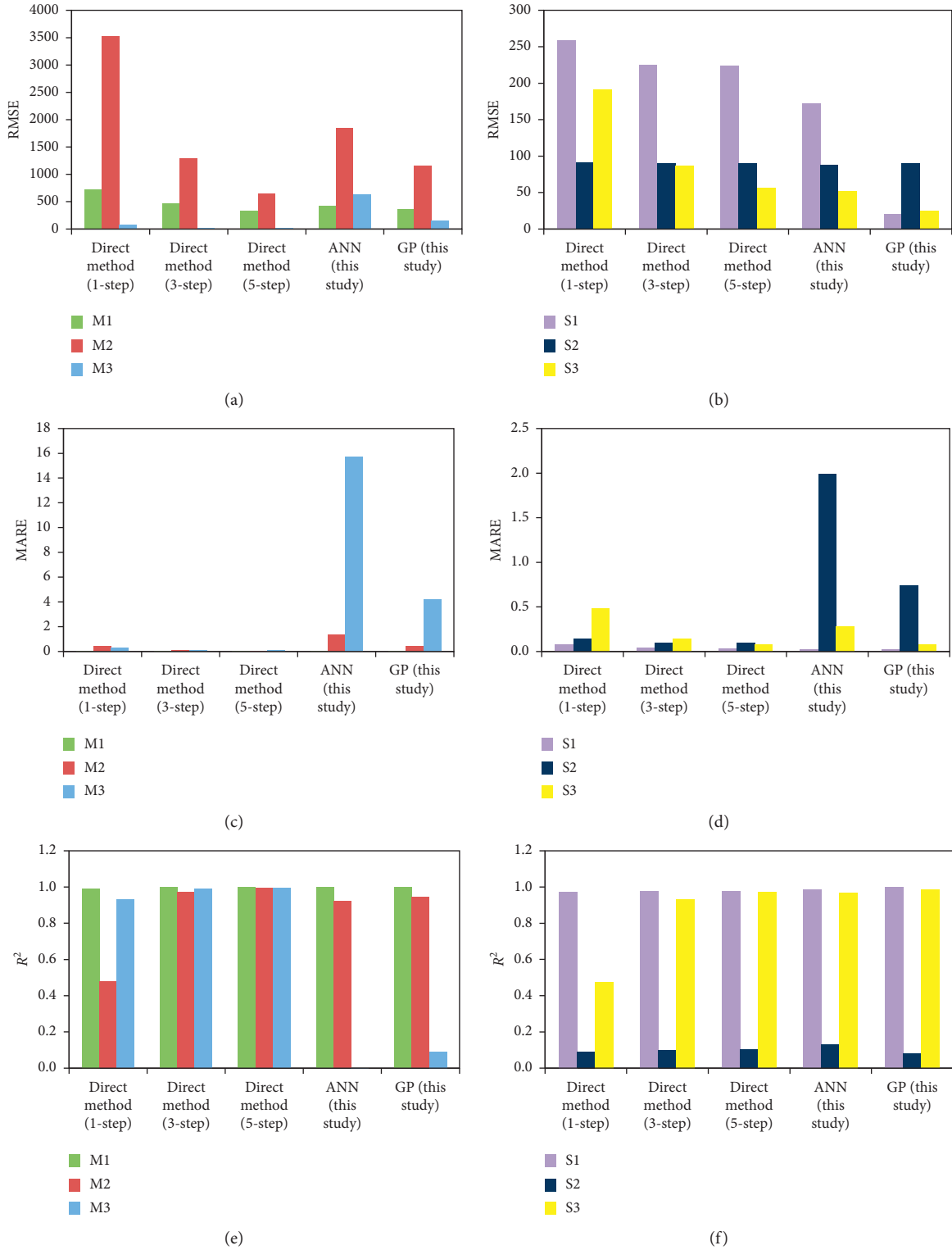
Figure 2: Performances of different methods for calculating distance of different GVF profiles: (a) RMSE for mild slopes, (b) RMSE for steep slopes, (c) MARE for mild slopes, (d) MARE for steep slopes, (e) R^2 for mild slopes, and (f) R^2 for steep slopes.
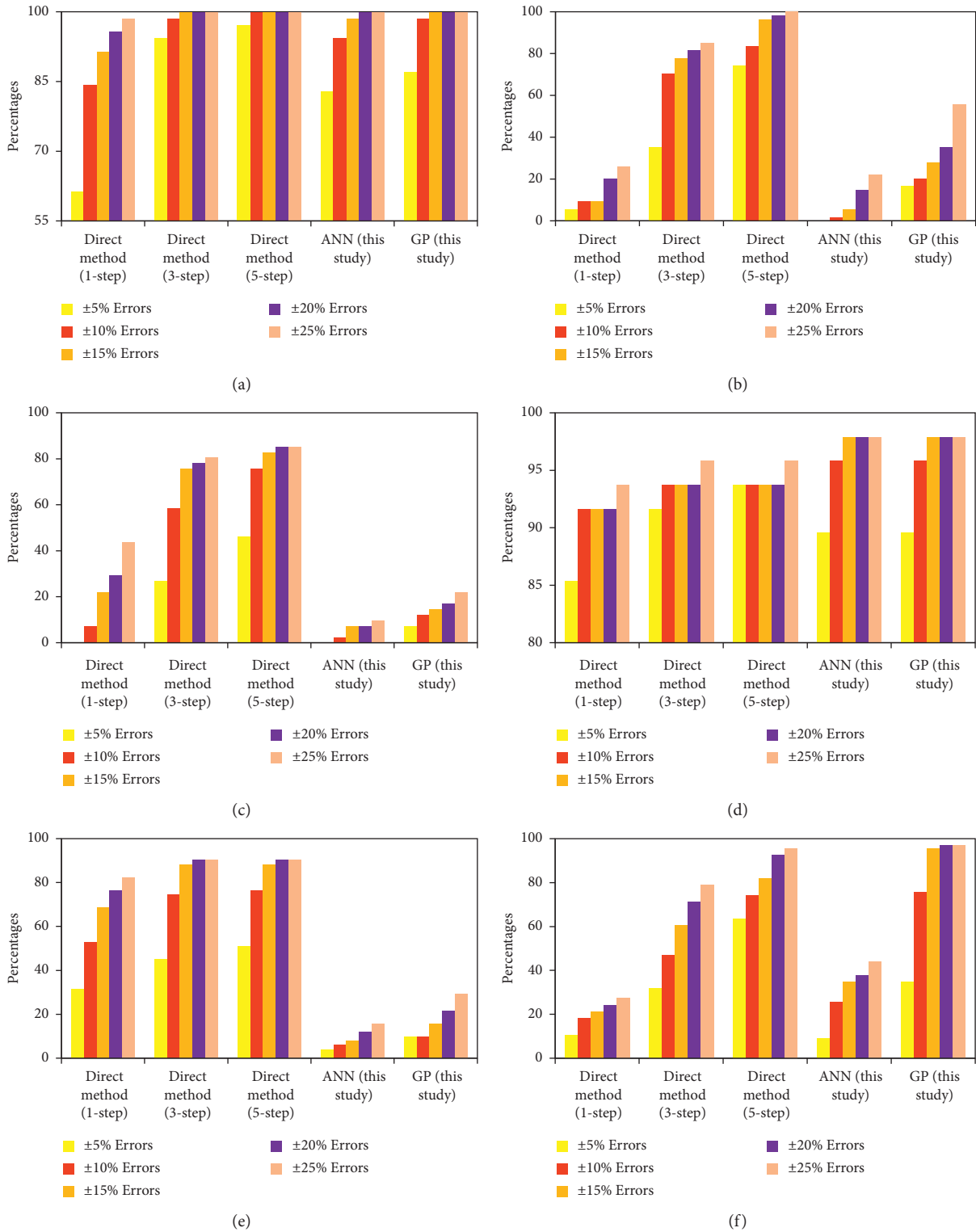
