Data-driven-based determination of influential parameters on local energy loss of slope-tapered culvert

Xiaowen Cai, Feng Ye and Ghazaleh Nassaji Matin

ABSTRACT

Local energy loss is among the essential parameters of culvert design, in which uncertainty and nonlinearity is controversial. In the present study, seven models were developed with the aid of the experimental data of slope-tapered culverts, and the efficiency of gene expression programing and Gaussian process regression as a kernel-based approach was assessed in predicting the entrance loss coefficient of a slope-tapered culvert. Also, one-at-a-time (OAT) sensitivity analysis was performed to determine the impact of input parameters. The results of both GEP and GPR methods with the performance criteria of R = 0.847, DC = 0.777, RMSE = 0.2 and R = 0.76, DC = 0.718, RMSE = 0.25 showed that the model with input parameters of Froude number (*Fr*), ratio of headwater to culvert diameter (H_w/D) and ratio of reducer length to barrel length (L_r/L) is the superior model. Although the accuracy of GEP method was slightly higher than GPR, obtained results proved the capability of the applied methods (i.e., high correlation coefficient (R) and coefficient of determination R^2 (DC) and low RMSE). Furthermore, OAT sensitivity analysis revealed that Froude number has the most impact on local loss coefficient and could cause a significant increment in model efficiency.

Key words | entrance loss coefficient, Gaussian process regression, gene expression programing, kernel-based approach, slope-tapered culvert

Xiaowen Cai
Feng Ye (corresponding author)
TaiZhou Keyuan Building Materials Testing Co., Ltd,
TaiZhou, ZheJiang, 318020,
China
and
TaiZhou Water Conservancy Project Quality and Safety Affairs Center,
TaiZhou, ZheJiang, 318000,
China
E-mail: *j.feng_ye@yahoo.com*; *zjyefengig@yeah.net*

Ghazaleh Nassaji Matin
Department of Water Resource Engineering,
Faculty of Civil Engineering,
University of Tabriz,
Tabriz,
Iran

HIGHLIGHTS

- The entrance loss coefficient of slope-tapered culvert was predicted via artificial intelligence approaches.
- The current study utilizes Gene Expression Programming (GEP) and Gaussian Process Regression (GPR) as artificial approaches.
- The obtained results improve the understanding of complex behavior of local energy loss in slope-tapered culverts.

INTRODUCTION

Culverts provide an efficient solution for flowing water through man-made cross barriers (e.g., roads). There is a wide variety of solutions of the inlet part of the culverts, among which, slope-tapered inlet is the focus of this paper. Hydraulic characteristic and sizing of culverts are complex problems for several reasons (Khalil & Zein 1995; Graziano *et al.* 2001a, 2001b; Jones *et al.* 2006; Mohammed 2016; Aly 2017; Wilhere *et al.* 2017; Jaeger *et al.* 2019a, 2019b, 2019c; Rao & Raju 2019; Zeng *et al.* 2019; Nakhostin *et al.* 2020). The optimal culvert could be considered the one in which the head loss is minimal and has the ability to transfer the design flow rate. Culverts are among the drainage structures in which proper and desirable performance is essential. Therefore, the failure of culvert structures is a catastrophic phenomenon. As a result, a detailed study of the causes of failure is recommended (Rigby *et al.* 2002; French & Jones

doi: 10.2166/hydro.2020.189

2018; Günal *et al.* 2019). One of the most recent comprehensive guidelines is *Hydraulic Design Series Number 5* (HDS-5) (Norman *et al.* 2001). Jaeger *et al.* (2019a, 2019b, 2019c) presented findings regarding how different inlet designs increase the discharge rate. The flume flow experiments showed that a slope-tapered inlet can have a greater capacity than a conventional culvert with a square edge. This type of inlet is designed to increase the culvert performance by providing a depression and a more efficient control section at the throat (Graziano 2001a, 2001b; Khaled 2004). In order to prevent flooding, estimation of water levels with reasonable accuracy to determine free-board and top elevation of culverts is essential. Therefore, several studies have tested a variety of culvert configurations in laboratory models and other research has investigated loss coefficient, aquatic organism passage, and culvert blockage (Apelt 1983; Graziano *et al.* 2001a, 2001b). It was found that culvert local losses accrue for a variety of reasons, such as outlet loss, inlet loss, or bend loss (Tullis *et al.* 2005; Malone & Parr 2008; Tullis 2012). Previous studies on the determination of the energy dissipation at culvert outlets has focused on simplified methods to estimate the head loss coefficient. As a general guideline, the head loss is considered to be comparable to the velocity head within the conduit (Simmons 1964; Liu & Zhu 2000; Tullis & Robinson 2008; Schall 2012; Habibzadeh & Rajaratnam 2016). Also, the evolution of the effective parameters on energy losses of buried-invert culverts via kernel-based approaches was carried out by Roushangar *et al.* (2019). Examining the bend loss of culverts showed that the parameters affecting the energy loss were the Froude number and bend angle whereas the parameters affecting the inlet loss were Froude number and depth ratio.

The structural choice of a culvert and corresponding inlet is based on environmental considerations, risk to property, cost of construction and maintenance and also esthetic considerations. Many U.S. agencies and researchers require that culverts be designed and engineered to meet specific federal, state, or local regulations and guidelines to ensure proper function and to protect against culvert failures (Smith 1957; Chen 1970; Simons & Stevens 1972; Rigby *et al.* 2002; Environmental Protection Agency EPA Management 2003; Alberta Transportation 2004; Federal Highway Administration Department of Interior Bureau of Land Management 2006; Günal *et al.* 2019). Defects in the conveying of water in culverts can lead to overflows and structural damage, therefore, to mitigate these risks, the capacity of many culverts needs to be increased. It is clear that there are many tools for enhancing the discharge efficiency of culverts (Harrison *et al.* 1972; Normann 1975). However, many researchers believe that the best way to increase culvert capacity is to improve inlet function. Relying on this belief, the use of inclined headwalls or misaligned culverts has been recommended (Aly 2017; Jaeger 2019a, 2019b, 2019c). Others researchers believed that utilizing improved inlets, such as slope-tapered ones, enhances the flow transient (McGrath & Heger 1983; Graziano *et al.* 2001a, 2001b). It is worth noting that energy loss is among one of the reasons affecting culvert capacity (Kotowski *et al.* 2011). Furthermore, head loss falls into two categories, general and local losses, which in culverts due to their short length the local loss is outstanding. Experimental study of bend loss in rectangular culverts has been developed by Malone & Parr (2008). From another viewpoint, there are other studies that have focused on outlet loss of culverts and have examined various types of culvert systems, such as elliptical, circular, and rectangular (Liu & Zhu 2000; Larson 2004; Robinson 2005; Habibzadeh & Rajaratnam 2016). On the other hand, several experimental studies on inlet loss have stated that the inlet loss coefficient is significantly affected by the inlet configuration and usually has a constant value (Graziano *et al.* 2001a, 2001b; Jones *et al.* 2006; Tullis 2012). However, in contrast, taking advantage of intelligence methods has shown that the local loss coefficient is a dependent parameter, which in addition to geometric parameters, is influenced by hydraulic parameters (Roushangar *et al.* 2019).

