Predicting peak breach discharge due to embankment dam failure

Jasna Duricic, Tarkan Erdik and Pieter van Gelder

ABSTRACT

Predicting peak breach discharge due to embankment dam failure is of vital importance for dam failure prevention and mitigation. Because, when dams fail, property damage is certain, but loss of life can vary depending on flood area and population. Many parametric breach models based on regression techniques have been developed so far. In this study, an efficient model is proposed to forecast peak discharge from the height of the water and volume of water behind the dam at failure, respectively, by using the Kriging approach. The previous studies, which consist of 13 numerical models, are used as a benchmark for testing the proposed new model, by employing five different error criteria. Moreover, a new database is compiled by extending the previous one. In addition, it is demonstrated that R^2, which only quantifies the dispersion between measurements and predictions, should not be considered alone for checking the model capabilities. At least, the other criteria should be employed together with R^2. As a result, it is shown that one can forecast the peak flow discharge with more significant accuracy by the proposed model than other previous models developed so far.

Key words | dambreak, embankment dam, kriging, peak discharge

Jasna Duricic
Tarkan Erdik (corresponding author)
Pieter van Gelder
Section of Hydraulic Structure and Flood Risk,
Faculty of Civil Engineering and Geosciences,
Delft University of Technology,
P.O. Box 5048,
2600 GA Delft,
The Netherlands
E-mail: *tarkanerdik@hotmail.com*;
erdik@itu.edu.tr

INTRODUCTION

Embankments, including both earth dams and sea dikes, are constructed worldwide for protecting the people residing near and along the coast. Currently, there are about 45,000 dams higher than 15 m worldwide, of which about 73% were built in the last 50 years (International Commission on Large Dams (ICOLD) 1997). However, the embankment failure can cause devastating disasters if it cannot resist when extreme nature strikes. In 1889, overtopping of the South Fork Dam, Pennsylvania, USA, induced over 2,200 deaths and large property losses (Singh & Scarlatos 1988). In February 1953, in the Netherlands, dike breaches at about 900 places due to heavy storm surge, led to one of the biggest natural disasters in Dutch history, causing 1,835 people to lose their lives and a direct economic loss of about 14% of the Dutch gross domestic product (GDP) (Huisman *et al.* 1998). Finally, in August 1975, the catastrophic heavy rainstorms (maximum 6-hour rainfall 830 mm) in central China caused disastrous failures of the Banqiao and the Shimantan Reservoir Dams, giving rise to 26,000 deaths (Pan 2000).

As described above, embankment failure can cause cataclysmic experiences. Therefore, understanding the physics of breach growth and predicting peak discharge due to embankment dam failure in determining the safety of dams are gaining growing attention throughout the world with the intention of reducing loss of life and damage to downstream properties. The earth embankment dams fail gradually due to breaching by overtopping which erodes their materials by water flow or wave action (American Society of Civil Engineers/Environmental & Water Resources Institute (ASCE/EWRI) Task Committee on Dam/Levee Breaching 2011).

In the literature, different classifications exist concerning embankment dam breaching (Wahl 1998; D'Eliso 2007). The most recent ASCE/EWRI Task Committee on Dam/Levee Breaching has divided embankment breach models as parametric, simplified physically-based, or detailed physically-based, considering the model formulation and approximation of physical processes. These can be further classified into empirical, analytical, and

doi: 10.2166/hydro.2013.196

numerical models in terms of their solution method employed. Parametric models, which are the focus of this research, are usually empirical, physically-based models derived by using regression models (RMs).

Before the late 1970s, peak discharge calculations were rarely used. During the 1980s, many researchers began gathering detailed information about dam breaching in order to develop models that are able to predict the effects and mechanisms of breach and estimate peak outflows. Kirkpatrick (1977) is among the first researchers to develop the peak discharge formula by employing a curve fitting technique. Since then, many researchers have followed him by compiling and/or using case studies data compiled by others to develop empirical equations relating peak breach outflow to dam height, reservoir storage volume, or combinations of both. Regarding the complexity of the dam breach phenomenon, existent models are often insufficient. Table 1 summarizes the types of relations proposed by various researchers from Kirkpatrick (1977) up to now.

The approaches published so far in Table 1 are based on RM incorporating important relevant parameters, which are practical applications aimed to forecast peak discharge. Recently, Wahl (1998) presented a database of 108 composite case studies by various investigators. Pierce *et al.* (2010) reviewed the previous efforts that suggested empirical regression equations for predicting peak discharge due to dambreak between the period 1977 and 1995. Moreover, they collected 44 case studies and added to the 43 entries of Wahl (1998) that include height (H) and volume (V) of water behind the dam at failure, hence resulting in 87 case studies, from which they developed enhanced relationships based on regression analysis.

In this study, 16 discreet investigations presented so far are compiled and included in Table 1, where Q_p = peak outflow due to breaching (m^3/s), H_w = height of the water behind the dam at failure (m), H_d = height of the dam (m), S = reservoir storage at normal pool (m^3), V_w = volume of the water behind the dam at failure (m^3), L = dam length, W_{awe} = average breach width (Figure 1). In Table 1, the height of the water, H, behind the dam at failure, is used to combine the H_w and H_d. Similarly, V, the volume of water behind the dam at failure, is used to combine the terms V_w and S (Pierce *et al.* 2010). It should be noted that

Table 1 | Brief description of previous regression equations for predicting peak discharge from breached embankment dams