FIGURE 3: Percentages of estimated distance in error ranges for different GVF profiles: (a) M1, (b) M2, (c) M3, (d) S1, (e) S2, and (f) S3.
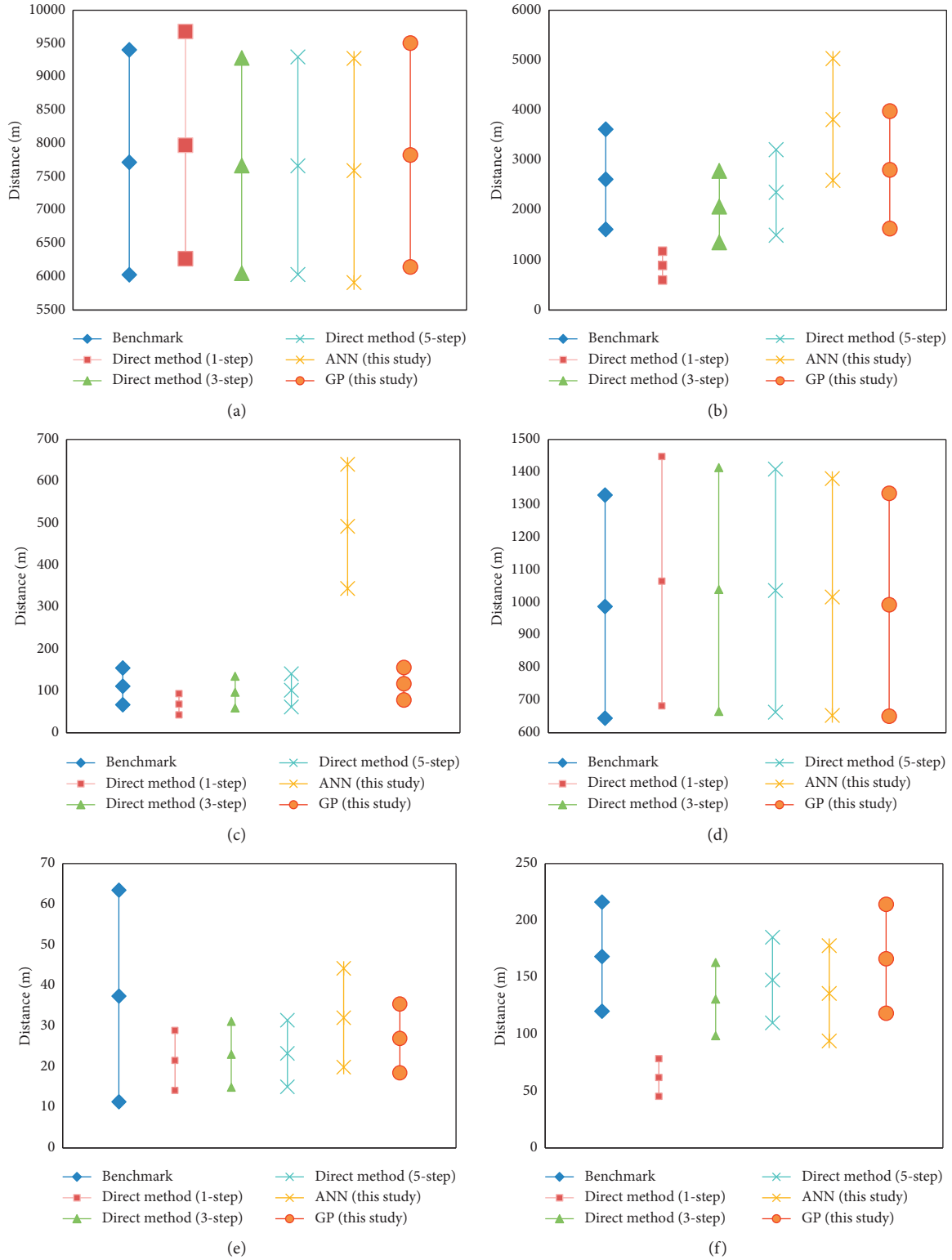
Figure 4: Confidence limits of estimated distances for different GVF profiles: (a) M1, (b) M2, (c) M3, (d) S1, (e) S2, and (f) S3.