Many studies have considered the entrance local energy loss coefficient of the slope-tapered culvert to be constant at 0.2 (Norman *et al.* 2001). However, some researchers have considered this coefficient as a function of geometric parameters and flow characteristics (Graziano *et al.* 2001a, 2001b; Roushangar *et al.* 2019). Energy losses are a complex process in which the existing regression models do not show favorable accuracy and the results are associated with large error. Therefore, the present study proposed artificial intelligence approaches to predict the entrance loss coefficient and evaluate the best input variables with the most impact on inlet loss coefficient of slope-tapered culverts. In the past decades, the application of artificial intelligence approaches such as artificial neural networks (ANNs), neuro-fuzzy models (NF), genetic programming (GP), gene expression programming (GEP), support vector machine (SVM), and Gaussian process regression (GPR) have become popular in water resources engineering, leading to numerous publications in this field (Liriano & Day 2001; Hazi & Ghani 2011; Hazi & Haque 2012; Najafzadeh 2016; Huang *et al.* 2017; Carvalho *et al.* 2018; Amaranto *et al.* 2018; Tayyebi *et al.* 2018; Owen & Liuzzo 2019; Poursorkhabi & Ghasempour 2019; Roushangar *et al.* 2019; Tinelli & Juran 2019; Zhu *et al.* 2019; Roushangar & Shahnazi 2020). Detailed review of previous studies (Graziano *et al.* 2001a, 2001b; Tullis 2012) revealed that the local loss coefficient of slope-tapered culverts was monitored experimentally and, to the best of the authors' knowledge, there is a lack of comprehensive research on estimating the entrance loss coefficient using artificial intelligence. As a result, this study aims to evaluate the efficiency of GPR and GEP for predicting the entrance loss coefficient of slope-tapered culverts. Also, a new equation to predict the local loss coefficient was provided utilizing the GEP method and the most important input variables were identified using OAT sensitivity analysis.

MATERIALS AND METHODS

In order to investigate the parameters affecting the local loss coefficient in the culverts with a slope-tapered inlet, the laboratory dataset of Graziano *et al.* (2001a, 2001b) was used. All of the laboratory testing was done at the Federal Highway Administration's Turner-Fairbank Highway Research Center located in McLean, Virginia, USA and inlet control design constants and entrance loss coefficients were calculated for the circular slope-tapered culverts with various reducer lengths. The experimental setup was constructed primarily of plywood and consisted of a 2.43 m long by 2.43 m wide headbox, and a 1.21 m wide by 2.43 m long tailbox, which was located 4.5 m downstream of the headbox. The slope-tapered inlet and culvert barrel spanned the 4.5 m between the headbox and the tailbox. A sketch of the slope-tapered culvert is shown in Figure 1. Fourteen pressure ports were inserted along the bottom of the culvert setup to measure hydraulic depth. The total entrance loss (H_{etot}) was measured for outlet control tests by projecting the energy grade line (EGL) in the headbox and the EGL for the culvert barrel to a common plane at the upstream end of the culvert barrel.

SIMULATION AND MODEL DEVELOPMENT

Data characterization

Providing an appropriate dataset is a critical step in the prediction of local loss coefficient via artificial intelligence methods. Examination of models showed that considering 75% of the dataset for training goals and the remaining 25% for testing goals leads to more accurate results. Owing to more accurate estimate of out-of-sample accuracy and more efficient use of data, as every observation is used for both training and testing, and in order to avoid model bias, the train and test dataset was divided using v-fold cross validation which was developed in STATISTICA

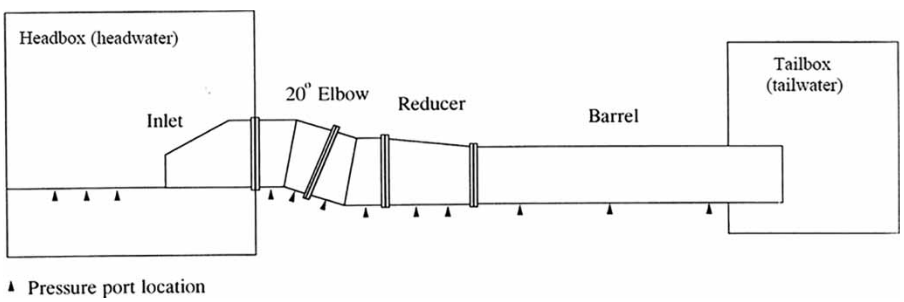

Figure 1 | Slope-tapered culvert.

software (v. 8). The separated dataset is then utilized in further model developments. V-fold cross validation is a standard approach for model selection and the main idea of cross validation is data splitting (Arlot & Lerasle 2016).

Input variables

Appropriate allocation of input parameters is a principal step during the modeling process. Based on Graziano *et al.* (2001a, 2001b) and Roushangar *et al.* (2019), the most important parameters in predicting the energy loss coefficient (Ke) of slope-tapered culverts included the upstream velocity (V), culvert diameter (D), gravity acceleration (g), flow depth (H_w), and reducer length (L_r). Therefore, the considered dimensionless input parameters are expressed in Equation (1):

$$Fr,\ \frac{H_W}{D},\ \frac{L_r}{L} \tag{1}$$

where $Fr = V/(g \times H_w)^{0.5}$ is the Froude number, a dimensionless value that describes flow regimes, H_w/D is the ratio of headwater to culvert diameter, and $\frac{L_r}{L}$ is the ratio of reducer length to barrel length. In order to develop the local loss models in slope-tapered circular culverts, seven models with different input variables were considered (Table 1).

Gaussian process regression

Kernel-based approaches (such as GPR) are one of the common methods for solving the nonlinear problems which are based on statistical learning theory, and the appropriate selection of kernel type in these approaches is the most important step due to its direct impact on the training and classification precision. Availability of sufficient input data enables these models to predict any variable. They are also fairly robust against overfitting, especially in high-dimensional space (Roushangar *et al.* 2019).