Equation type	Equation number	Researcher(s)	Developed equation(s)	Best-fit or envelope
Height of water equations	(1)	Kirkpatrick (1977)	$Q_p = 1.268\ (H_w + 0.3)^{2.5}$	Best-fit
	(2)	Soil Conservation Service (SCS) (1981) for dams >31.4 m	$Q_p = 16.6\ (H_w)^{1.85}$	Envelope (as suggested by Wahl (1998))
	(3)	US Bureau of Reclamation (1982)	$Q_p = 19.1\ (H_w)^{1.85}$	Envelope
	(4)	Singh & Snorrason (1982)	$Q_p = 13.4\ (H_d)^{1.89}$	Best-fit
	(5)	Pierce *et al.* (2010) – linear	$Q_p = 0.784\ (H)^{2.668}$	Best-fit
	(6)	Pierce *et al.* (2010) – curvilinear	$Q_p = 2.325\ \ln(H)^{6.405}$	Best-fit
Storage equations	(7)	Singh & Snorrason (1984)	$Q_p = 1.776(S)^{0.47}$	Best-fit
	(8)	Evans (1986)	$Q_p = 0.72(V_w)^{0.53}$	Best-fit
	(9)	Pierce *et al.* (2010)	$Q_p = 0.00919(V)^{0.745}$	Best-fit
Height of water	(10)	Hagen (1982)	$Q_p = 1.205(V_w H_w)^{0.48}$	Envelope
and storage equations	(11)	MacDonald & Langridge-Monopolis (1984)	$Q_p = 1.154(V_w H_w)^{0.412}$	Best-fit
	(12)	MacDonald & Langridge-Monopolis (1984)	$Q_p = 3.85(V_w H_w)^{0.411}$	Envelope
	(13)	Costa (1985)	$Q_p = 0.763(V_w H_w)^{0.42}$	Best-fit
	(14)	Froehlich (1995)	$Q_p = 0.607(V_w^{0.295} H_w^{1.24})$	Best-fit
	(15)	Pierce *et al.* (2010)	$Q_p = 0.0176(VH)^{0.606}$	Best-fit
	(16)	Pierce *et al.* (2010)	$Q_p = 0.038(V^{0.475} H^{1.09})$	Best-fit

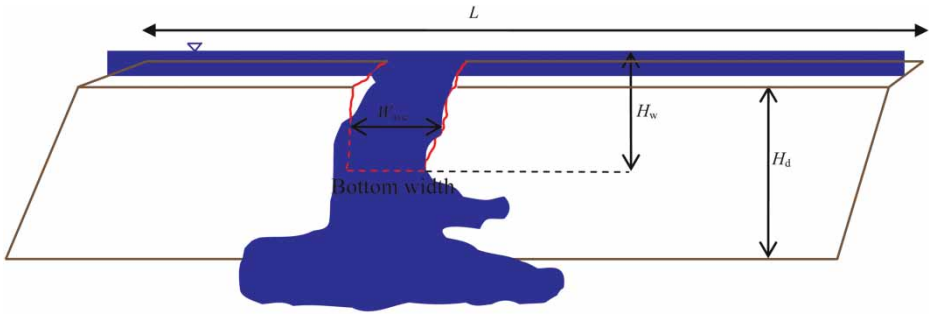

Figure 1 | Dam breaching parameters.

since relationships are derived from a limited number of databases and the method for determining peak flows are unknown in some cases, inconsistency is inherited into the system. In this research, the Kriging model (KM) is developed by following the notation of Pierce *et al.* (2010).

This study objectives are to: (1) examine previous research studies for predicting peak discharge due to embankment dam failure based on regression equations; (2) extend the database of Pierce *et al.* (2010); and (3) construct a new model for peak discharge estimation and compare it with others in the literature.

From the detailed literature review, it is seen that this paper is the first to apply the KM, or a new approach rather than RMs, in peak discharge calculation by embankment dam failure. The KM was first introduced for geostatistical calculations (Matheron 1963). Since then, it has been applied for various scientific areas such as drought analysis by Sirdaş & Şen (2003), level fluctuations in Lake Van, Turkey, by Altunkaynak *et al.* (2003), ocean engineering studies by Altunkaynak (2005), wave height and period prediction by Ozger & Sen (2007), hydrological researches by Altunkaynak (2009), suspended solid concentrations by Altunkaynak & Wang (2010) and finally estimation of missing precipitation records by Teegavarapu (2009).

RESTRICTIVE ASSUMPTIONS IN REGRESSION ANALYSIS

The equations developed so far in Table 1 benefited from regression analysis, which is a widely-used approach that could provide acceptable results. However, this approach bears restrictive assumptions. Direct application of regression analysis ignoring the restrictive assumptions might cause pitfalls and biased calculations. The restrictive assumptions of regression analysis such as Normality, Means of conditional distributions, and Homoscedasticity Independence, are briefly described in the papers by Uyumaz *et al.* (2006), Erdik (2009) and Erdik *et al.* (2009).

If any one of these assumptions is not applied, then the regression analysis estimations are under suspicion. In order to check the validity of the regression analysis, the total data of 87 case studies of Pierce *et al.* (2010) are considered. The first violation in the assumption that the independent variable must be normally distributed is shown in Figure 2. It is clear from the Figure 2 that both H and V data of Pierce *et al.* (2010) do not fit the normal probability plot. The great discrepancy is clearly visible in rectangular blocks. In addition, in order to measure the discrepancy existing between observed and theoretical frequencies of both H and V, the Kolmogorov–Smirnov test is applied at a 0.05 significance level. The computed Kolmogorov–Smirnov values for both H and V are 0.3621 and 0.2035, respectively. However, the critical value is 0.1436 for both, at 5% significance level. Since the values 0.3621 and 0.2035 are greater than 0.1436, the hypothesis that the observed data fit the normal distribution is violated.

The second violation occurs in the assumption that conditional distribution of the residuals should be normally distributed. As seen in Figure 3, the residuals for both Equations (15) and (16) by Pierce *et al.* (2010) do not abide by normal distribution function.

The third violation occurs in the assumption 'Means of conditional distributions', stating that mean average errors must be zero. However, these values are calculated as 1,319.2 and 1,295.5 by using the Equations (15) and (16) of Pierce *et al.* (2010) in Table 1, using 87 case studies.

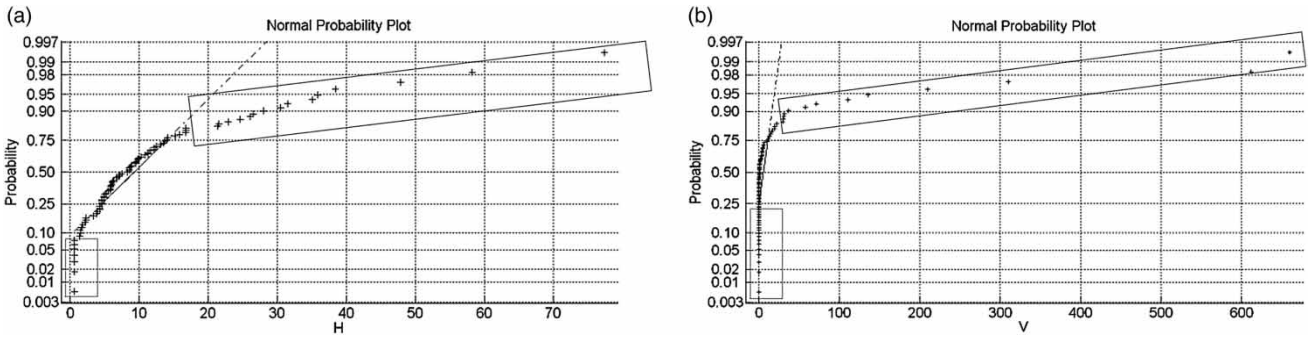

Figure 2 | Normal probability curves for independent variables (a) H and (b) V, using 87 data of Pierce *et al.* (2010).

