



5th International Conference of the International
Commission of Agricultural and Biosystems Engineering
(CIGR)

Hosted by the Canadian Society for Bioengineering
(CSBE/SCGAB) Virtually from Québec City, Canada – May
11-14, 2021



ÉVALUATION DE LA PERFORMANCE DE LA PROGRAMMATION DE L'EXPRESSION GÉNIQUE DANS LA PRÉVISION DU COEFFICIENT DE DÉCHARGE D'UN ORIFICE CIRCULAIRE LATÉRAL

LAURA DUMONT-BOYER¹, ISA EBTEHAJ², HOSSEIN BONAKDARI^{3*}

¹ Département des sols et de génie agroalimentaire, Université Laval, Québec, G1V 0A6, Canada.
<laura.dumont-boyer.1@ulaval.ca>

² Département des sols et de génie agroalimentaire, Université Laval, Québec, G1V 0A6, Canada
<isa.ebtehaj.1@ulaval.ca>

³ Département des sols et de génie agroalimentaire, Université Laval, Québec, G1V 0A6, Canada
<hossein.bonakdari@fsaa.ulaval.ca>

ABSTRACT L'utilisation de l'irrigation pour une agriculture durable passe par le contrôle et la quantification des livraisons d'eau afin de rendre le système sensible aux besoins de la ferme. Cet objectif est atteint en utilisant des dispositifs pratiques de mesure de l'eau pour réguler la surface d'écoulement dans les canaux d'irrigation et les réseaux de drainage. Les orifices latéraux sont largement utilisés comme structures de dérivation dans les réseaux de drainage et d'irrigation et d'autres structures hydrauliques. En raison de l'importance de ce type de structure de dérivation, plusieurs études théoriques et expérimentales ont été menées par différents chercheurs sur le comportement d'écoulement et le rejet d'eau passant par les orifices latéraux. Le calcul souple a été largement utilisé dans la modélisation de systèmes non linéaires, la reconnaissance de formes et la résolution de problèmes complexes dans différentes sciences au cours des dernières années. Compte tenu de la capacité du calcul souple à résoudre des problèmes complexes, cette méthode est souvent appliquée dans différents domaines pratiques. L'objectif principal de cette étude est d'introduire une nouvelle équation pour prédire le coefficient de décharge d'un orifice latéral circulaire à crête pointue situé sur le côté principal d'un canal rectangulaire droit. Une nouvelle équation simple est présentée pour calculer le coefficient de décharge à l'aide des paramètres de validité et de l'intelligence artificielle. L'exactitude de l'équation présentée sera en outre comparée à l'exactitude des équations existantes. De plus, l'effet de chaque paramètre sans dimension utilisé dans

l'équation proposée est étudié en présentant cinq modèles différents. La tendance du coefficient de décharge par rapport à chaque variable indépendante dans l'équation proposée est également étudiée par analyse partielle de sensibilité dérivée. Le modèle simple proposé dans cette étude peut servir de méthode alternative en ingénierie appliquée

Keywords: Irrigation, détournement de l'écoulement, coefficient de décharge, calcul souple, analyse de sensibilité

INTRODUCTION

Ramamurthy et coll. (1986) ont mené une étude expérimentale sur l'écoulement passant par un orifice latéral rectangulaire dans un chenal ouvert. Ils ont obtenu une équation pour déterminer la décharge à travers un orifice latéral rectangulaire en supposant une vitesse constante dans le déversoir et une vitesse moyenne dans le chenal principal. En outre, selon leurs résultats, la décharge passant par un orifice latéral est fonction de la longueur de l'orifice, de la largeur du canal principal et du rapport entre la vitesse moyenne dans le canal principal et la vitesse du jet de sortie du déversoir. Gill (1987) a étudié l'écoulement dans les orifices latéraux rectangulaires relativement courts des conduites d'eaux usées comme étant un type particulier d'écoulement spatialement varié. Swamee et coll. (1993) ont obtenu une équation pour calculer le rejet passant par un orifice latéral à l'aide du concept élémentaire du coefficient de décharge. Ils ont présenté le coefficient de décharge en fonction de la profondeur d'écoulement à l'intérieur du canal principal et de la hauteur de crête de l'orifice latéral. Ojha et Subbaiah (1997) ont utilisé le concept du coefficient élémentaire de décharge pour suggérer une équation servant à calculer le coefficient de décharge d'un orifice latéral situé sur le côté latéral d'un chenal ouvert en fonction de la géométrie et de la hauteur de crête de l'orifice latéral. Bryant et coll. (2008) ont étudié un modèle et les détails de l'écoulement en amont d'un orifice latéral à l'aide de la théorie du potentiel. Hussain et coll. (2010) ont mené une étude expérimentale sur les spécifications de l'écoulement passant par un orifice latéral circulaire situé sur le côté d'un chenal principal rectangulaire. Ils ont utilisé une analyse dimensionnelle pour calculer le coefficient de décharge d'une équation d'orifice circulaire en fonction du nombre de Froude et du rapport entre le diamètre de l'orifice et la largeur du canal principal. La relation de décharge classique pour un petit débit d'un orifice circulaire avec distribution constante de pression est:

$$Q = C_d \sqrt{2gH} \cdot \frac{\pi}{4} D^2 \quad (1)$$

où Q est la décharge dans l'orifice, C_d est le coefficient de décharge, g est l'accélération gravitationnelle, H est la hauteur d'eau au-dessus de l'axe central de l'orifice et D est le diamètre de l'orifice. Pour calculer le coefficient de décharge à l'orifice, Swamee et Swamee (2010) ont exprimé analytiquement l'équation suivante :

$$C_d = 0.611 \left[87 \left(\frac{v}{D\sqrt{gH}} \right)^{1.43} + \left(1 + \frac{4.5v}{D\sqrt{gH}} \right)^{-1.26} \right]^{-0.7} \quad (2)$$