GPR models are based on the assumption that adjacent observations should convey information about one another. Thus, GP regression models are able to understand the predictive distribution corresponding to the test inputs, and due to prior knowledge about the data and functional dependencies, no validation process is required for generalization (Rasmussen & Williams 2006). $x \times y$ accounts for the domains of input and output, respectively, wherein n pairs of (x_i, y_i) are drawn independently and distributed identically. By considering $y \subseteq \mathrm{Re}$, a GP on x is defined by a mean function $\mu{:}x \to \mathrm{Re}$. The main assumption of GP considered y equal to $y = f(x) + \xi$, in which $\xi \sim \mathrm{N}\ (0, \sigma^2)$. Every input x in GP regression has an associated random variable $f(x)$, which is the value of the stochastic function f at that location. In the present study, it is assumed that observational error ξ is normally independent and identically distributed with mean value of zero ($\mu(x) = 0$), a variance of σ^2 and $f(x)$ drawn from the Gaussian process on x specified by k (Equation (2)):

$$y = (y_1,\ \dots, y_n) \sim N(0,\ k + \sigma^2 I) \tag{2}$$

in which $K_{ij} = k(x_i, x_j)$ and I is the identity matrix. For a given vector of test data X^*, the predictive distribution of the corresponding output $Y*/(X, Y),\ X* \sim N(\mu,\ \Sigma)$ is Gaussian where:

$$\mu = K(X*, X*)(K(X, X) + \sigma^2 I)^{-1} Y \tag{3}$$

$$\Sigma = K(X*, X*) - \sigma^2 I - K(X*,\ X)(K(X, X) + \sigma^2 I)^{-1} K(X, X*) \tag{4}$$

$K(X, X*)$ represents the $n \times n*$ matrix of covariance, if there are n training data and $n*$test data. This is also true for the other values of $K(X, X), K(X*, X)$ and $K(X*, X*)$; here, X and Y are the vectors of the training data labels y_i,

Table 1 | Model development

Model	Input variables
M(I)	Fr, H_w/D, L_r/L
M(II)	Fr, H_w/D
M(III)	Fr, L_r/L
M(IV)	H_w/D, L_r/L
M(V)	Fr
M(VI)	H_w/D
M(VII)	L_r/L

whereas $X*$ is the vector of the test data. A specified covariance function is required to develop a positive semi-defined covariance matrix K, where $K_{ij} = k\ (x_i, x_j)$. Choosing the appropriate covariance function and parameters is essential during the training process of GP regression models. In the case of GP regression with a fixed value of Gaussian noise, a GP model can be trained by applying Bayesian inference which leads to the minimization of negative log-posterior (Equation (5)):

$$P(\sigma^2, k) = \frac{1}{2}y^T(K + \sigma^2 I)^{-1}y + \frac{1}{2}log|K + \sigma^2 I| - \log p(\sigma^2) - \log p(k) \quad (5)$$

To assess the hyper-parameters, the partial derivation of Equation (4) can be obtained with respect to σ^2 and k. For more detailed discussion of GP regression, see the study of Kuss (2006). The optimal value of capacity constant (c) and the size of error-intensive zone (ε) in GPR are required due to their high impact on the accuracy of the mentioned regression approaches. The optimum values of these parameters were obtained after the trial-and-error process. In order to develop the GPR approach, a code written in MATLAB software was executed.

Gene expression programming (GEP)

Gene expression programming was developed by Ferreira (2001) using fundamental principles of the genetic algorithms (GA) and GP. The strength of the proposed approach includes the simplicity of creating genetic diversity, and a unique and multi-genic nature which allows the evaluation of more complex programs composed of several subprograms. GEP as GA mimics the biological evolution to create a computer program for simulating a specified phenomenon. A GEP algorithm begins by selecting five elements, including the function set, terminal set, fitness function, control parameters, and stopping condition. There is a comparison between predicted values and actual values in each subsequent step. When the desired results are obtained in accordance with previously selected error criteria, the GEP process is terminated. After the desired fitness score is achieved, the process terminates and then the chromosomes are decoded for the best solution of the problem. The most important advantages of GEP are as follows (Ferreira 2001): (1) the chromosomes are simple entities and (2) the expression trees are exclusively the expression of their respective chromosomes. It should be noted that the software used to develop this approach is GeneXproTools 4.0 Release 2.

Performance criteria

In the current study, the model's performance was evaluated using three statistical parameters: correlation coefficient (R), coefficient of determination R^2 (DC), and root mean square error (RSME). It is worth noting that the RMSE has the same unit of the target parameter (K_e) which is a non-dimensional parameter. The expressions for performance criteria are presented in Equation (6):

$$DC = 1 - \frac{\sum_{i=1}^{N}(I_0 - I_P)^2}{\sum_{i=1}^{N}(I_0 - \overline{I_P})^2},\ R = \frac{\sum_{i=1}^{N}(I_0 - \overline{I_0}) \times (I_P - \overline{I_P})}{\sqrt{\sum_{i=1}^{N}(I_0 - \overline{I_0})^2 \times (I_P - \overline{I_P})^2}},\ RMSE = \sqrt{\sum_{i=1}^{N}\frac{(I_0 - I_P)^2}{N}} \quad (6)$$

where I_0,I_P, $\bar{I}_0$, $\overline{I_P}$, N, respectively, are the measured values, predicted values, mean measured values, mean predicted values, and number of data samples.

GPR and GEP model development

The design of GP-based regression approach involves use of the concept of kernel functions. In order to select the best kernel function, model M (I) was predicted using various kernels. According to the statistical parameters, using the kernel function of squared-exponential led to more accurate prediction. Figure 2 indicates the results of the statistical parameters of different kernels for model M (I).

Furthermore, GEP was trained for local loss coefficient prediction with basic arithmetic operators of (+, − ,*, /) and several mathematical functions (exp, x^2, x3,$\sqrt{}$) as the function set. Different combinations of chromosomes' structure presented in Table 2 were tested. Then the model was run

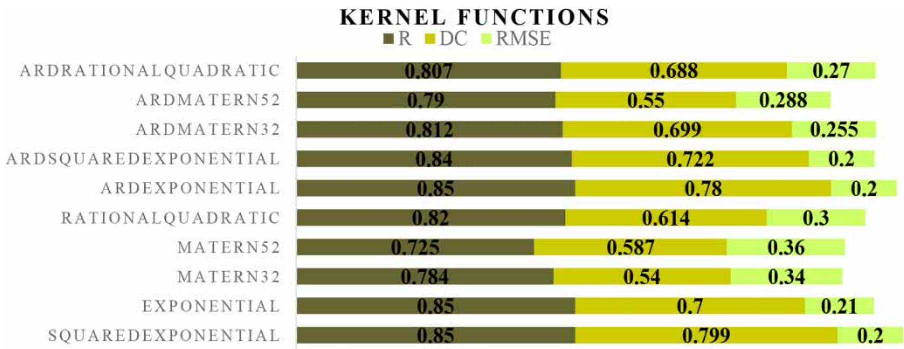

Figure 2 | Statistics parameters via GPR kernel function types for a testing set of model M (I).