models perform much better in many cases in comparison with other methods considered in this study.

4. Conclusions

Although varied flow function provides accurate estimation of the lengths between any pair of known water depths in steady GVF profiles, it requires four complicated integral forms to be calculated which makes it inadequate in practice in comparison to numerical methods such as the direct step method. In this regard, two AI models called the ANN and GP were trained with 5154 data that contained varied flow function. The performances of the ANN and GP in predicting the length between two specified water depths were compared with the direct method having one, three, and five steps, while the results of varied flow function were set as benchmark solutions. The test data consist of 165 data for mild slope and 165 data for steep slope. According to the results, the accuracy of the direct method increases considering more number of intervals considered by the compensation of increasing computational efforts. Also, the comparison clearly demonstrates that GP outperformed others for M1, S1, and S3 profiles. Furthermore, high accurate results were obtained by the ANN for M1 and S1 profiles. However, the results reveal better performance of 3-step and 5-step direct methods for M2, M3, and S2 profiles, while the 1-step direct method failed to estimate precise lengths for M2, M3, and S3 profiles. Finally, GP and the ANN are suggested for estimation of GVF lengths when water depth is larger than normal and critical depths in mild and steep slopes, respectively. The shortcomings and improvements of the AI models in estimating GVF lengths can be beneficial to water surface modeling in future studies.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] P. K. Swamee, "Length of gradually varied flow profiles," *ISH Journal of Hydraulic Engineering*, vol. 9, no. 2, pp. 70–79, 2003.