Le calcul souple a été largement utilisée dans la modélisation de systèmes non linéaires, la reconnaissance de tendances et la résolution de problèmes complexes dans différentes sciences au cours des dernières années. Compte tenu de la capacité du calcul souple à résoudre des problèmes complexes, il a souvent été appliquée dans différents domaines, comme l'hydrologie (Nayak et coll., 2004; Kisi, 2005), le génie fluvial (Agil et coll., 2007), le génie hydraulique (Bonakdari et coll., 2011; Baghalian et coll., 2012; Ebtehaj et coll., 2014) et le transport des sédiments (Aytek et Kisi, 2008; Kisi et Guven, 2010; Ebtehaj et Bonakdari, 2013, 2014a, 2014b). Khorchani et Blanpain (2005) qui ont utilisé des réseaux de neurones artificiels et présenté un modèle pour calculer le coefficient de décharge d'un déversoir latéral situé sur le côté d'un canal rectangulaire. Emiroglu et coll. (2010) ont utilisé le système d'inférence floue basé sur un réseau adaptatif (ANFIS) qui a prédit le coefficient de décharge d'un orifice latéral triangulaire à labyrinthe situé sur un canal rectangulaire droit en régime sous-critique. Ils ont calculé l'équation du coefficient de décharge de l'orifice latéral triangulaire du labyrinthe en fonction de la longueur du déversoir latéral jusqu'à la largeur du chenal principal, de la longueur latérale du déversoir jusqu'à la longueur latérale de la crête du déversoir et de la hauteur du déversoir latéral jusqu'à la profondeur d'écoulement en amont de l'orifice; le déversoir latéral de labyrinthe a inclus l'angle et le nombre de Froude en amont du déversoir latéral. Emiroglu et coll. (2011) ont prédit la capacité de décharge hydraulique d'un déversoir latéral triangulaire à labyrinthe à l'aide de réseaux de neurones artificiels (ANN). Ils ont comparé les résultats artificiels du réseau neuronal avec la régression non linéaire et les résultats de régression linéaire multiple. Kisi et coll. (2012) ont prédit les détails de l'écoulement de dérivation passant par un orifice latéral triangulaire à l'aide de réseau de fonctions de base radiale (RBNN), de réseaux neuronaux de régression généralisée (GRNN), d'algorithmes de programmation d'expression génétique (GEP) et d'analyses de régression linéaire et non linéaire.

L'un des plus récents modèles présentés sur le sujet du calcul souple est la programmation d'expression génétique (GEP). Azamathulla et Ab. Ghani (2010) ont prédit

la profondeur d'affouillement des pipelines à l'aide de ce modèle et ont conclu que par rapport aux modèles existants, le modèle proposé donne de meilleurs résultats. Ab. Ghani et Azamathulla (2010) ont développé des relations fonctionnelles en utilisant la programmation d'expression génétique pour étudier le transport des sédiments dans les réseaux de conduites d'égout. Khan et coll. (2012) ont utilisé la programmation d'expression génétique pour prédire l'affouillement de la jetée du pont. Chang et coll. (2012) ont utilisé des techniques de calcul souple, y compris la programmation d'expression génétique, les réseaux de neurones avancés (FFNN) et l'ANFIS pour prédire la charge totale en matériaux de lit de trois rivières différentes en Malaisie. Leurs résultats ont indiqué que le niveau d'exactitude des techniques de calcul souple est acceptable. De plus, ils ont conclu que, comparativement à d'autres modèles, la programmation d'expression génétique est plus précise. En utilisant la programmation de l'expression génétique, Baylar et coll. (2014) ont mis au point deux modèles pour estimer l'efficacité de l'aération et l'entraînement de l'air des déversoirs. Pour surmonter les faiblesses des équations basées sur la régression, Ebtehaj et coll. (2015a) ont développé une nouvelle équation pour prédire le coefficient de décharge des déversoirs latéraux rectangulaires à l'aide de la programmation d'expression génétique, qui est très précise dans différentes conditions hydrauliques.

L'objectif principal de cette étude est d'introduire une nouvelle équation pour prédire le coefficient de décharge d'un orifice latéral circulaire à crête pointue situé sur le côté principal d'un canal rectangulaire droit. La nouvelle équation vise à surmonter les faiblesses des équations basées sur la régression. Par conséquent, l'analyse dimensionnelle est principalement appliquée pour déterminer les facteurs influençant l'estimation du coefficient de décharge. Une nouvelle équation simple est présentée pour calculer le coefficient de décharge à l'aide des paramètres efficaces et de la technique de la programmation d'expression génétique (GEP). L'exactitude de l'équation présentée sera en outre comparée à l'exactitude des équations existantes. De plus, l'effet de chaque paramètre sans dimension utilisé dans l'équation proposée est étudié en présentant cinq modèles différents. La tendance du coefficient de décharge par rapport à chaque variable indépendante dans l'équation proposée est également étudiée par analyse partielle de sensibilité dérivée.

LA PROGRAMMATION DE L'EXPRESSION DES GÈNES

La programmation génétique (GP) est parmi les modèles informatiques souples qui impliquent des principes darwiniens de sélection naturelle pour résoudre différents problèmes. Koza (1992) a présenté GP en développant l'algorithme génétique (GA). La principale différence entre GP et GA est la façon dont les résultats sont présentés. GA

utilise une chaîne de nombres pour présenter les résultats, tandis que GP présente les relations entre les paramètres en tant que structures arborescentes.

GEP est une variante linéaire de la programmation génétique introduite par Ferreira (2001). GEP contient des chromosomes linéaires de longueur fixe hérités d'algorithmes génétiques et d'arbres analytiques provenant de la programmation génétique de formes et de tailles variées (Ferreira, 2001). Bien que le phénotype et le génotype aient été combinés dans un système simple de répliquats, GEP est un système avancé de phénotype/génotype dans lequel le génotype et le phénotype sont complètement séparés. Par conséquent, le système évolué de phénotype/génotype GEP est supérieur de 100 à 60000 fois comparé au système de GP (Ferreira, 2001). GEP comprend cinq composantes principales : une fonction de conditionnement physique, un ensemble de fonctions, des paramètres de contrôle, un ensemble de bornes et une condition de terminaison. Contrairement aux arbres d'analyse dans la programmation génétique, GEP utilise une chaîne de détails avec une longueur constante pour présenter les résultats comme des arbres d'analyse de différentes tailles et formes. Un des avantages de la méthode GEP est la variété génétique extrêmement simple qui fonctionne à différents niveaux chromosomiques en tant qu'opérateur génétique. Une autre spécification unique de la GEP est sa nature multigénétique, qui présente un programme complexe en utilisant un certain nombre de sous-programmes différents (Alavi et al., 2011).

Dans le processus de programmation génétique, les fonctions essentielles pour créer le modèle et l'ensemble terminal sont d'abord sélectionnées. À l'étape suivante, l'ensemble de données existant est appelé pour prédire les paramètres prévus et les comparer avec les valeurs réelles. Ensuite, les chromosomes sont générés au hasard pour présenter la population initiale. À l'étape suivante, le programme est exécuté pour la population comprenant les chromosomes existants et l'aptitude de la fonction cible est vérifiée. Le programme s'arrête lorsque la condition d'arrêt est remplie. Sinon, la fonction cible est évaluée à l'aide d'une nouvelle population et des nouveaux chromosomes qui ont été corrigés par l'opérateur génétique. Ce processus se poursuit jusqu'à ce que la condition d'arrêt soit remplie. La figure 1 montre un schéma du processus de modélisation GEP.