It is rather obvious that these assumptions are not satisfied collectively in many problems such as for the data in the research by Pierce *et al.* (2010). Although the equations developed by RM yield better results than others in terms of error statistics, they must be cautiously employed. The KM, which has no restrictive assumptions, is suggested in this paper for better and robust application and subsequent predictions.

DATA USED

Pierce *et al.* (2010) extended the previous 43 case study data of Wahl (1998, 2004) by including 44 case studies, hence resulting in a composite database of 87 case studies. In this study, 11 additional case study data for embankment dam failure due to overtopping of Xu & Zhang (2009) are added to this database. The resulting database consists of 98 case study data (Table 2). The total database is first checked for outliers. To this aim H, V and Q_p data are analyzed separately and data which cannot satisfy the criteria defined such as $H = (H\text{-mean}(H)) <= 3*\text{std}(H)$ are excluded. Hence, four data out of 98, such as Banqiao – China, Euclides de Cunha – Brazil, and Oros-Brazil, Teton Id. – USA, are removed from the database.

THEORETICAL BASIS OF KRIGING METHODOLOGY

In this section, brief information is given about the KM, which is an optimum interpolation method employing the empirical sample Semivariogram (SV) to weight sample points based on their locations in space relative to the point value that is to be estimated. Kriging interpolation technique is built on the assumption that things that are close to one another are more alike than those farther away (Şen 2009). The SV is a measure to quantify this relationship that pairs which are close in distance should have a smaller measurement difference than those farther away.

The KM consists of different alternative types such as Simple Kriging, Ordinary Kriging, Universal Kriging, and Block Kriging. In this study, Ordinary Kriging methodology, which is similar to multiple linear regression and interpolates values based on point estimates, is applied. Ordinary Kriging

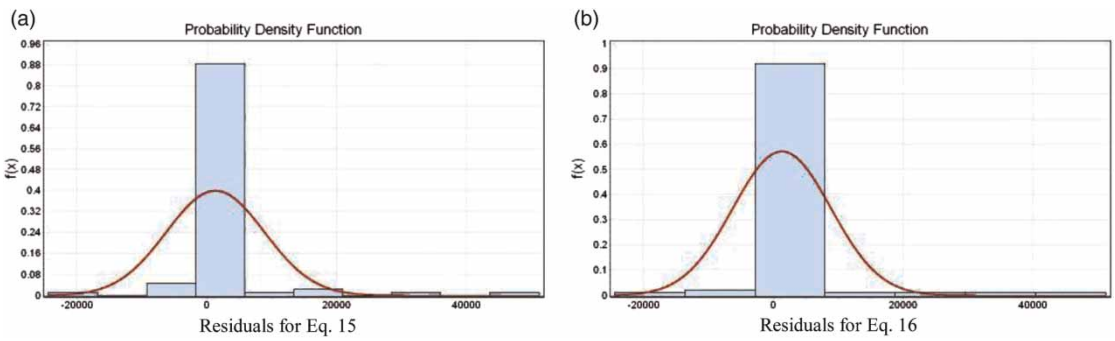

Figure 3 | Normal distribution functions for residuals (a) H and (b) V, using 87 data of Pierce *et al.* (2010).

Table 2 | Data collected

Site	V (10^6 m^3)	H (m)	Q_p (m^3/s)	Data by	Remark
Apishapa, Colo.	22.20	28.00	6,850.00	Pierce *et al.* (2010)	Included
Baldwin Hills, Calif.	0.91	12.20	1,130.00	Pierce *et al.* (2010)	Included
Banqiao, China	612.00	26.11	78,000.00	Pierce *et al.* (2010)	Removed
Big Bay Dam, Miss.	17.50	13.59	4,160.00	Pierce *et al.* (2010)	Included
Boydstown, Pa.	0.36	8.96	65.13	Pierce *et al.* (2010)	Included
Break Neck Run	0.05	7.00	9.20	Pierce *et al.* (2010)	Included
Buffalo Creek, W.Va.	0.48	14.02	1,420.00	Pierce *et al.* (2010)	Included
Butler, Ariz.	2.38	7.16	810.00	Pierce *et al.* (2010)	Included
Caney Coon Creek, Okla.	1.32	4.57	16.99	Pierce *et al.* (2010)	Included
Castlewood, Okla.	4.23	21.34	3,570.00	Pierce *et al.* (2010)	Included
Castlewood, Colo.	6.17	21.60	3,570.00	Pierce *et al.* (2010)	Included
Cherokee Sandy, Okla.	0.44	5.18	8.50	Pierce *et al.* (2010)	Included
Colonial #4, Pa.	0.04	9.91	14.16	Pierce *et al.* (2010)	Included
Dam Site #8, Miss.	0.87	4.57	48.99	Pierce *et al.* (2010)	Included
Danghe, China	10.70	24.50	2,500.00	Xu & Zhang (2009)	Included
Davis Reservoir. Calif.	58.00	11.58	510.00	Pierce *et al.* (2010)	Included
Dells, USA	13.00	18.30	5,440.00	Xu & Zhang (2009)	Included
DMAD, Utah	19.70	8.80	793.00	Pierce *et al.* (2010)	Included
Dongchuankou, China	27.00	31.00	21,000.00	Xu & Zhang (2009)	Included
Euclides de Cunha, Brazil	13.60	58.22	1,020.00	Pierce *et al.* (2010)	Removed
Field Test 1-1, Norway	0.07	6.10	190.00	Pierce *et al.* (2010)	Included
Field Test 1-2, Norway	0.06	5.90	113.00	Pierce *et al.* (2010)	Included
Field Test 1-3, Norway	0.06	5.90	242.00	Pierce *et al.* (2010)	Included
Field Test 2–2, Norway	0.04	5.00	74.00	Pierce *et al.* (2010)	Included
Field Test 2-3, Norway	0.07	6.00	174.00	Pierce *et al.* (2010)	Included
Field Test 3-3, Norway	0.02	4.30	170.00	Pierce *et al.* (2010)	Included
Frankfurt, Germany	0.35	8.23	79.00	Pierce *et al.* (2010)	Included
Fred Burr, Mont.	0.75	10.20	654.00	Pierce *et al.* (2010)	Included
French Landing, Mich.	3.87	8.53	929.00	Pierce *et al.* (2010)	Included
Frenchman Creek, Mont.	16.00	10.80	1,420.00	Pierce *et al.* (2010)	Included
Frias, Argentina	0.25	15.00	400.00	Xu & Zhang (2009)	Included
Goose Creek, SC	10.60	1.37	565.00	Pierce *et al.* (2010)	Included
Hatchtown, Utah	14.80	16.80	3,080.00	Pierce *et al.* (2010)	Included
Hatfield	12.30	6.80	3,400.00	Pierce *et al.* (2010)	Included
Haymaker, Mont.	0.37	4.88	26.90	Pierce *et al.* (2010)	Included
Hell Hole, Calif.	30.60	35.10	7,360.00	Pierce *et al.* (2010)	Included
Horse Creek #2, Colo.	4.80	12.50	311.49	Pierce *et al.* (2010)	Included
HR Wallingford Test 10, UK	0.00	0.60	0.31	Pierce *et al.* (2010)	Included
HR Wallingford Test 11, UK	0.00	0.60	0.34	Pierce *et al.* (2010)	Included
HR Wallingford Test 12, UK	0.00	0.60	0.53	Pierce *et al.* (2010)	Included