Table 2 | Optimized parameters of GEP models

Parameter	**Setting**
Function set	$+, -, \times, /, x^2, x^z, \sqrt{}$
Chromosomes	25, **30**, 35
Head size	7, **8**
Linking function	Addition
Fitness function	Root mean square error

for a number of generations and stopped when there was no significant change in the fitness function value and coefficient of correlation. Accordingly, the models with 30 chromosomes head size of 8 and 3 numbers of genes led to the better results. In addition, one of the important steps in preparing the GEP model is choosing the appropriate set of genetic operators, therefore, a combination of all genetic operators was also tested. Characteristics of optimized GEP models are shown in Table 2.

RESULTS AND DISCUSSION

GEP-based developed models

In order to evaluate the impact of various parameters on inlet loss coefficient of slope-tapered culverts, several models were developed based on flow characteristics (Froude number) and geometric parameters (reducer's length and depth ratio). For predicting the inlet loss coefficient, all models were GPR and GEP models which were trained and tested. The obtained results are listed in Table 3 as well as Figures 2 and 3. From the R, DC, and RMSE viewpoints (highest R and DC and lowest RMSE), it can be concluded that among the seven models, the model with input parameters of Fr, H_w/D, and L_r/L leads to more accurate results (0.847, 0.777, 0.2) compared to others, and therefore, it was identified as the best model. However, it can still be seen that models M (II) and M (III) produce an acceptable result with R, DC, and RMSE equal to 0.827, 0.747, 0.224 and 0.815, 0.69, 0.24, respectively.

Moreover, although model M (V) did not lead to undesirable results, examination of the superior model showed that adding the geometric parameters of L_r/L and H_w/D to the input combination caused an increment in model efficiency. Also, the performance criteria (R, DC, RMSE) of models

Table 3 | Statistical parameters of the GPR and GEP models

		Evaluation criteria					
		Train				**Test**	
Models		**R**	**DC**	**RMSE**	**R**	**DC**	**RMSE**
M (I)	GPR	0.826	0.83	0.098	0.76	0.718	0.254
	GEP	0.942	0.912	0.140	0.847	0.777	0.200
M (II)	GPR	0.827	0.827	0.100	0.760	0.530	0.250
	GEP	0.928	0.887	0.160	0.827	0.747	0.227
M (III)	GPR	0.840	0.750	0.257	0.720	0.514	0.162
	GEP	0.924	0.860	0.206	0.815	0.682	0.240
M (IV)	GPR	0.672	0.490	0.903	0.312	0.251	0.405
	GEP	0.459	0.420	0.390	0.415	0.450	0.360
M (V)	GPR	0.820	0.710	0.250	0.520	0.320	0.160
	GEP	0.843	0.770	0.224	0.642	0.56	0.3
M (VI)	GPR	0.527	0.46	0.15	0.508	0.330	0.320
	GEP	0.421	0.38	0.25	0.415	0.345	0.340
M (VII)	GPR	0.398	0.340	0.360	0.333	0.209	0.290
	GEP	0.360	0.286	0.160	0.310	0.180	0.506

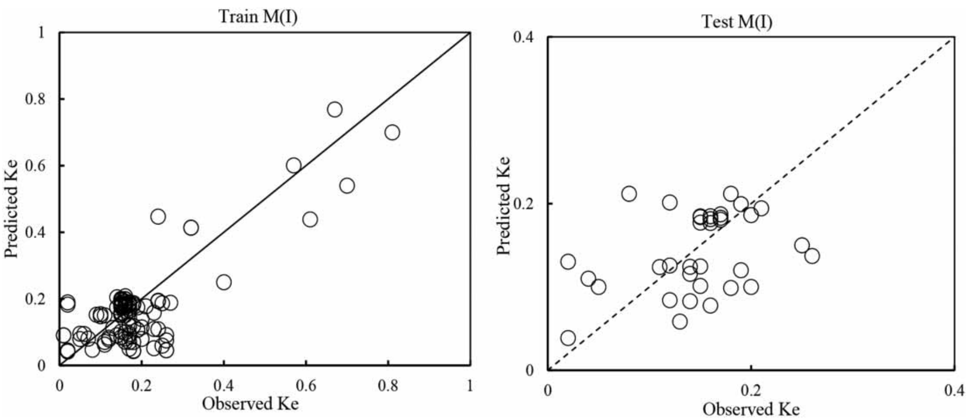

Figure 3 | Comparison of observed and predicted local loss for GEP best model (M (I)).

M (IV), M (VI), and M (VII) showed unfavorable efficiency in predicting the local loss of slope-tapered culverts. This issue confirms that, in the presence of insufficient information, the inlet loss coefficient of slope-tapered culverts could be estimated using input variables of F_r, H_w/D, and L_r/L. The comparisons of observed and predicted local loss coefficient of GEP models are presented in Figures 3 and 4. Also, the strength point of GEP models is the ability of models to develop an explicit expression of the relationship between the variables. Since the local loss coefficient is a dependent parameter of Froude number, relative depth, and ratio of reducer's length to the barrel length, the mathematical expression of GEP for the best model is presented in Equation (7). It was observed that the equation has three numbers of genes with plus operator as presented in Table 2. In addition, investigating the obtained equation revealed that the input parameter of Froude number has the major effect in predicting the local loss coefficient of the culvert.

$$K_e = \left[\frac{0.12Fr^2}{Fr^2 - 0.12\dfrac{H_w}{D}}\right] + \left[\frac{\dfrac{L_r}{L}}{5.6Fr - 13.2}\right] + \left[\frac{2}{Fr^2}\right] \quad (7)$$

GPR-based developed models

The results obtained from GPR model developments are illustrated in Table 3 and Figures 5 and 6. The results were similar to the GEP models, in which the superior model was M (I) with performance criteria (R, DC, RMSE) equal to 0.76, 0.718, and 0.25. By comparing R, DC, and RMSE in GEP and GPR models of this case, it seems that the accuracy of GEP models was slightly higher.