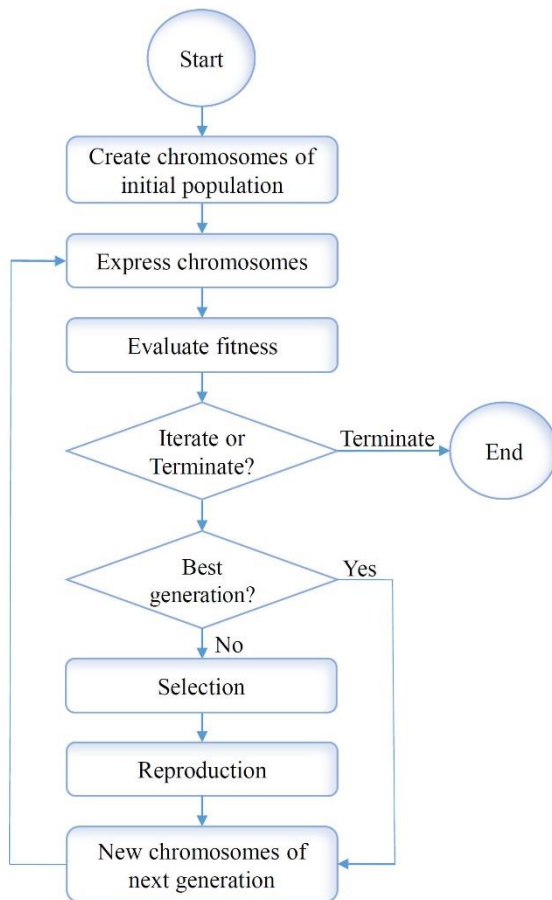


Figure 1. Schéma du processus de modélisation GEP

COLLECTE DES DONNÉES

Dans cette étude, les résultats expérimentaux obtenus par Hussain et coll. (2010) ont été utilisés pour la prédiction du coefficient de décharge. Hussain et coll. (2010) ont mené des expériences dans un chenal principal rectangulaire de 9,15 m de longueur, 0,5 m de largeur et 0,6 m de profondeur avec un orifice latéral situé à 5,18 m en amont du chenal principal. Ils ont utilisé une barrière d'écluse à l'extrémité du canal principal pour régler la profondeur d'écoulement. Le déversement passant par le déversoir latéral a traversé un canal de dérivation de 3.8m de longueur, 0,26 m de largeur et 0,41 m de profondeur, puis est retourné au chenal principal. Dans les expériences de Hussain et coll. (2010), les orifices avaient un diamètre de 5, 10 et 15 cm et avaient des hauteurs de crête de 0,05; 0,10; 0,15; 0,20 et 0,25 m. Différents débits de décharge (Q_m) ont été utilisés dans le canal principal pour chaque diamètre d'orifice latéral et hauteur de crête réglée, et les différentes profondeurs d'écoulement ont été maintenues pour chaque décharge dans le

canal principal en les réglant avec la porte d'écluse. Leurs expériences ont été exécutées avec un écoulement libre passant par un orifice sans formation de vortex près de l'orifice. L'éventail des données appliquées dans cette étude est présenté au Tableau 1.

Tableau 1. Gamme des données collectées mesurées par Hussain et al. (2010).

	B/D	W/D	Y _m /D	Y _m /W	Re	Fr
Max	10.00	3.00	11.78	10.08	1.12E+05	0.52
Min	3.33	0.33	1.74	1.50	3.96E+03	0.04

ANALYSE ADIMENSIONNELLE

Selon les études sur les structures de dérivation (déversoirs et orifices latéraux), les paramètres qui influencent la détermination du coefficient de décharge sont la vitesse dans le canal principal (V), la hauteur de crête (W), la profondeur d'écoulement dans le canal principal (Y_m), la largeur du canal principal (B), la densité massique (ρ), l'accélération due à la gravité (g) et la viscosité (μ). Par conséquent, les facteurs qui influencent la prédiction du coefficient de décharge sont énoncés dans l'équation suivante :

$$C_d = f_1(B, W, V, D, Y_m, \rho, \mu, g) \quad (3)$$

En considérant que le paramètre du coefficient de décharge est adimensionnel, l'équation ci-dessus peut être réécrite comme suit pour rendre les paramètres adimensionnels:

$$C_d = f_2\left(\frac{B}{D}, \frac{Y_m}{D}, \frac{W}{D}, \frac{\rho V D}{\mu}, \frac{V}{\sqrt{g Y_m}}\right) \quad (4)$$

où le paramètre $V / \sqrt{g Y_m}$ indique le nombre Froude (Fr) et $\rho V D / \mu$ représente le nombre de Reynolds. Étant tenu compte du fait que l'effet du nombre de Reynolds n'est pas significatif dans les canaux ouverts, ce paramètre peut être supprimé de l'équation ci-dessus et l'équation réécrite comme suit (Hussain et al., 2010) :

$$C_d = f_3\left(\frac{B}{D}, \frac{Y_m}{D}, \frac{W}{D}, Fr\right) \quad (5)$$

DÉRIVATION DU COEFFICIENT DE DÉCHARGE BASÉ SUR GEP

L'équation présentée dans cette section implique la méthode GEP pour calculer le coefficient de décharge dans un orifice latéral circulaire. Pour présenter le modèle, 80 % de toutes les données (172 données) sont sélectionnées parmi les 215 données existantes comme ensemble de formation et les 20 % de données restantes (43 données) constituent l'ensemble d'essais. On peut prétendre que l'environnement d'entraînement du système est défini après avoir sélectionné l'ensemble de formation. Après avoir catégorisé les données, différents paramètres doivent être définis à des fins de modélisation. Les individus des chromosomes multigéniques, qui comprennent trois gènes, sont utilisés pour créer la population initiale. La taille initiale de la population doit maintenant être déterminée. Compte tenu de la suggestion de Ferreira (2011) d'utiliser une taille de population dans la fourchette de 30-100 pour produire de bons résultats, la population utilisée dans cette étude comprend 50 chromosomes trouvés par essais et erreurs. Les individus sont évalués après avoir sélectionné la taille de la population et leur fonction de conditionnement physique est calculée à l'aide de MSE comme suit:

$$f_i = \frac{1000}{1 + MSE_i} \quad (6)$$

où O_{ij} est la valeur observée pour le cas de forme physique j , P_{ij} est la quantité prévue utilisant le chromosome individuel i pour le cas j de forme physique, et MSE_i est l'erreur carrée moyenne (MSE) qui est calculée comme suit :

$$MSE = \frac{1}{n} \sum_{i=1}^n (O_i - P_i) \quad (7)$$

Le meilleur état est lorsque $E_{ij}=0$ est obtenu. Cela signifie que la quantité prévue à l'aide du chromosome i individuel pour le cas de forme physique j est égale à la quantité observée pour le cas de forme physique j ($P_{ij}=O_{ij}$). L'ensemble des bornes et l'ensemble des fonctions doivent être déterminés pour chaque gène du chromosome après avoir sélectionné la fonction de conditionnement physique. Les fonctions de conditionnement physique utilisées dans cette étude incluent $\{+, -, \times, /, X2, 4Rt, 5Rt, Pow\}$ alors que l'ensemble des bornes est comme suit :

$$T = \left\{ C_d, \frac{B}{D}, \frac{Y_m}{D}, \frac{W}{D}, Fr \right\} \quad (8)$$

Par la suite, le nombre de gènes et les longueurs de la tête et de la queue sont déterminés pour chaque gène du chromosome. Les quatre gènes de chaque chromosome utilisé dans CIGR 5th International Conference – Virtually from Québec City, Canada – May 11-14, 2021 8

cette étude ont été sélectionnés par essais et erreurs. La longueur de tête sélectionnée est de 5 ($h=5$), le nombre maximal d'arguments par fonction est de 2 ($n_{max}=2$) et la longueur de queue obtenue est de 6 ($t=5 \times (2-1)+1$). Maintenant, le taux de différents opérateurs génétiques est déterminé. Des opérateurs génétiques, tels que l'inversion de mutation, le transport (IS, RIS, transport de gènes), la recombinaison ou le croisement (un point, deux points, recombinaison de gènes) sont employés. Les taux des paramètres mentionnés sont présentés au Tableau 2.