(*continued*)

Table 2 | continued

Site	V (10^6 m^3)	H (m)	Q_p (m^3/s)	Data by	Remark
HR Wallingford Test 14, UK	0.00	0.60	0.28	Pierce *et al.* (2010)	Included
HR Wallingford Test 15, UK	0.00	0.60	0.35	Pierce *et al.* (2010)	Included
HR Wallingford Test 16, UK	0.00	0.60	0.43	Pierce *et al.* (2010)	Included
HR Wallingford Test 17, UK	0.00	0.60	0.61	Pierce *et al.* (2010)	Included
Ireland No. 5, Colo.	0.16	3.81	110.00	Pierce *et al.* (2010)	Included
Johnstown, Pa. (South Fork)	18.90	24.60	8,500.00	Pierce *et al.* (2010)	Included
Kelly Barnes, Ga.	0.78	11.30	680.00	Pierce *et al.* (2010)	Included
Kodaganar, India	12.30	11.50	1,280.00	Xu & Zhang (2009)	Included
Lake Avalon, NM	31.50	13.70	2,320.00	Pierce *et al.* (2010)	Included
Lake Latonka, Pa.	4.09	6.25	290.00	Pierce *et al.* (2010)	Included
Lake Tanglewood, Texas	4.85	16.76	1,351.00	Pierce *et al.* (2010)	Included
Laurel Run, Pa.	0.56	14.10	1,050.00	Pierce *et al.* (2010)	Included
Lawn Lake, Colo.	0.80	6.71	510.00	Pierce *et al.* (2010)	Included
Lijiaju, China	1.14	25.00	2,950.00	Xu & Zhang (2009)	Included
Lily Lake, Colo.	0.09	3.35	71.00	Pierce *et al.* (2010)	Included
Little Deer Creek, Utah	1.36	22.90	1,330.00	Pierce *et al.* (2010)	Included
Little Wewoka, Okla.	0.99	9.45	42.48	Pierce *et al.* (2010)	Included
Liujiatai, China	40.54	35.90	28,000.00	Xu & Zhang (2009)	Included
Lower Latham, Colo.	7.08	5.79	340.00	Pierce *et al.* (2010)	Included
Lower Reservoir, Maine	0.60	9.60	157.44	Pierce *et al.* (2010)	Included
Lower Two Medicine, Mont.	29.60	11.30	1,800.00	Pierce *et al.* (2010)	Included
Mammoth, USA	13.60	21.30	2,520.00	Xu & Zhang (2009)	Included
Martin Cooling Pond Dike, Fla.	136.00	8.53	3,115.00	Pierce *et al.* (2010)	Included
Middle Clear Boggy, Okla.	0.44	4.57	36.81	Pierce *et al.* (2010)	Included
Mill River, Mass.	2.50	13.10	1,645.00	Pierce *et al.* (2010)	Included
Murnion, Mont.	0.32	4.27	17.50	Pierce *et al.* (2010)	Included
Nanaksagar, Ind.	210.00	15.85	9,700.00	Pierce *et al.* (2010)	Included
North Branch, Pa.	0.02	5.49	29.40	Pierce *et al.* (2010)	Included
Oros, Brazil	660.00	35.80	9,630.00	Pierce *et al.* (2010)	Removed
Otto Run	0.01	5.79	60.00	Pierce *et al.* (2010)	Included
Owl Creek, Okla.	0.12	4.88	31.15	Pierce *et al.* (2010)	Included
Peter Green, NH	0.02	3.96	4.42	Pierce *et al.* (2010)	Included
Prospect, Colo.	3.54	1.68	116.00	Pierce *et al.* (2010)	Included
Puddingstone, Calif.	0.62	15.20	480.00	Pierce *et al.* (2010)	Included
Qielinggou, China	0.70	18.00	2,000.00	Xu & Zhang (2009)	Included
Quail Creek, Utah	30.80	16.70	3,110.00	Pierce *et al.* (2010)	Included
Salles Oliveira, Brazil	71.50	38.40	7,200.00	Pierce *et al.* (2010)	Included
Sandy Run, Pa.	0.06	8.53	435.00	Pierce *et al.* (2010)	Included
Schaeffer, Colo.	4.44	30.50	4,500.00	Pierce *et al.* (2010)	Included
Shimantan, China	111.00	26.55	30,000.00	Pierce *et al.* (2010)	Included

(*continued*)