Sensitivity analysis

The impact of the different employed parameters on inlet loss coefficient of a slope-tapered culvert is evaluated using sensitivity analysis. There are different approaches of sensitivity analysis, including local or global, quantitative or qualitative, or one-at-a-time (OAT). One of the simplest and most common approaches is that of OAT, to determine what effects this produces on the output. OAT sensitivity analysis essentially consists of selecting an initial parameter setting (nominal set) and varying one parameter at a time while keeping all the other parameters fixed. Consequently, OAT reveals a form of the relationship between the varied parameter and the output, given that all other parameters have their nominal values (Holvoet *et al.* 2005).

The current study utilized the OAT approach to evaluate the impact of each parameter. To do so, the models were run with all input parameters, then, one of the parameters was eliminated and the process was repeated. Based on the results from Figure 7, it could be deduced that in predicting the inlet loss, with eliminating L_r/L, H_w/D, and Froude, the statistical parameters of R and DC decreased to 0.82, 0.74, 0.815, 0.69, and 0.43, 0.45,

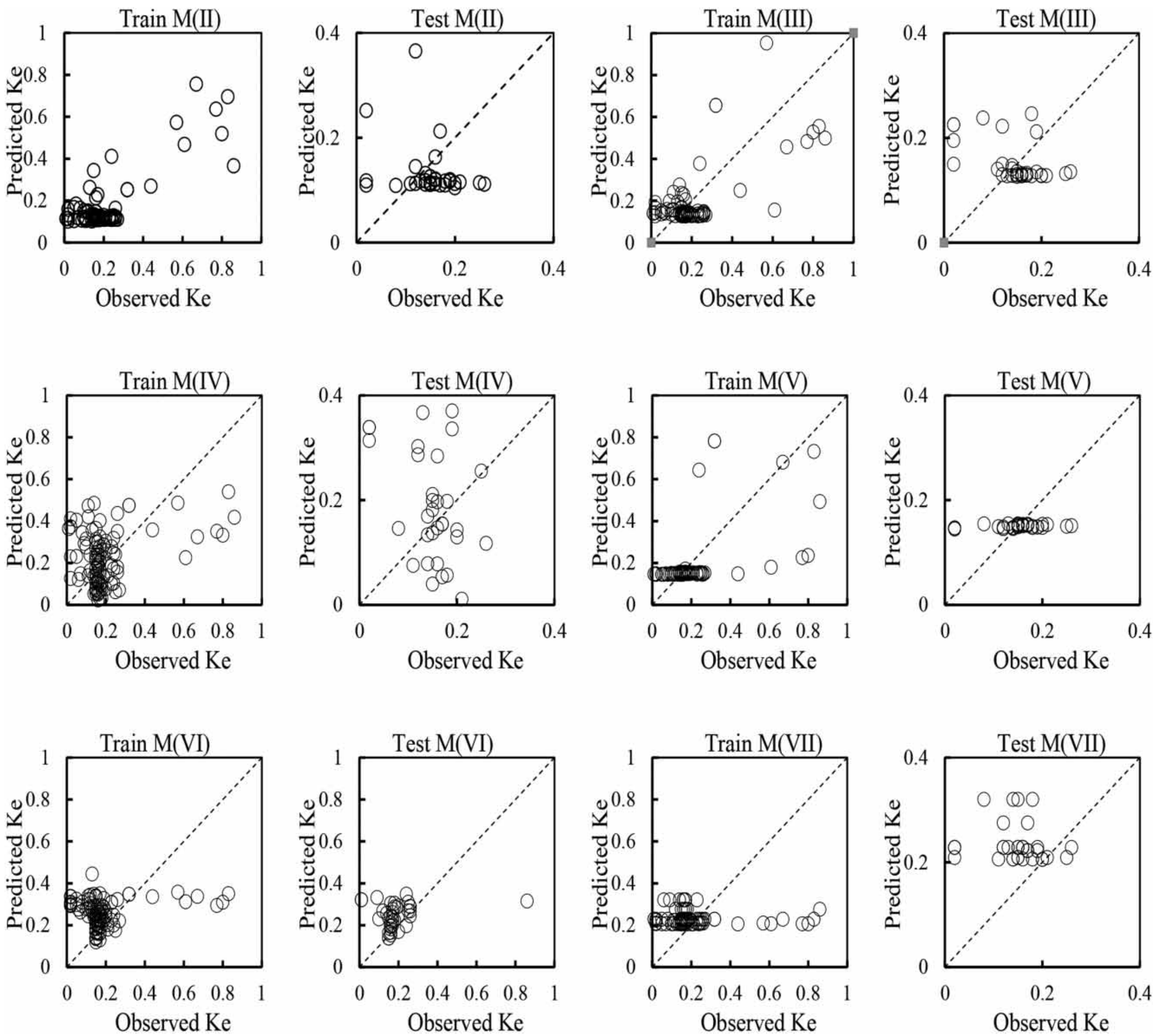

Figure 4 | Comparison of observed and predicted local loss GEP models.

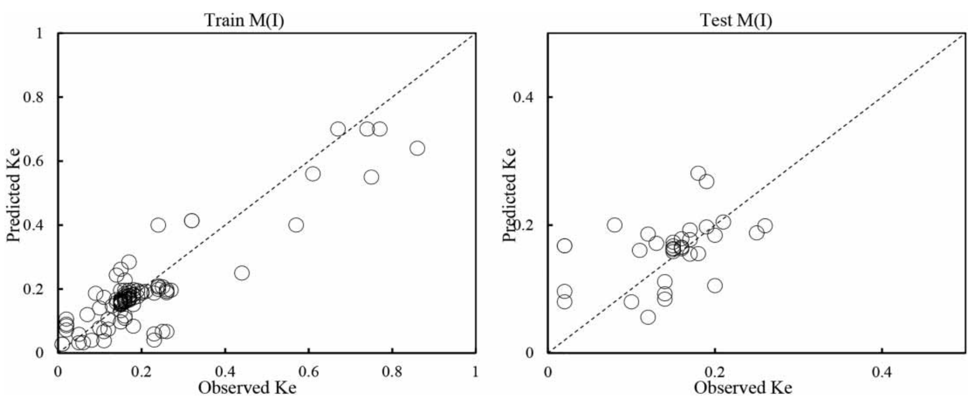

Figure 5 | Comparison of observed and predicted local loss for GPR best model (M (I)).