[2] A. R. Vatankhah and S. M. Easa, "Accurate gradually varied flow model for water surface profile in circular channels," *Ain Shams Engineering Journal*, vol. 4, no. 4, pp. 625–632, 2013.

[3] A. S. Ramamurthy, S. F. Saghravani, and R. Balachandar, "A direct integration method for computation of gradually varied flow profiles," *Canadian Journal of Civil Engineering*, vol. 27, no. 6, pp. 1300–1305, 2000.

[4] R. G. Patil, V. N. Diwanji, and R. M. Khatsuria, "Integrating equation of gradually varied flow," *Journal of Hydraulic Engineering*, vol. 127, no. 7, pp. 624-625, 2001.

[5] M. Venutelli, "Direct integration of the equation of gradually varied flow," *Journal of Irrigation and Drainage Engineering*, vol. 130, no. 1, pp. 88–91, 2004.

[6] C.-D. Jan and C.-l. Chen, "Use of the Gaussian hypergeometric function to solve the equation of gradually-varied flow," *Journal of Hydrology*, vol. 456-457, pp. 139–145, 2012.

[7] A. R. Vatankhah, "Analytical solution of gradually varied flow equation in circular channels using variable Manning coefficient," *Flow Measurement and Instrumentation*, vol. 43, pp. 53–58, 2015.

[8] R. Szymkiewicz and A. Szymkiewicz, "Method to solve the non-linear systems of equations for steady gradually varied flow in open channel network," *Communications in Numerical Methods in Engineering*, vol. 20, no. 4, pp. 299–312, 2004.

[9] B. Kaya, "Investigation of gradually varied flows using differential quadrature method," *Scientific Research and Essays*, vol. 6, no. 13, pp. 2630–2638, 2011.

[10] W. Artichowicz and R. Szymkiewicz, "Computational issues of solving the 1D steady gradually varied flow equation," *Journal of Hydrology and Hydromechanics*, vol. 62, no. 3, pp. 226–233, 2014.

[11] S. Ghazizadeh and M. J. Abedini, "Gradually varied flow modeling: how to choose a reach with maximum information content?," *Flow Measurement and Instrumentation*, vol. 56, pp. 56–60, 2017.

[12] C. Sivapragasam, P. Saravanan, K. Ganeshmoorthy, A. Muhil, S. Dilip, and S. Saivishnu, "Mathematical modeling of gradually varied flow with genetic programming: a lab-scale application," in *Information and Communication Technology for Intelligent Systems*, pp. 409–416, Springer, Berlin, Germany, 2019.

[13] B. W. Karney, A. Malekpour, and H. Salehi, "An exploratory approach to teaching gradually varied flow," *Journal of Hydro-Environment Research*, vol. 4, no. 2, pp. 175–180, 2010.

[14] H. Damangir and M. J. Abedini, "System identification and subsequent discharge estimation based on level data alone-gradually varied flow condition," *Flow Measurement and Instrumentation*, vol. 36, pp. 24–31, 2014.

[15] E. Gadissa and A. Teshome, "Estimation of open channel flow parameters by using genetic algorithm," *Open Journal of Optimization*, vol. 7, no. 3, p. 51, 2018.

[16] F. Motaman, G. R. Rakhshandehroo, M. R. Hashemi, and M. Niazkar, "Application of RBF-DQ method to time-dependent analysis of unsaturated seepage," *Transport in Porous Media*, vol. 125, no. 3, pp. 543–564, 2018.

[17] M. Niazkar and N. Talebbeydokhti, "Comparison of explicit relations for calculating Colebrook friction factor in pipe network analysis using H-based methods," *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, vol. 44, no. 1, p. 231, 2020.

[18] M. Najafzadeh, F. Saberi-Movahed, and S. Sarkamaryan, "NF-GMDH-based self-organized systems to predict bridge pier scour depth under debris flow effects," *Marine Georesources & Geotechnology*, vol. 36, no. 5, pp. 589–602, 2018.

[19] M. Najafzadeh, "Evaluation of conjugate depths of hydraulic jump in circular pipes using evolutionary computing," *Soft Computing*, vol. 23, no. 24, pp. 13375–13391, 2019.

[20] M. Najafzadeh and G. Oliveto, "Riprap incipient motion for overtopping flows with machine learning models," *Journal of Hydroinformatics*, vol. 22, no. 4, pp. 749–767, 2020.

[21] M. Niazkar and M. Zakwan, "Assessment of artificial intelligence models for developing single-value and loop rating curves," *Complexity*, vol. 2021, Article ID 6627011, 21 pages, 2021.

[22] M. Niazkar and S. H. Afzali, "Optimum design of lined channel sections," *Water Resources Management*, vol. 29, no. 6, pp. 1921–1932, 2015.