Table 2 | continued

Site	V (10^6 m^3)	H (m)	Q_p (m^3/s)	Data by	Remark
Site Y-30–95, Miss.	0.14	7.47	144.42	Pierce *et al.* (2010)	Included
Site Y-31A-5, Miss.	0.39	9.45	36.98	Pierce *et al.* (2010)	Included
Site Y-36–25, Miss.	0.04	9.75	2.12	Pierce *et al.* (2010)	Included
South Fork Tributary, Pa.	0.00	1.83	122.00	Pierce *et al.* (2010)	Included
Stevens Dam, Mont.	0.08	4.27	5.92	Pierce *et al.* (2010)	Included
Swift, Mont.	37.00	47.85	24,947.00	Pierce *et al.* (2010)	Included
Taum Sauk Reservoir, Mo.	5.39	31.46	7,743.00	Pierce *et al.* (2010)	Included
Teton, Id.	310.00	77.40	65,120.00	Pierce *et al.* (2010)	Removed
Upper Clear Boggy, Okla.	0.86	6.10	70.79	Pierce *et al.* (2010)	Included
Upper Red Rock, Okla.	0.25	4.57	8.50	Pierce *et al.* (2010)	Included
USDA-ARS Test #1, Okla.	0.00	2.29	6.50	Pierce *et al.* (2010)	Included
USDA-ARS Test #3, Okla.	0.00	2.29	1.80	Pierce *et al.* (2010)	Included
USDA-ARS Test #4, Okla.	0.01	1.50	2.30	Pierce *et al.* (2010)	Included
USDA-ARS Test #6, Okla.	0.01	1.50	1.30	Pierce *et al.* (2010)	Included
USDA-ARS Test #7, Okla.	0.00	2.13	4.20	Pierce *et al.* (2010)	Included
Wheatland Reservoir #1, Wyo.	11.50	12.19	566.34	Pierce *et al.* (2010)	Included
Zhugou, China	18.43	23.50	11,200.00	Xu & Zhang (2009)	Included
Zuocun, China	40.00	35.00	23,600.00	Xu & Zhang (2009)	Included

is widely used because it is statistically the best linear unbiased estimator (BLUE). Ordinary Kriging is linear because its estimates are linear combinations of the available data. It is unbiased because it attempts to keep the mean residual to be zero. Finally, it is called best because it tries to minimize the residual variance (Şen 2009).

This technique has three assumptions:

1. The spatial sampling points are representatives of the random variable (ReV) at a set of given locations with measurement values.
2. The ReV is considered as a second-order random field variable with mean, variance, and SV.
3. The mean of ReV is assumed to be constant regionally but unknown.

The KM needs two variables as independants and one variable as dependant. x–y axes in the classical geostatistical approach are replaced by the two independent variables within the system. In this study, H, height of the water behind the dam at failure, and V, volume of the water behind the dam at failure, are taken as inputs respectively, whereas Q_p, peak outflow due to dam breaching, is an output. The contour maps to identify and relate the influence of H and V on Q_p are also plotted.

The main principle of the Kriging (optimum interpolation) methodology is to set up a valid theoretical variogram model (γ) that can interpret and characterize the structural relationships of natural phenomenon. In Ordinary Kriging techniques, the theoretical variogram matrix is given as follows:

$$\gamma = \begin{bmatrix} \gamma(z_1, z_1) & \gamma(z_1, z_2) & \dots & \gamma(z_1, z_n) & 1 \\ \gamma(z_2, z_1) & \gamma(z_2, z_2) & \dots & \gamma(z_2, z_n) & 1 \\ . & . & \dots & . & . \\ . & . & \dots & . & . \\ . & . & \dots & . & . \\ \gamma(z_n, z_1) & \gamma(z_n, z_2) & \dots & \gamma(z_n, z_n) & 1 \\ 1 & 1 & & 1 & 0 \end{bmatrix}$$

B is the semivariance matrix between the estimation site and the other sites around, described as:

$$B = \begin{bmatrix} \gamma(z_E, z_1) \\ \gamma(z_E, z_2) \\ . \\ . \\ \gamma(z_E, z_n) \\ 1 \end{bmatrix}$$

The weights λ are calculated by performing simple matrix algebra as $\lambda = \gamma^{-1}B$, in which, γ^{-1} is the inverse matrix of γ.

$$\lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ . \\ . \\ \lambda_n \\ -\mu \end{bmatrix}$$

The Ordinary Kriging estimate of the ReV at location E can be obtained as $Z_E = M^T \times \lambda$, in which $M =$ measurement vector and $M^T =$ transpose matrix of M.

In general, a Kriging study covers two steps: (i) obtaining theoretical SV (γ) and (ii) solving the prediction problem using the equation, Z_E.

The flow chart of the Kriging methodology is given in Figure 4.

APPLICATION

The Pierce *et al.* (2010) study did not split up the data into two parts as constructing and testing while performing regression analysis. Similarly, it is not required to split up the data in the KM. However, in this study, the data, described in Table 2, are divided into two sets for better determining the efficiency of the model. Out of 94 data, 76 data are employed for constructing and the remaining 18 for testing.

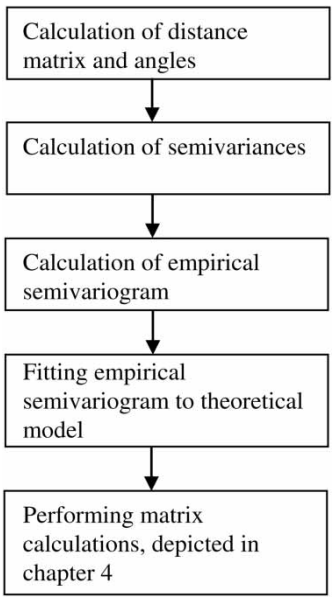

Figure 4 | The flow chart of the Kriging methodology.

The principle of the KM is to establish a valid variogram model that can interpret and characterize the structural relationships of natural phenomenon. By doing so, the variogram model can be used as a simple and reliable statistical tool to interpret the behavior of a random field (Altunkaynak & Wang 2010).

Variogram models can be grouped with and without sill. The sill is, by definition, the maximum value of the variogram that can become stable. The SV has the lowest value at the smallest lag distances and increases with distance and finally levels off at the sill. The sill generally approaches the variance of the field data (Ozger & Sen 2007). The critical horizontal distance where the sill is reached is called range (or the radius of influence). The points, whose separation distances are larger than the range, are not correlated. Another important point in a variogram is the nugget effect. In theory, the SV value at the origin (0 lag distance) should be zero whereas, in practical applications, this criterion cannot be met due to the measurement errors. The nugget effect, which is a jump at the origin, includes discontinuity in the data (Figure 5). The term 'Scale' is the vertical distance, which is the total vertical sill minus the nugget effect. In some cases, the theoretical SV does not reach the sill, as is the case with linear, logarithmic, and power variograms (Şen 2009).