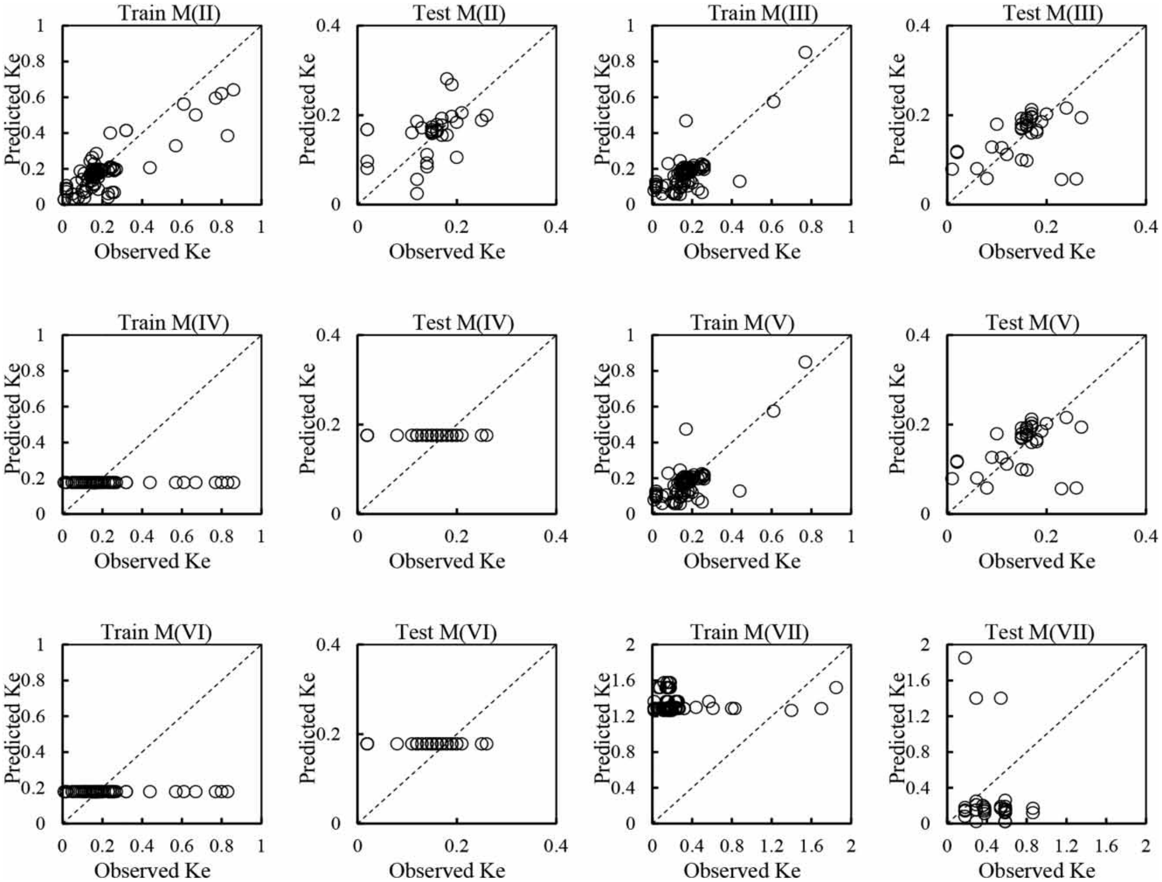

Figure 6 | Comparison of observed and predicted local loss GPR models.

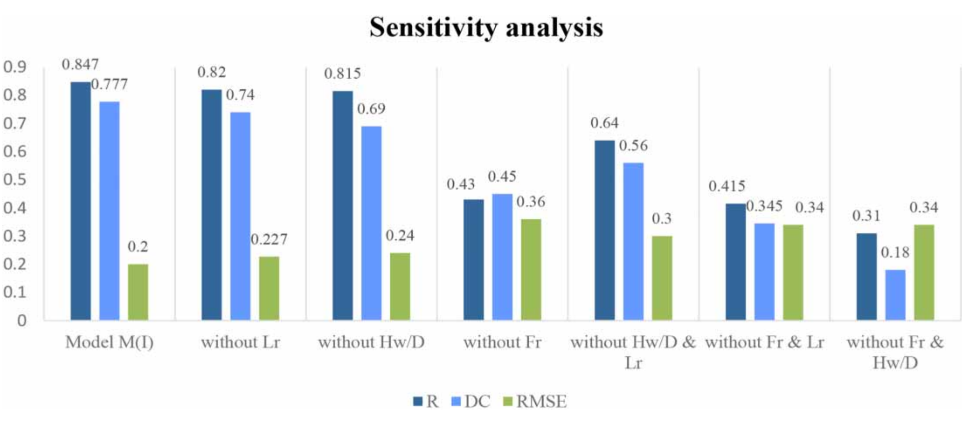

Figure 7 | Relative significance of each of the input parameters of the best model (M (I)).

and the amount of RMSE error increased to 0.227, 0.24, and 0.36, respectively. Consequently, variable *Fr*, would be the most significant impact in the current study and model M (I) with input parameters of *Fr*, L_r/L, H_w/D would be the superior model. Also, investigating the developed models revealed that variables H_w/D and L_r/L had scant impact on local loss coefficient and the impact of H_w/D was slightly more than L_r/L.

CONCLUSION

Inlet edge configuration is one of the prime factors influencing the performance of a culvert. As the inlet configuration is improved, the flow contraction is reduced, thus, a slope-tapered inlet improves culvert performance by providing a more efficient control section. The defections due to the flow of water in culvert systems can cause problems such as overflow of water on the road. The optimum design of culverts depends on many factors, including but not limited to, the entrance loss coefficient, which is the focus in this paper. Since many previous studies have pointed out the dependency of local loss coefficient on geometric and hydraulic parameters, it seems that having sufficient knowledge of parameters affecting the entrance loss coefficient could be useful in the accurate calculation of inlet energy loss of culvert systems and the upcoming optimum design of culverts. In the present study, the capability of artificial intelligence methods including GEP and GPR in the estimation of entrance local energy loss coefficient of a slope-tapered culvert was assessed. For this purpose, laboratory data of Graziano *et al.* (2001a, 2001b) of a slope-tapered culvert were used. In order to perform a detailed investigation on entrance loss coefficient, seven models with various input parameters were selected.

According to the results of the GEP approach, it was found that model M (I) with input variables of Fr, H_w/D, and L_r/L leads to better results, and it was identified as the superior model (R = 0.847, DC = 0.777, and RMSE = 0.2). Unacceptable results of model M (IV) showed that adding Froude numbers increases the efficiency, significantly. Also, by comparing model M (V) with model M (I), it was observed that parameters H_w/D and L_r/L had a slight contribution in increasing the accuracy of the model. Moreover, an explicit expression of the relationship of variables was developed using the GEP method.

The results obtained using the GPR method also provided similar results, in which model M (I) with performance criteria of 0.76, 0.718, and 0.254 was the best model in estimating the entrance loss coefficient of the slope-tapered culvert. Also, comparison of evaluation criteria (R, DC, RMSE) of the GEP and GPR methods illustrated a relatively high accuracy of the GEP method in predicting the entrance loss.