[23] M. Niazkar, "Revisiting the estimation of colebrook friction factor: a comparison between artificial intelligence models and C-W based explicit equations," *KSCE Journal of Civil Engineering*, vol. 23, no. 10, pp. 4311–4326, 2019.

[24] M. Niazkar, G. R. Rakhshandehroo, and S. H. Afzali, "Deriving explicit equations for optimum design of a circular channel incorporating a variable roughness," *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, vol. 42, no. 2, pp. 133–142, 2018.

[25] M. Niazkar, "Assessment of artificial intelligence models for calculating optimum properties of lined channels," *Journal of Hydroinformatics*, vol. 22, no. 5, p. 1410, 2020.

[26] M. Niazkar, N. Talebbeydokhti, and S. H. Afzali, "Development of a new flow-dependent scheme for calculating grain and form roughness coefficients," *KSCE Journal of Civil Engineering*, vol. 23, no. 5, pp. 2108–2116, 2019a.

[27] A. R. Vatankhah, "Exact sensitivity equation for one-dimensional steady-state shallow water flow (application to model calibration)," *Journal of Hydrologic Engineering*, vol. 15, no. 11, pp. 939–945, 2010.

[28] L. Arneson, L. Zevenbergen, P. Lagasse, and P. Clopper, "Evaluating scour at bridges (HEC-18)," Technical Rep No HIF-12 3, Federal Highway Administration, Washington, DC, USA, 2012.

[29] M. Niazkar, N. Talebbeydokhti, and S. H. Afzali, "One dimensional hydraulic flow routing incorporating a variable grain roughness coefficient," *Water Resources Management*, vol. 33, no. 13, pp. 4599–4620, 2019c.

[30] M. Niazkar, N. Talebbeydokhti, and S. H. Afzali, "Novel grain and form roughness estimator scheme incorporating artificial intelligence models," *Water Resources Management*, vol. 33, no. 2, pp. 757–773, 2019b.

[31] V. T. Chow, *Open Channel Hydraulics*, McGraw-Hill, New York, NY, USA, l959.

[32] K. Subramanya, *Flow in Open Channels*, Tata McGraw-Hill Education, New York, NY, USA, Third edition, 1982.

[33] C.-D. Jan, *Gradually-varied Flow Profiles in Open Channels*, Springer, Berlin, Germany, 2014.

[34] M. Niazkar and S. H. Afzali, "Application of new hybrid method in developing a new semicircular-weir discharge model," *Alexandria Engineering Journal*, vol. 57, no. 3, pp. 1741–1747, 2018.

[35] M. Niazkar, N. Talebbeydokhti, and S.-H. Afzali, "Bridge backwater estimation: a comparison between artificial intelligence models and explicit equations," *Scientia Iranica*, 2020.

[36] H. R. Niazkar and M. Niazkar, "COVID-19 outbreak: application of multi-gene genetic programming to country-based prediction models," *Electronic Journal of General Medicine*, vol. 17, no. 5, p. em247, 2020.

[37] F. D. Francone, *Discipulus Owner's Manual*, Machine Learning Technologies, Inc, Littleton, CO, USA, 1998.

[38] F. M. Henderson, *Open Channel Flow*, Macmillan, New York, NY, USA, 1966.

[39] M. Niazkar and S. H. Afzali, "Streamline performance of excel in stepwise implementation of numerical solutions," *Computer Applications in Engineering Education*, vol. 24, no. 4, pp. 555–566, 2016.

[40] M. Niazkar and S. H. Afzali, "New nonlinear variable-parameter muskingum models," *KSCE Journal of Civil Engineering*, vol. 21, no. 7, pp. 2958–2967, 2017.

[41] M. Najafzadeh and S. Sarkamaryan, "Extraction of optimal equations for evaluation of pipeline scour depth due to currents," in *Proceedings of the Institution of Civil Engineers-Maritime Engineering*, vol. 171, no. 1, pp. 1–10, 2018.

[42] S. Keesstra, J. P. Nunes, P. Saco et al., "The way forward: can connectivity be useful to design better measuring and modelling schemes for water and sediment dynamics?" *Science of the Total Environment*, vol. 644, pp. 1557–1572, 2018.

[43] S. D. Keesstra, J. Davis, R. H. Masselink, J. Casalí, E. T. H. M. Peeters, and R. Dijksma, "Coupling hysteresis analysis with sediment and hydrological connectivity in three agricultural catchments in Navarre, Spain," *Journal of Soils and Sediments*, vol. 19, no. 3, pp. 1598–1612, 2019.