In this study, the KM is only constructed by using 76 case study data, namely the constructing data set. Since the different input variables (H, V) and the output variable (Q_p) have different ranges and units, it is necessary to employ a standardization procedure in order to restrict the data range to the interval of zero-to-one. Hence, both inputs and the output are divided by their maximum such

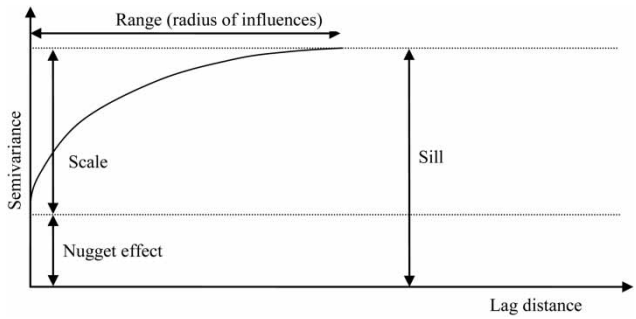

Figure 5 | Variogram elements.

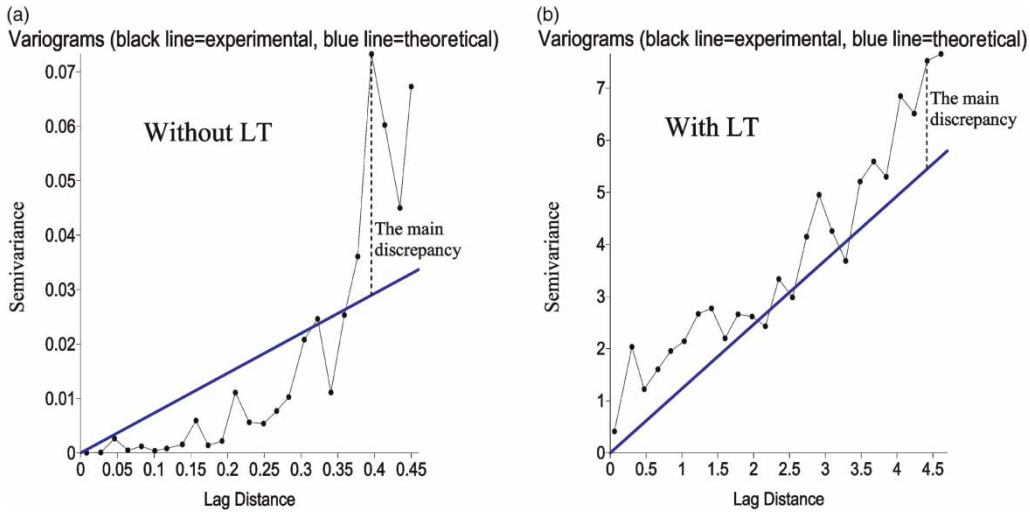

Figure 6 | Semivariograms (a) without logarithmic transformation and (b) with logarithmic transformation.

as H/H_{max}, V/V_{max}, Q_p/Q_{pmax}, in which $H_{max} = 47.85$ m, $V_{max} = 210 \times 10^6$ m^3, and $Q_{max} = 30{,}000$ m^3/s. Then, logarithmic transformation (LT) is applied to each column in order to reduce the discrepancies of the SV. In Figure 6, the SVs created using the same dataset (a) without LT and (b) with LT are demonstrated. It is clearly seen in the figures that LT reduces the main discrepancy between the experimental and the theoretical SVs.

The theoretical variogram is suggested as a linear model (Figure 6(b)) because the SV values increase linearly with lag distances. The maximum lag distance for calculating the theoretical variogram in Figure 6(b) is taken as 4.7, which is approximately 1/3 of the maximum distance for the whole data as a rule of thumb (Şen 2009). The number of lags and lag width are determined as 25 and 0.188, respectively $(4.7 \times 0.188 = 25)$. The least square approximation is then applied for fitting the data. In this study, an isotropic condition, which assumes that the semivariance is only dependent on distance and not on direction, is applied.

The triple diagram (TD) by using Kriging approach is plotted, considering H and V as inputs for axis 'x' and 'y' and Q_p as output for axis 'z', which helps to make optimum interpolation. The design or site engineer can make easy and reliable predictions by reading these maps. TD for Q_p prediction due to the embankment dam failure is depicted in Figure 7 with color scale. Numerical prediction results from the maps can be obtained by entering $\mathrm{Ln}(H/H\,\mathrm{max})$ and $\mathrm{Ln}(V/V\,\mathrm{max})$ on x and y axes, respectively. The desired prediction value can be either read from these maps approximately, or calculated by using Kriging prediction equations. The Q_p, namely the axis z, corresponds Ln (Q_p/Q_{pmax}), which is within the range of 0 to -12.

It is noticeable that there exists a nonlinear relationship between the inputs and output variables in Figure 7. In the diagram, the bended curves show the uncertainties inherited in the system, which are shown in rectangular blocks. These nonlinearities are nearly impossible to reflect by using simple regression approaches.

Logical interpretations can also be made by using Figure 7, such as the following:

1. As V and H increase, Q_p increases.
2. The derivation of Q_p is more dependent on V than H.
3. The uncertainties are mostly situated in the green region of Figure 7, which is within the range of [-6 to -8], namely [75–10 m^3/s], $(30{,}000 \times \exp(-6) = 75,\ 30{,}000 \times \exp(-8) = 10)$.

EVALUATION OF THE PROPOSED MODEL

Pierce *et al.* (2010) employed only the coefficient of determination, R^2, for comparing their models with others in the

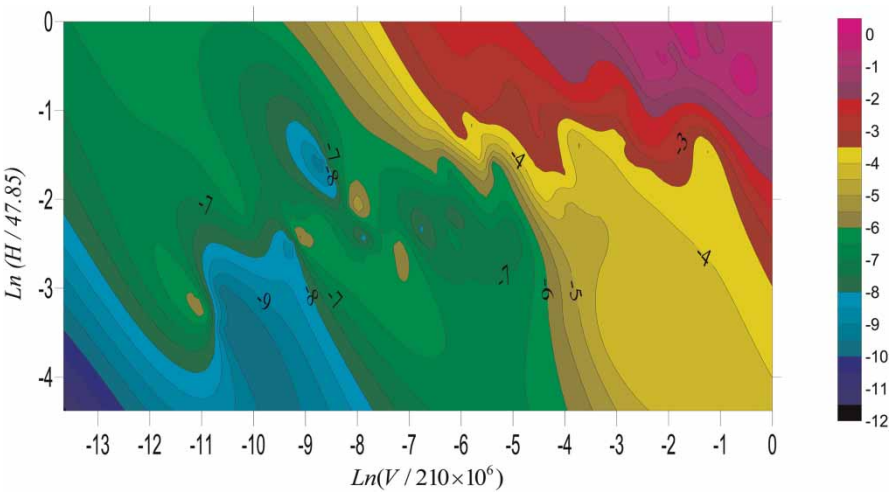

Figure 7 | TD for Q_p prediction. Please refer to the online version of this paper to see this figure in color: http://www.iwaponline.com/jh/toc.htm.