From the obtained results of OAT sensitivity analysis, it was found that the correlation coefficient between K and Fr (in the state of entrance loss of a slope-tapered culvert) was significantly higher than other parameters. Therefore, Froude number had the most impact on entrance loss coefficient. By investigating M (II) to M (IV) models, it can be deduced that eliminating input parameters of L_r/L, H_w/D, and Fr increases the amount of RMSE error up to 0.227, 0.24, and 0.36, respectively. Also, scrutinizing models M (VI) and M (VII) revealed that, in the absence of Froude number, input variables of H_w/D and L_r/L are the most affecting parameters, respectively.

It was concluded that the proposed approaches were found to be able to predict the entrance loss coefficient of the slope-tapered culvert successfully. This issue confirms that in the presence of insufficient information, the inlet loss coefficient of the slope-tapered culvert could be estimated via these approaches. However, it is worth noting that the selected models are data-driven ones, hence, in order to discover the merits of the models to estimate the entrance loss coefficient in real conditions of flow, further studies using data ranges beyond the scope of this study are essential.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

REFERENCES

Alberta Transportation 2004 *Design Guidelines for Bridge Size Culverts*. PDF Original Document 1995 Alberta Transportation, Technical Standards Branch, Government of the Province of Alberta.

Aly, T. E. 2017 Improving the pipe culvert efficiency by using inclined headwalls. *International Water Technology Journal* **7**, 1.

Amaranto, A., Munoz-Arriola, F., Corzo, G., Solomatine, D. P. & Meyer, G. 2018 Semi-seasonal groundwater forecast using multiple data-driven models in an irrigated cropland. *Journal of Hydroinformatics* **20** (6), 1227–1246.

Apelt, C. J. 1983 Minimum energy culverts and bridge waterways. In *NAASRA BEC 1983 Seminar, Waterway Analysis and Design*, Adelaide, Australia.

Arlot, S. & Lerasle, M. 2016 Choice of V for V-fold cross-validation in least-squares density estimation. *The Journal of Machine Learning Research* **17** (1), 7256–7305.

Carvalho, J., Santos, J. P. V., Torres, R. T., Santarém, F. & Fonseca, C. 2018 Tree-based methods: concepts, uses and limitations under the framework of resource selection models. *Journal of Environmental Informatics* **32** (2), 112–124.

Chen, Y. H. 1970 *Scour at Outlets of Box Culverts*. PhD Thesis, Colorado State University, Fort Collins, CO, USA.

Department of Interior Bureau of Land Management 2006 Culvert use, installation, and sizing. Chapter 8 (PDF). *Low Volume Engineering Journal*. blm.gov/bmp

Environmental Protection Agency EPA Management 2003 *Culverts-Water. NPS Unpaved Roads Chapter 3 (PDF)*, 'CULVERTS' epa.gov.

Ferreira, C. 2001 *Gene Expression Programming: A New Adaptive Algorithm for Solving Problems. arXiv preprint cs/0102027*.

French, R. & Jones, M. 2018 Design for culvert blockage: the ARR 2016 guidelines. *Australasian Journal of Water Resources* **22** (1), 84–87.

Graziano, F., Martin, B., Stein, S. & Umbrell, E. 2001a *Hydraulics of Iowa DOT Slope-Tapered Pipe Culverts (No. FHWA-RD-01-077)*. Turner-Fairbank Highway Research Center, McLean, VA, USA.

Graziano, F., Martin, B., Stein, S. & Umbrell, E. 2001b *South Dakota Culvert Inlet Design Coefficients (No. FHWA-RD-01-076)*. Turner-Fairbank Highway Research Center, McLean, VA, USA.

Günal, M., Günal, A. Y. & Osman, K. 2019 Simulation of blockage effects on scouring downstream of box culverts under unsteady flow conditions. *International Journal of Environmental Science and Technology* **16** (9), 5305–5310.

Habibzadeh, A. & Rajaratnam, N. 2016 Turbulent jet approach to predict expansion head loss at submerged outlets. *Journal of Hydraulic Engineering* **142** (10), 06016011.

Harrison, L. J., Morris, J. L., Normann, J. M. & Johnson, F. L. 1972 Hydraulic design of improved inlets for culverts. *Hydraulic Engineering Circular* (13), 184 pp.

Hazi, M. A. & Ghani, A. A. 2011 ANFIS-based approach for predicting the scour depth at culvert outlets. *Journal of Pipeline Systems Engineering and Practice* **2** (1), 35–40.

Hazi, M. A. & Haque, A. A. M. 2012 Prediction of scour depth at culvert outlets using gene-expression programming. *International Journal of Innovative Computing, Information and Control* **8** (7B), 5045–5054.

Holvoet, K., van Griensven, A., Seuntjens, P. & Vanrolleghem, P. A. 2005 Sensitivity analysis for hydrology and pesticide supply towards the river in SWAT. *Physics and Chemistry of the Earth, Parts A/B/C* **30** (8–10), 518–526.

Huang, F., Huang, J., Jiang, S. H. & Zhou, C. 2017 Prediction of groundwater levels using evidence of chaos and support vector machine. *Journal of Hydroinformatics* **19** (4), 586–606.

Jaeger, R., Jacobs, C., Tondera, K. & Tindale, N. 2019a Improving flows in misaligned culverts. *Water* **11** (9), 1932.

Jaeger, R., Tondera, K., Jacobs, C., Porter, M. & Tindale, N. 2019b Numerical and physical modeling to improve discharge rates in open channel infrastructures. *Water* **11** (7), 1414.

Jaeger, R., Tondera, K., Pather, S., Porter, M., Jacobs, C. & Tindale, N. 2019c Flow control in culverts: a performance comparison between inlet and outlet control. *Water* **11** (7), 1408.

Jones, J. S., Kerenyi, K. & Stein, S. 2006 *Effects of Inlet Geometry on Hydraulic Performance of box Culverts (No. FHWA-HRT-06-138)*. United States Federal Highway Administration, Office of Infrastructure Research and Development, Washington, DC, USA.

Khaled, A. A. 2004 *Improving the Hydraulic Performance of Highway Culverts*. PhD Thesis, Helwan University, Helwan, Egypt.

Khalil, M. B. & Zein, S. 1995 An improvement to culvert performance and capacity. *Engineering Research Journal* 422–435.

Kotowski, A., Szewczyk, H. & Ciezak, W. 2011 Entrance loss coefficients in pipe hydraulic systems. *Environment Protection Engineering* **37** (4), 105–117.