Tableau 2. Paramètres utilisés dans la modélisation GEP.

Parameter	Setting
Number of generations	400000
Number of chromosomes	50
Number of genes	1
function set	+, -, ×, /, X2, 4Rt, 5Rt, Pow
Mutation rate	0.005
Inversion rate	0.15
IS transposition rate	0.15
RIS transposition rate	0.15
Gene transposition rate	0.15
One-point recombination rate	0.35
Two-point recombination rate	0.35
Gene recombination rate	0.15

Finalement, la fonction de connexion doit être déterminée. En tenant compte du fait qu'il y a 4 gènes avec 3 sous-expressions différentes dans cette étude, ces gènes doivent être connectés pour actualiser la réponse finale. Par conséquent, l'opérateur {+} est utilisé

dans la présente étude pour la fonction de connexion entre les gènes. La simulation de modèle suit la détermination des paramètres essentiels. Le critère de terminaison est la fonction de conditionnement physique maximale, qui est, en soi, une fonction de *MSE*. Après avoir modélisé et atteint la condition d'arrêt, l'équation (9) est appliquée pour prédire le coefficient de décharge en utilisant les paramètres introduits dans l'équation (3). La figure 2 indique l'arbre d'expression du modèle présenté.

$$C_d = 0.65 \times \left(\left(\frac{B}{D} \right)^{0.2} + \left(\frac{W}{D} \right)^{0.013} - \left(\frac{Y_m}{D} \right)^{0.026} - Fr^{0.25} \right)^{0.1} \quad (9)$$

APPLICATION ET RÉSULTATS

Afin de vérifier la faisabilité du modèle proposé, différents indices statistiques sont appliqués. Les indices statistiques R-squared (R^2), mean absolute percentage error (*MAPE*), root mean square error (*RMSE*), scatter index (*SI*) et *BIAS* sont définis ci-dessous et sont utilisés pour étudier quantitativement le coefficient de décharge dans cette étude.

$$R^2 = \left[\frac{\sum_{i=1}^n (C_{d_EXPi} - \overline{C_{d_EXPi}})(C_{d_GEPi} - \overline{C_{d_GEPi}})}{\sqrt{\sum_{i=1}^n (C_{d_EXPi} - \overline{C_{d_EXPi}})^2 \sum_{i=1}^n (C_{d_GEPi} - \overline{C_{d_GEPi}})^2}} \right]^2 \quad (10)$$

$$MAPE = \left(\frac{100}{n} \right) \sum_{i=1}^n \left(\frac{|C_{d_EXPi} - C_{d_GEPi}|}{C_{d_EXPi}} \right) \quad (11)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (C_{d_EXPi} - C_{d_GEPi})^2} \quad (12)$$

$$BIAS = \frac{\sum_{i=1}^n (C_{d_EXPi} - C_{d_GEPi})}{n} \quad (13)$$

$$SI = \frac{RMSE}{\frac{1}{n} \sum_{i=1}^n C_{d_EXPi}} \quad (14)$$

La figure 3 indique les résultats obtenus pour la prédiction du coefficient de décharge à l'aide de GEP dans l'entraînement du modèle lorsque C_d est un paramètre dépendant du rapport entre la largeur du canal principal et le diamètre de l'orifice (B/D), le rapport de la hauteur de crête au diamètre de l'orifice (W/D), le rapport de la profondeur d'écoulement dans le canal principal au diamètre de l'orifice (Y_m/D) et le nombre de Froude (Fr) (paramètres indépendants). La figure montre que les coefficients de décharge prévus sont assez précis, car la majorité des quantités prévues sont conformes aux quantités observées. Les indices statistiques mentionnés ci-dessus sont également utilisés pour mener une enquête quantitative sur la prédiction du coefficient de décharge à l'aide de GEP. GEP a prédit le coefficient de décharge avec $R^2=0,98$ et $MAPE=0,24$ qui indiquent la grande précision de GEP en prédisant le coefficient de décharge dans les orifices latéraux circulaires. Les autres indices statistiques utilisés pour le modèle présenté dans cette étude sont $RMSE=0,0019$, $BIAS=0,00004$ et $SI=0,003$. Comme ces montants sont proches de zéro, ils signifient l'exactitude du modèle proposé à l'aide de GEP. La figure 4 montre les valeurs d'erreur relatives des coefficients de décharge prévus à l'aide de GEP et différents nombres de Froude pour trois rapports différents largeur de canal principal/diamètre d'orifice (B/D). La figure indique que l'erreur relative maximale dans la prédiction du coefficient de décharge est inférieure à 1,2 % lorsque $B/D=10$. Il est également évident que lorsque $B/D<10$, les erreurs relatives des coefficients de décharge prévus n'atteignent pas 1%, ce qui est une indication de la grande précision de prédiction GEP.

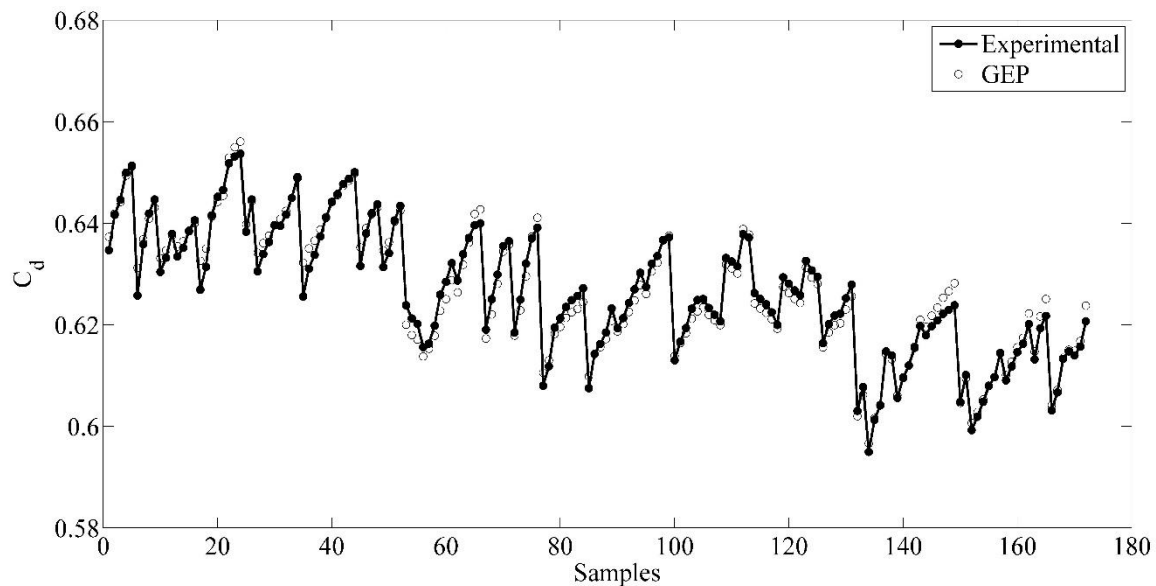


Figure 3. Proposition du modèle de vérification à l'aide des résultats expérimentaux de Hussain et coll. (2010).