literature. However, in order to come to a sound conclusion, consideration of only R^2 is not sufficient. In this paper, it is demonstrated that other error statistics should be considered together with R^2. In addition, Pierce *et al.* (2010) implied that R^2 of the proposed models of Equations (15) and (16) are 0.84 and 0.85 respectively, using all the data in their database, which is a total of 87. However, these values are calculated as 0.65 and 0.64 respectively, in this research with the same data they employed. It has been discovered that Pierce *et al.* (2010) used LT of the Q_p for both observed and predicted values before or while applying R^2. This caused the R^2 values to be higher than they should be. Şen (1978) demonstrated that transformation in the results can absolutely change the reality and gives biased predictions. In this study, no transformation is applied to the prediction and observed results (Q_p) of the models.

The performance of the proposed model is achieved in terms of graphical representations and five different numerical error criteria (Equations (17)–(20)), which are correlation coefficient (CC), coefficient of efficiency (COE), root mean squared error (RMSE), mean absolute error (MEAE), maximum absolute error (MAAE), in which CC and COE are the measures of how much of the variation and trends in the observed data are predicted by the model, whereas RMSE, MEAE and MAAE indicate quantitative information of the model error with the characteristic that larger errors receive greater attention than smaller ones (Erdik *et al.* 2009).

$$\mathrm{CC} = \frac{\sum_{i=1}^{n}\left(Q_{pm}^{i} - \bar{Q}_{pm}\right) - \left(Q_{pp}^{i} - \bar{Q}_{pp}\right)}{(n-1)s_{Q_{pm}}s_{Q_{pp}}} \tag{17}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left[\left(Q_{pm}^{i}\right) - \left(Q_{pp}^{i}\right)\right]^2}{n}} \tag{18}$$

$$\mathrm{COE} = 1 - \frac{\sum_{i=1}^{n}\left[\left(Q_{pm}^{i}\right) - \left(Q_{pp}^{i}\right)\right]^2}{\sum_{i=1}^{n}\left[\left(Q_{pm}^{i}\right) - \left(\bar{Q}_{pm}\right)\right]^2} \tag{19}$$

$$\mathrm{MEAE} = \frac{1}{n}\sum_{i=1}^{n}\left|Q_{pm}^{i} - \bar{Q}_{pp}\right| \tag{20}$$

where subscripts m and p indicate measured and predicted values respectively, the variable with a bar over it represents the average of that variable and n is total number of data, which are 76, 18 and 94 for constructing, testing and total data, respectively. $s_{Q_{pm}}$ and $s_{Q_{pp}}$ represent standard deviation of Q_{pm} and Q_{pp}, respectively.

The fact that only dispersion is quantified is one of the major drawbacks of CC if it is considered alone (Krause

et al. 2005). A model which systematically over or underpredicts might still result in good CC values close to 1.0 even if all predictions are wrong, this is the reason why five different numerical error criteria are used in this study.

In this study, the proposed model results are compared with 13 RMs developed in the literature. Although Pierce et al. (2010) suggested five different RMs for Q_p prediction (Table 1), only Equations (15) and (16) are employed, because of the fact that both are superior to other developed models by Pierce et al. (2010). The prediction results for 76 constructing dataset are given in Table 3. As is clearly demonstrated, our proposed model yields encouraging results, which supersede all the models developed so far in terms of five error statistics. In the case of 76 constructing data, RMSE, MEAE, CC, MAAE and COE values are 775.789, 245.281, 0.997, 5,096.259 and 0.983 respectively. These are 4,584.951, 1,559.312, 0.774, 21,715.403 and 0.412, and 4,327.416, 1,424.701, 0.843, 21,014.498 and 0.476 for Equations (15) and (16) of Pierce et al. (2010).

In Table 4, the best three models together are also given, considering the five different error statistics. In Table 4, the superiority of the KM is rather obvious in terms of the five different error statistics. Both the US Bureau of Reclamation (1982) and Froehlich (1995) approaches are placed second in terms of two error statistics out of five. Table 4 supports the assertion of Krause et al. (2005) that using only one error criteria as R^2 can result in wrong judgements. The approaches of the US Bureau of Reclamation (1982), Froehlich (1995) and Evans (1986) outnumbered Equation (16) of Pierce et al. (2010) in Table 4, in terms of the 76 constructing dataset.

The scatter diagram of the best three models around a 45° perfect model line is shown in Figure 8. This figure

Table 3 | The prediction performances for 76 constructing data

Models	RMSE	MEAE	CC	MAAE	COE
Costa (1985), Equation (13) in Table 1	4,849.563	1,636.993	0.761	22,761.188	0.342
Evans (1986), Equation (8) in Table 1	4,738.346	2,227.858	0.610	20,246.266	0.372
Froehlich (1995), Equation (14) in Table 1	4,126.519	1,402.969	0.844	21,636.290	0.524
Hagen (1982), Equation (10) in Table 1	8,530.507	4,721.747	0.769	35,140.078	−1.035
Kirkpatrick (1977), Equation (1) in Table 1	4,252.267	1,482.308	0.764	25,263.255	0.494
MacDonald & Langridge-Monopolis (1984) Best-fit, Equation (11) in Table 1	4,511.731	1,662.448	0.760	21,123.732	0.431
MacDonald & Langridge-Monopolis (1984) Envelope, Equation (12) in Table 1	6,065.432	3,669.943	0.760	21,956.853	−0.029
Pierce et al. (2010), Equation (15) in Table 1	4,584.951	1,559.312	0.774	21,715.403	0.412
Pierce et al. (2010), Equation (16) in Table 1	4,327.416	1,424.701	0.843	21,014.498	0.476
Singh & Snorrason (1982), Equation (4) in Table 1	3,953.305	1,583.617	0.784	23,414.530	0.563
Singh & Snorrason (1984), Equation (7) in Table 1	4,768.883	2,138.153	0.616	21,314.278	0.364
Soil Conservation Service (SCS) (1981) for dams >31.4 m, Equation (2) in Table 1	3,862.289	1,626.312	0.785	22,844.706	0.583
US Bureau of Reclamation (1982), Equation (3) in Table 1	3,740.591	1,697.128	0.785	21,767.102	0.609
Proposed model	775.789	245.281	0.997	5,096.259	0.983