Kuss, M. 2006 *Gaussian Process Models for Robust Regression, Classification, and Reinforcement Learning*. PhD Thesis, Echnische Universität Darmstadt, Darmstadt, Germany.

Larson, E. A. 2004 *Energy Dissipation in Culverts by Forcing A Hydraulic Jump at the Outlet*. PhD Thesis, Department of Civil and Environmental Engineering, Washington State University, Pullman, WA, USA.

Liriano, S. L. & Day, R. A. 2001 Prediction of scour depth at culvert outlets using neural networks. *Journal of Hydroinformatics* **3** (4), 231–238.

Liu, M. & Zhu, D. Z. 2000 Study of tunnel outlet head-loss coefficient. *Canadian Journal of Civil Engineering* **27** (6), 1306–1310.

Malone, T. R. & Parr, A. D. 2008 *Bend Losses in Rectangular Culverts (No. K-TRAN: KU-05-5)*. Depatment of Transportation, Kansas, USA.

McGrath, T. J. & Heger, F. J. 1983 *Structural Design Manual for Improved Inlets and Culverts (No. FHWA-IP-83-6 Final Rpt.)*. US Department of Transportation, Federal Highway Administration, Turner-Fairbank Highway Research Center, McLean, VA, USA.

Mohammed, M. S. 2016 Characteristics and prediction of simultaneous flow over broad-crested weirs and through culverts. *The Egyptian International Journal of Engineering Sciences & Technology* **6** (1), 133–137.

Najafzadeh, M. 2016 Neurofuzzy-based GMDH-PSO to predict maximum scour depth at equilibrium at culvert outlets. *Journal of Pipeline Systems Engineering and Practice* **7** (1), 06015001.

Nakhostin, E., Kenny, S. & Sivathayalan, S. 2020 Numerical performance assessment of buried corrugated metal culvert subject to service load conditions. *Canadian Journal of Civil Engineering*. https://doi.org/10.1139/cjce-2019-0316

Norman, J. M., Houghtalen, R. J. & Johnston, W. J. 2001 *Hydraulic Design Series Number 5 (HDS-5), Hydraulic Design of Highway Culverts*. Federal Highway Administration (FHWA), Washington, DC, USA.

Normann, J. M. 1975 Improved design of highway culvert. *Civil Engineering – ASCE* **45** (3), 70–73.

Owen, N. E. & Liuzzo, L. 2019 Impact of land use on water resources via a Gaussian process emulator with dimension reduction. *Journal of Hydroinformatics* **21** (3), 411–426.

Poursorkhabi, R. V. & Ghasempour, R. 2019 Possibilities to use the meta model and classical approaches to evaluate the impact of hydraulic conditions in prediction of the critical submergence depth ratio. *Water Supply* **19** (4), 1055–1065.

Rao, K. R. & Raju, P. M. 2019 Special considerations for design of storm water drainage system – a case study. In: *Water Resources and Environmental Engineering* (M. Rathinassmy, S. Chandramouli, K. B. V. N. Phanindra & U. Nanduri, eds). Springer, Singapore, pp. 43–55.

Rasmussen, C. E. & Williams, C. K. 2006 *Gaussian Processes for Machine Learning*. The MIT Press, Cambridge, MA, USA.

Rigby, E. H., Boyd, M. J., Roso, S., Silveri, P. & Davis, A. 2002 Causes and effects of culvert blockage during large storms. In *Global Solutions for Urban Drainage: Proceedings of the Ninth International Conference on Urban Drainage*, Portland, OR, USA.

Robinson, S. C. 2005 *Hydraulic Characteristics of A Buried Invert Elliptical Culvert Inlet and Quantification of Culvert Exit Loss*. PhD Thesis, Utah State University, Logan, UT, USA.

Roushangar, K. & Shahnazi, S. 2020 Prediction of sediment transport rates in gravel-bed rivers using Gaussian process regression. *Journal of Hydroinformatics* **22** (2), 249–262.

Roushangar, K., Matin, G. N., Ghasempour, R. & Saghebian, S. M. 2019 Evaluation of the effective parameters on energy losses of rectangular and circular culverts via kernel-based approaches. *Journal of Hydroinformatics* **21** (6), 1014–1029.

Schall, J. D. 2012 *Hydraulic Design of Highway Culverts (No. FHWA-HIF-12-026)*. United States. Federal Highway Administration, Washington, DC, USA.

Simmons, W. P. 1964 Transitions for canals and culverts. *Journal of the Hydraulics Division* **90** (3), 115–153.

Simons, D. B. & Stevens, M. A. 1972 Scour control in rock basins at culvert outlets. *River Mechanics* **2**, 24–62.

Smith, G. L. 1957 *An Analysis of Scour Below Culvert Outlets*. PhD Thesis, Colorado State University, Fort Collins, CO, USA.

Tayyebi, A., Tayyebi, A. H., Pekin, B. K., Omrani, H. & Pijanowski, B. C. 2018 Modeling historical land use changes at a regional scale: applying quantity and locational error metrics to assess performance of an artificial neural network-based back-cast model. *Journal of Environmental Informatics* **31** (2), 74–84.

Tinelli, S. & Juran, I. 2019 Artificial intelligence-based monitoring system of water quality parameters for early detection of non-specific bio-contamination in water distribution systems. *Water Supply* **19** (6), 1785–1792.

Tullis, B. P. 2012 *Hydraulic Loss Coefficients for Cculverts*, Vol. 734. Transportation Research Board, Washington, DC, USA.

Tullis, B. P. & Robinson, S. C. 2008 Quantifying culvert exit loss. *Journal of Irrigation and Drainage Engineering* **134** (2), 263–266.

Tullis, B. P., Robinson, S. C. & Young, J. C. 2005 Hydraulic characteristics of buried-invert elliptical culverts. *Transportation Research Record* **1904** (1), 104–112.

Wilhere, G. F., Atha, J. B., Quinn, T., Tohver, I. & Helbrecht, L. 2017 Incorporating climate change into culvert design in Washington State, USA. *Ecological Engineering* **104**, 67–79.

Zeng, J., Ansar, M., Rakib, Z., Wilsnack, M. & Chen, Z. 2019 Applications of computational fluid dynamics to flow rating development at complex prototype hydraulic structures: case study. *Journal of Irrigation and Drainage Engineering* **145** (12), 05019009.

Zhu, S., Luo, X., Xu, Z. & Ye, L. 2019 Seasonal streamflow forecasts using mixture-kernel GPR and advanced methods of input variable selection. *Hydrology Research* **50** (1), 200–214.

First received 16 May 2020; accepted in revised form 13 August 2020. Available online 23 October 2020