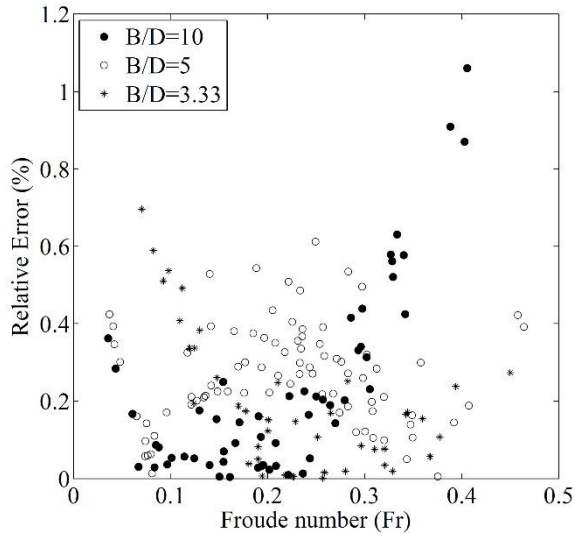


Figure 4. Erreur relative du coefficient de décharge prédit par GEP pour différents nombres de Froude et B/D.

La figure 5 indique les valeurs du coefficient de décharge prédites à l'aide de l'équation proposée (Eq. 9) et d'un modèle existant (Swamee et Swamee, 2010) par rapport aux résultats expérimentaux observés. Les quantités expérimentales observées font partie de 20 % des données de ce chiffre qui n'ont joué aucun rôle dans la prédiction avec l'équation (Eq. 9). Selon la figure, les quantités prévues à l'aide de GEP sont bien conformes aux résultats observés, car l'erreur relative du coefficient de décharge prédit à l'aide de GEP est d'environ 1 %. Contrairement au modèle présenté dans cette étude, le modèle de Swamee et Swamee (2010) n'est pas très précis pour prédire le coefficient de rejet. Comme on l'a observé, les coefficients de décharge prévus avec ce modèle sont inférieurs aux quantités réelles (sous-estimation) pour les coefficients de décharge inférieurs à 0,62. À mesure que le coefficient de rejet augmente, les valeurs prédites à l'aide du modèle de Swamee et Swamee (2010) sont surestimées. De plus, à mesure que le coefficient de décharge augmente, l'erreur de prédiction à l'aide de cette équation augmente également.

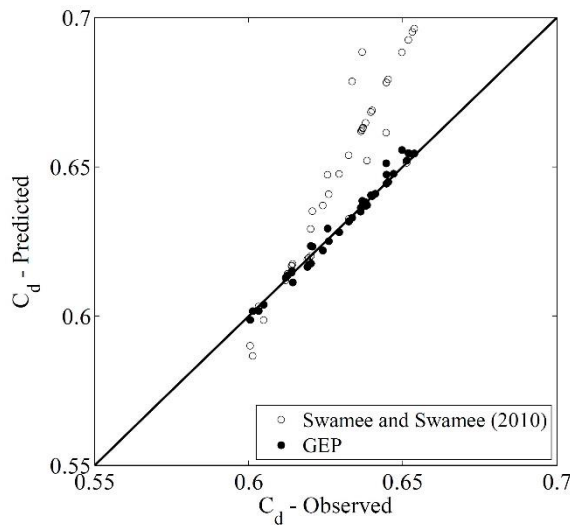


Figure 5. Proposition du modèle de vérification à l'aide des résultats expérimentaux de Hussain et al. (2010) (volet test).

Différents indices statistiques ont été utilisés pour comparer quantitativement les résultats de la prédiction du coefficient de décharge à l'aide du modèle GEP présenté dans cette étude (équation 9) avec une équation existante (Swamee et Swamee, 2010). Ces valeurs d'indice pour le modèle proposé (équation 9) et le modèle existant (Swamee et Swamee, 2010) sont présentées au tableau 3.

Tableau 3. Étude quantitative du modèle GEP présenté dans cette étude et équation de Swamee et Swamee (2010) utilisant différents indices statistiques.

Models	R^2	MAPE	RMSE	BIAS	SI
GEP	0.98	0.20	0.0017	0.0001	0.0028
Swamee and Swamee (2010)	0.81	3.33	0.025	0.0197	0.039

Le tableau indique que le modèle GEP présenté produit de meilleurs résultats pour tous les indices statistiques de ce tableau que le modèle de Swamee et Swamee (2010). Selon le tableau, GEP prédit le coefficient de décharge avec $R^2= 0,98$ (coefficient de corrélation) tandis que R^2 est d'environ 0,81 pour le modèle de Swamee et Swamee (2010). Les valeurs

des indices *RMSE*, *MAPE* et *SI* pour le modèle de Swamee et Swamee (2010) sont près de 10 fois celles de l'équation GEP présentée, ce qui indique la précision supérieure de prédiction du coefficient de décharge par GEP par rapport aux équations existantes.

L'analyse de sensibilité est également appliquée dans cette étude pour étudier l'effet de chacun des paramètres sur l'estimation du coefficient de décharge. Par conséquent, en supprimant l'un des paramètres efficaces sur l'estimation du coefficient de décharge d'un orifice latéral circulaire dans chaque état (équation 5), les résultats de sortie GEP sont évalués. Les résultats de l'analyse indiquent que le fait de ne pas utiliser le nombre de Froude (*Fr*) comme paramètre efficace dans la prédiction du coefficient de décharge entraîne une plus grande diminution de la précision de prédiction par rapport aux autres paramètres. Le coefficient de corrélation minimal ($R^2=0,82$) et les valeurs maximales de différents indices statistiques ($MAPE=0,699$, $RMSE =0,0057$, $BIAS=0,0044$ et $SI=0,009$) appartiennent à cet état (modèle 2). En outre, le rapport entre la largeur du canal principal et le paramètre sans dimension du diamètre de l'orifice (B/D) a le moins d'effet sur la prédiction du coefficient de décharge dans un orifice latéral circulaire, avec $MAPE=0,27$, $RMSE=0,0028$, $BIAS=0,002$ et $SI=0,005$. Le fait de ne pas utiliser le rapport entre la hauteur de la crête et le diamètre de l'orifice (W/D) et le rapport entre la profondeur d'écoulement dans le canal principal et le diamètre de l'orifice (Y_m/D) qui sont des paramètres sans dimension a un effet similaire sur les résultats de prédiction du coefficient de décharge à l'aide de GEP. Les résultats de l'analyse de sensibilité effectuée sont présentés au tableau 4.

Tableau 4. Résultats de l'analyse de sensibilité à l'aide de différents indices statistiques pour le modèle GEP

Model's No.	dependent parameter	Independent parameters	R^2	MAPE	RMSE	BIAS	SI
1	C_d	$B/D, Y_m/D, W/D, Fr$	0.98	0.23	0.0019	0.00	0.003
2	C_d	$B/D, Y_m/D, W/D$	0.82	0.699	0.0057	0.0044	0.009
3	C_d	$B/D, Y_m/D, Fr$	0.94	0.359	0.0035	0.0022	0.0045
4	C_d	$B/D, W/D, Fr$	0.94	0.398	0.0031	0.0025	0.005
5	C_d	$Y_m/D, W/D, Fr$	0.95	0.27	0.0028	0.002	0.0038

La figure 6 montre l'effet de sortie du modèle GEP sur la distribution du nombre de Froude. L'exactitude du modèle suggéré dans cette étude pour différents nombres de Froude est étudiée à l'aide du rapport d'écart (DR) dans cette figure, qui inclut les montants prévus et réels. La meilleure prédiction est lorsque $DR=1$, tandis que les montants prévus supérieurs ou inférieurs indiquent la surestimation du coefficient de décharge (ou la sous-estimation). Les résultats du modèle GEP pour différentes valeurs DR de différents nombres de froude (Fr) présentés dans cette étude sont présentés à la figure 6. Le DR maximal, moyen et minimum obtenu est de 1,01, 1,0006 et 0,995 avec le modèle GEP. La figure montre également que beaucoup de points sont proches de 1. Le modèle GEP a sous-estimé le coefficient de décharge à une valeur DR de 0,375 ($DR=0,995$) et surestimé le Cd à une valeur Fr de 0,155 ($DR=1,01$).

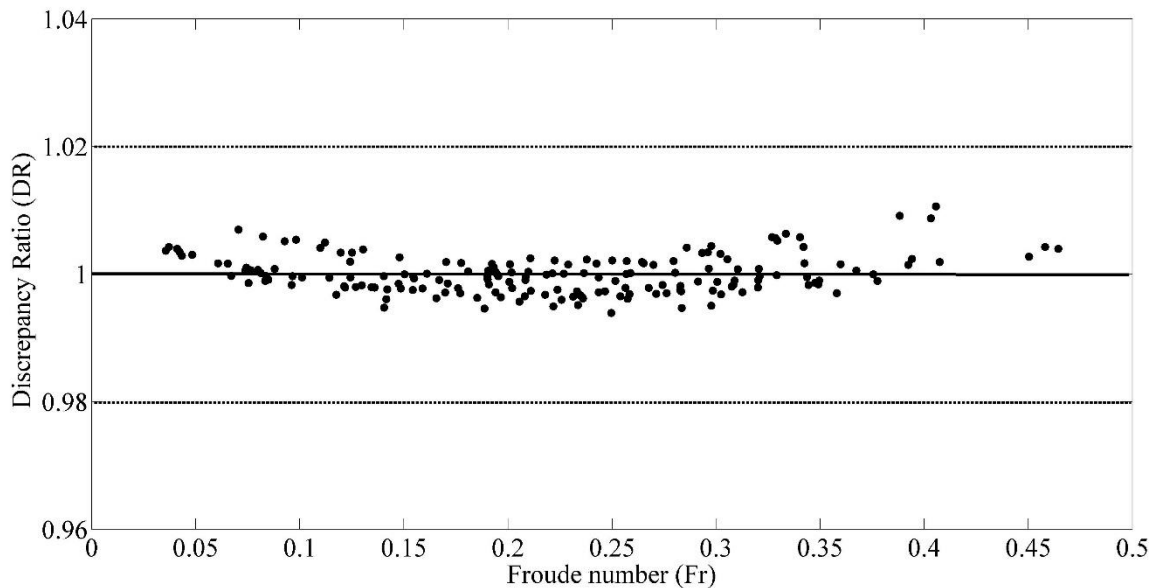


Figure 6. Valeurs DR versus Fr pour le modèle GEP.

Dans cette étude, la tendance de l'équation (9) concernant chaque paramètre indépendant est vérifiée à l'aide d'une analyse aux dérivées partielles de sensibilité (Ebtehaj et al., 2015b). La procédure de cette méthode consiste à calculer la dérivée partielle du modèle GEP liée aux variables B/D , W/D , Y_m/D et Fr , dont l'équation (9) dépend. Si la valeur dérivée partielle est plus élevée, l'effet du paramètre indépendant

sur l'estimation de la valeur C_d à l'aide de l'équation (9) est plus grand. Si la valeur dérivée partielle d'un paramètre est négative (ou positive), elle indique qu'une augmentation de la valeur de ce paramètre entraîne une diminution (ou une augmentation) de la valeur C_d à l'aide de l'équation (9). La figure 7 présente les résultats dérivés partiels de l'équation 9 en termes de variables B/D , W/D , Y_m/D et Fr . Avec une valeur dérivée partielle positive liée à B/D et W/D , une augmentation de la valeur de ces deux paramètres conduit à une augmentation de la valeur C_d calculée avec l'équation (9). Les résultats dérivés partiels de l'équation (9) liés aux deux autres paramètres (Y_m/D et Fr) ont une tendance très différente de B/D et W/D . De plus, les valeurs dérivées partielles calculées à la figure (7) indiquent que le paramètre Fr a le plus d'effet sur l'estimation de C_d avec l'équation (9).

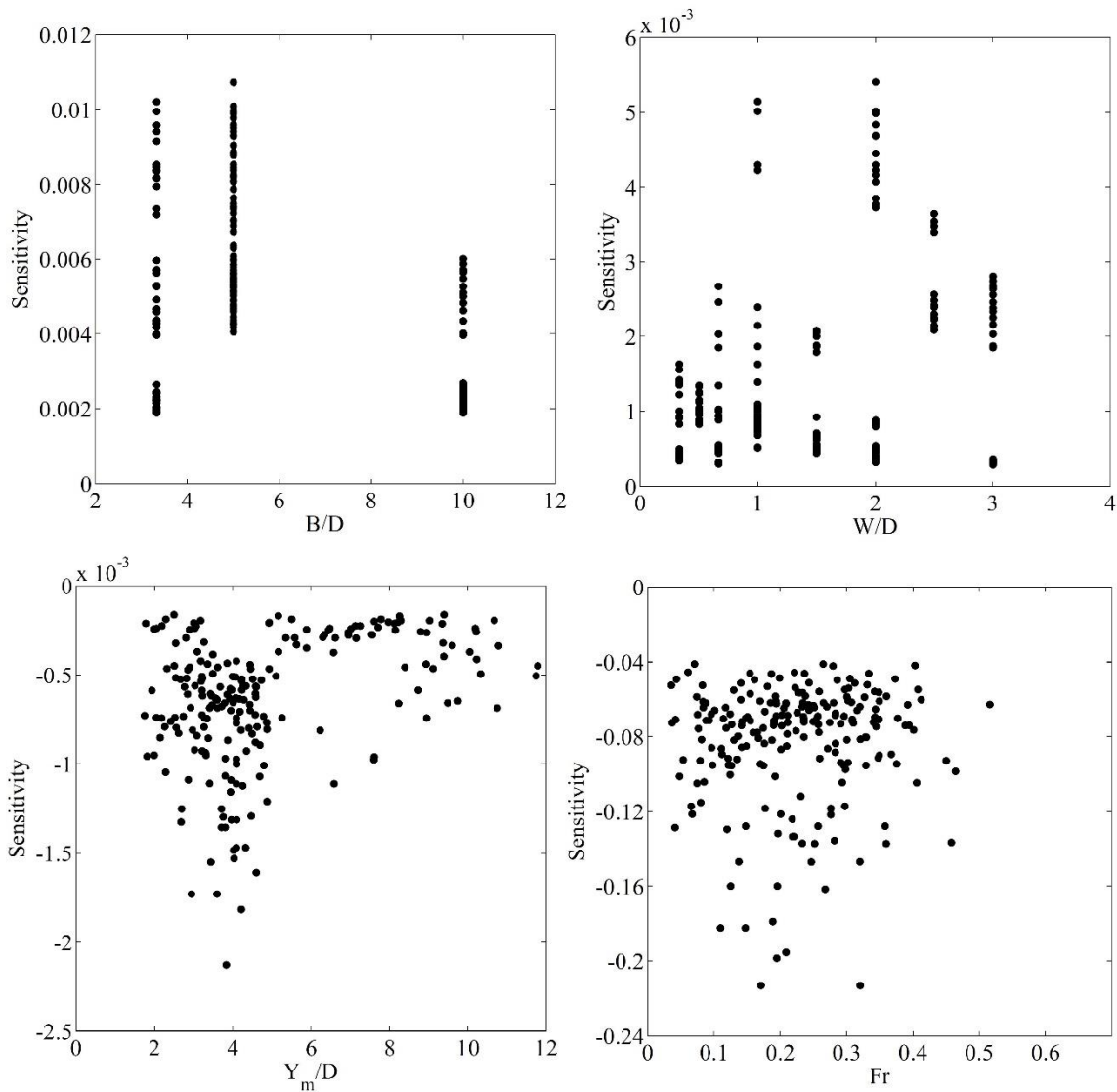


Figure 7. Résultats de l'analyse de sensibilité pour différentes variables d'entrée.

CONCLUSION

Dans la présente étude, la programmation d'expression génétique GEP a été employé pour prévoir le coefficient de décharge d'un orifice latéral circulaire. Pour ce faire, l'analyse dimensionnelle a d'abord été utilisée pour identifier les paramètres affectant la prédiction du coefficient de décharge. Les paramètres sans dimension sont le rapport entre la largeur du canal principal et le diamètre de l'orifice (B/D), le rapport entre la hauteur de la crête et le diamètre de l'orifice (W/D), le rapport entre la profondeur d'écoulement dans le canal principal et le diamètre de l'orifice (Y_m/D) et le nombre de

Froude (Fr). Le modèle GEP a été présenté sous forme d'équation (9) pour la prédiction à l'aide de 80 % de toutes les données expérimentales, tandis que 20 % des données qui n'avaient aucun rôle dans la prédiction du modèle ont servi à vérifier l'exactitude de la prédiction du coefficient de rejet par le modèle proposé. Les résultats de l'enquête indiquent la grande précision du modèle présenté ($R^2=0,97$, $MAPE=0,256$, $RMSE=0,0023$, $BIAS=0,0017$ et $SI=0,0037$). L'exactitude de la prédiction GEP a également été comparée à l'équation de Swamee et Swamee (2010), ce qui signifie la supériorité du modèle présenté. De plus, une analyse de sensibilité a été effectuée afin d'étudier l'effet de chacun des quatre paramètres sans dimension sur la précision de prédiction du coefficient de décharge. L'analyse a révélé la plus grande importance du paramètre du nombre de Froude (Fr) dans la prédiction du coefficient de décharge par rapport aux autres paramètres. On a également observé que l'utilisation du rapport entre la largeur du canal principal et le diamètre de l'orifice (B/D) qui est un paramètre adimensionnel a eu le moins d'effet sur la prédiction du coefficient de décharge. Par conséquent, compte tenu des explications fournies, GEP a été proposé comme une nouvelle méthode alternative pour estimer le coefficient de décharge dans les orifices latéraux circulaires. En outre, une analyse partielle de sensibilité dérivée a indiqué l'augmentation des valeurs de B/D et de W/D et la diminution de la valeur de Y_m/D , alors que Fr a mené au calcul amélioré de la valeur de C_d en utilisant l'équation (9). L'équation simple proposée avec GEP peut être utilisée comme une méthode alternative pour l'ingénierie pratique.

ACKNOWLEDGEMENTS

The authors acknowledged the financial support provided by Natural Science and Engineering Research Council of Canada (NSERC) Discover Grant (#RGPIN-2020-04583)

RÉFÉRENCES

- Ab. Ghani, A., and Azamathulla, H. Md. 2010. Gene-expression programming for sediment transport in sewer pipe systems. *J. Pipeline Syst. Eng. Pract.* 2(3): 102-106. [doi:10.1061/(ASCE)PS.1949-1204.0000076].
- Aqil, M., Kita, I., Yano, A., and Nishiyama, S. 2007. Analysis and prediction of flow from local source in a river basin using a Neuro-fuzzy modeling tool. *J. Environ. Manage.* 85(1): 215-223. [doi:10.1016/j.jenvman.2006.09.009].

Alavi, A. H., and Gandomi, A. H. 2011. A robust data mining approach for formulation of geotechnical engineering systems. *Eng. Comput.* 28(3): 242-274. [doi:10.1108/02644401111118132].

Aytek, A., and Kisi, O. 2008. A genetic programming approach to suspended sediment modelling. *J. Hydrol.* 351(3-4): 288-298. [doi:10.1016/j.jhydrol.2007.12.005]

Azamathulla, H. Md., and Ab Ghani, A. 2010. Genetic Programming to Predict River Pipeline Scour. *J. Pipeline Syst. Eng. Pract.* 1(3): 127-132. [doi:10.1061/(ASCE)PS.1949-1204.0000060].

Baghalian, S., Bonakdari, H., Nazari, F., and Fazli, M. 2012. Closed-form solution for flow field in curved channels in comparison with experimental and numerical analyses and Artificial Neural Network. *Eng. Appl. Comput. Fluid Mech.* 6(4): 514-526. [doi:10.1080/19942060.2012.11015439].

Baylar, A., Unsal, M., Ozkan, F., and Kayadelen, C. 2014. Estimation of Air Entrainment and Aeration Efficiencies of Weirs using Genetic Expression Programming. *KSCE J. Civil Eng.* 18 : 1632-1640. [doi:10.1007/s12205-014-1058-1].

Bryant, D. B, Khan A. A., Aziz, N. M. 2008. Investigation of flow upstream of orifices. *J. Hydraul. Eng.* 134(1): 98-104. [doi:10.1061/(ASCE)0733-9429(2008)134:1(98)].

Bonakdari, H., Baghalian, S. Nazari, F., and Fazli, M. 2011. Numerical analysis and prediction of the velocity field in curved open channel using Artificial Neural Network and Genetic Algorithm. *Eng. Appl. Comput. Fluid Mech.* 5(3): 384-396. [doi: 10.1080/19942060.2011.11015380].

Chang, C. K., Azamathulla, H. Md., Zakaria, N. A., and Ab Ghani, A. 2012. Appraisal of soft computing techniques in prediction of total bed material load in tropical rivers. *J. Earth Syst. Sci.* 121(1): 125-133. [doi: 10.1007/s12040-012-0138-1].

Ebtehaj, I., and Bonakdari, H. 2013. Evaluation of Sediment Transport in Sewer using Artificial Neural Network. *Eng. Appl. Comput. Fluid Mech.* 7(3): 382-392. [doi:10.1080/19942060.2013.11015479].

Ebtehaj, I., and Bonakdari, H. 2014a. Performance Evaluation of Adaptive Neural Fuzzy Inference System for Sediment Transport in Sewers. *Water Resour. Manage.* 28(13): 4765-4779. [doi:10.1007/s11269-014-0774-0].

Ebtehaj, I., and Bonakdari, H. 2014b. Comparison of genetic algorithm and imperialist competitive algorithms in predicting bed load transport in clean pipe. *Water Sci. Technol.*70(10): 1695-1701. [doi:10.2166/wst.2014.434].

Ebtehaj, I., Bonakdari, H., Khoshbin, F., and Azimi, H. 2014. Pareto Genetic Design of GMDH-type Neural Network for Predict Discharge Coefficient in Rectangular Side Orifices. *Flow Meas. Instrum.* 41: 67-74. [doi:10.1016/j.flowmeasinst.2014.10.016].

Ebtehaj, I., Bonakdari, H., Zaji, A. H., Azimi, H., and Sharifi, A. 2015a. Gene Expression Programming to Predict the Discharge Coefficient in Rectangular Side Weirs. *Appl. Soft Comput.*35, 618-628. [doi:10.1016/j.asoc.2015.07.003].

Ebtehaj, I., Bonakdari, H., Zaji, A. H., Azimi, H., and Khoshbin, F. 2015b. GMDH-type neural network approach for modeling the discharge coefficient of rectangular sharp-crested side weirs. *Eng. Sci. Technol. Int. J.* 18(4): 746:757. [doi:10.1016/j.jestch.2015.04.012].

Emiroglu, M. E., Bilhan, O., and Kisi, O. 2011. Neural networks for estimation of discharge capacity of triangular labyrinth side-weir located on a straight channel. *Expert Syst. Appl.* 38(1): 867-874. [doi:10.1016/j.eswa.2010.07.058].

Emiroglu, M.E., Kisi, O. and Bilhan, O. 2010. Predicting discharge capacity of triangular labyrinth side weir located on a straight channel by using an adaptive neuro-fuzzy technique. *Adv. Eng. Soft.* 41(2): 154-160. [doi:10.1016/j.advengsoft.2009.09.006].

Ferreira, C. 2001. Gene expression programming: a new adaptive algorithm for solving problems. *Complex Syst.* 13(2): 87-129.

Gill, M. A. 1987. Flow through side slots. *J. Environm. Eng.* 135, 1047-1057. [doi:10.1061/(ASCE)0733-9372(1987)113:5(1047)].

Hussain, A., Ahmad, Z. and Asawa, G. L. 2010. Discharge characteristics of sharp-crested circular side orifices in open channels. *Flow Meas. Instrum.* 21(3): 418-424. [doi:10.1016/j.flowmeasinst.2010.06.005].

Kisi, O. 2005. Suspended sediment estimation using neuro-fuzzy and neural network approaches. *Hydrolog. Sci. J.* 50(4): 683-696. [doi:10.1623/hysj.2005.50.4.683].

Kisi, O., Emiroglu, M. E. Bilhan, O., and Guven, A. 2012. Prediction of lateral outflow over triangular labyrinth side weirs under subcritical conditions using soft computing approaches. *Expert Syst. Appl.* 39: 3454-3460. [doi:10.1016/j.eswa.2011.09.035].

Kisi, O., and Guven, A. 2010. A machine code-based genetic programming for suspended sediment concentration estimation. *Adv. Eng. Soft.* 41(7-8): 939-945. [doi:10.1016/j.advengsoft.2010.06.001].

Khan, M., Azamathulla, H. Md., Tufail, M., and Ab Ghani, A. 2012. Bridge pier scour prediction by gene expression programming. *P. I. Civil Eng-Wat. Water M.* 165(9): 481-493. [doi:10.1680/wama.11.00008].

Khorchani, M., and Blanpain, O. 2005. Development of a discharge equation for side weirs using artificial neural networks. *J. Hydroinform.* 7(1): 31-39.

Koza, J. 1992. *Genetic Programming, on the Programming of Computers by Means of Natural Selection.* MIT Press, Cambridge, MA, USA.

Nayak, P. C., Sudheer, K. P., Rangan, D. M., and Ramasastri, K. S. 2004. A neuro-fuzzy computing technique for modeling hydrological time series. *J. Hydrol.* 291(1-2): 52-66. [doi:10.1016/j.jhydrol.2003.12.010].

Ojha, C. S. P., and Subbaiah, D. 1997. Analysis of flow through lateral slot. *J. Irrig. Drain. Eng.* 123(5), 402-405. [doi:10.1061/(ASCE)0733-9437(1997)123:5(402)].

Ramamurthy, A. S., Tim, U. S., and Sarraf, S. 1986. Rectangular lateral orifices in open channels. *J. Environ. Eng.* 112(2): 292-300. [doi:10.1061/(ASCE)0733-9372(1986)112:2(292)].

Swamee, P. K., Pathak, S. K., and Ali, M. S. 1993. Weir orifice units for uniform flow distribution. *J. Irrig. Drain. Eng.* 119(6): 1026-1035.

Swamee, P. K., and Swamee, N. 2010. Discharge equation of a circular sharp-crested orifice. *J. Hydraul. Res.* 48(1), 106-107. [doi:10.1080/00221680903568626].