Table 4 | Best three models for 76 constructing dataset

	RMSE	MEAE	CC	MAAE	COE
1	Proposed model	Proposed model	Proposed model	Proposed model	Proposed model
2	US Bureau of Reclamation (1982)	Froehlich (1995)	Froehlich (1995)	Evans (1986)	US Bureau of Reclamation (1982)
3	Soil Conservation Service (SCS) (1981) for dams >31.4 m	Pierce et al. (2010), Equation (16) in Table 1	Pierce et al. (2010), Equation (16) in Table 1	Pierce et al. (2010), Equation (16) in Table 1	Soil Conservation Service (SCS) (1981) for dams >31.4 m

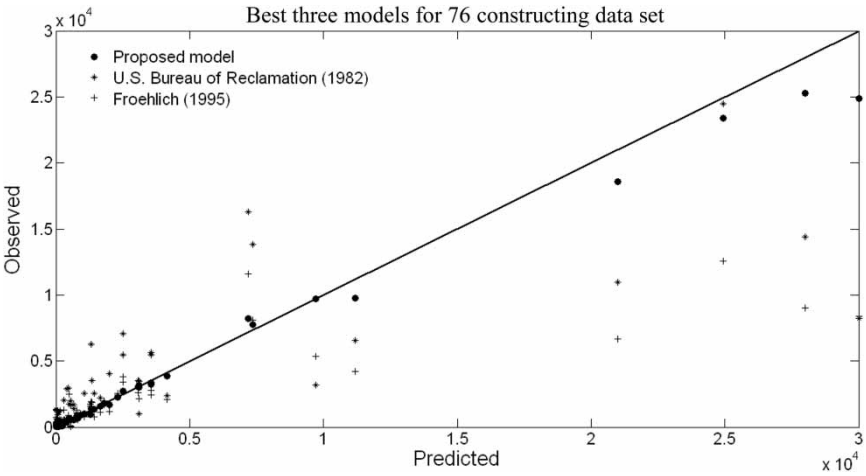

Figure 8 | Scatter diagrams around perfect line for best three models using 76 testing dataset.

shows that the proposed model follows the perfect line closely, but this is not valid for the others.

The prediction results for the 18 testing dataset are given in Table 5. It is clearly demonstrated in this table that the proposed model has the smallest RMSE, MEAE, MAAE and the highest COE compared to the others. In terms of CC, the prediction capability of Equations (15) and (16) of Pierce *et al.* (2010) is slightly higher than that of the proposed model. These values are 0.9375 and 0.9559 for Equations (15) and (16), whereas it is 0.9367 for the proposed model.

In Table 6, the best models are demonstrated, considering five different error statistics. As for the 18 testing data, our proposed model gives the best results according to four out of the five error statistics. Although our proposed model is third in CC, the comparative results are more or less the same with Equations (15) and (16) of Pierce *et al.* (2010). The formula of the US Bureau of Reclamation

Table 5 | The prediction performances for 18 testing data

Models	RMSE	MEAE	CC	MAAE	COE
Costa (1985), Equation (13) in Table 1	4,859.3587	2,422.2860	0.8987	18,304.6056	0.2447
Evans (1986), Equation (8) in Table 1	4,172.4861	2,074.7575	0.8190	15,901.1778	0.4431
Froehlich (1995), Equation (14) in Table 1	3,848.9408	1,768.1178	0.9359	14,881.5202	0.5262
Hagen (1982), Equation (10) in Table 1	5,253.6103	3,675.4076	0.9145	13,188.8286	0.1172
Kirkpatrick (1977), Equation (1) in Table 1	3,767.5381	1,823.4426	0.8427	14,212.3689	0.5460
MacDonald & Langridge-Monopolis (1984) Best-fit, Equation (11) in Table 1	4,374.1339	2,121.4600	0.8963	16,832.7759	0.3880
MacDonald & Langridge-Monopolis (1984) Envelope, Equation (12) in Table 1	3,731.3045	2,694.0697	0.8960	8,984.1135	0.5547
Pierce *et al.* (2010), Equation (15) in Table 1	4,782.2052	2,501.6770	0.9375	17,461.4435	0.2685
Pierce *et al.* (2010), Equation (16) in Table 1	4,375.2437	2,251.8585	0.9559	16,121.8140	0.3877
Singh & Snorrason (1982), Equation (4) in Table 1	3,373.2203	1,725.4516	0.8223	12,498.1999	0.6360
Singh & Snorrason (1984), Equation (7) in Table 1	4,424.7680	2,189.1137	0.8056	16,956.2825	0.3738
Soil Conservation Service (SCS) (1981) for dams >31.4 m, Equation (2) in Table 1	3,278.8656	1,789.4674	0.8207	11,670.1778	0.6561
US Bureau of Reclamation (1982), Equation (3) in Table 1	3,204.0875	1,923.7027	0.8207	9,873.5179	0.6716
Proposed model	2,173.9308	1,238.6242	0.9367	7,073.1954	0.8488

Table 6 | Best three models for 18 testing dataset

	RMSE	MEAE	CC	MAAE	COE
1	Proposed model	Proposed model	Pierce *et al.* (2010), Equation (16) in Table 1	Proposed model	Proposed model
2	US Bureau of Reclamation (1982)	Singh & Snorrason (1982)	Pierce *et al.* (2010), Equation (15) in Table 1	MacDonald & Langridge-Monopolis (1984) Envelope	US Bureau of Reclamation (1982)
3	Soil Conservation Service (SCS) (1981) for dams >31.4 m	Froehlich (1995)	Proposed model	US Bureau of Reclamation (1982)	Soil Conservation Service (SCS) (1981) for dams >31.4 m

(1982) remains in second place twice in terms of two error criteria out of five. The equations of Singh & Snorrason (1982), Equation (15) of Pierce *et al.* (2010), and finally MacDonald & Langridge-Monopolis (1984) Envelope equations appeared once in the second best place.

Scatter diagrams of best models for the 18 testing dataset are shown in Figure 9. It is clearly visible that the proposed model yields more realistic results than the others do.

The prediction results of Q_p for a total of 94 data are given in Table 7. As clearly noticed, the deviation decreases in the results of the proposed model, as compared to the others. In the case of the total 94 data, RMSE, MEAE, CC, MAAE and COE values are 1,179.6501, 435.4958, 0.9815, 7,073.1954 and 0.9604, respectively. However, these are 4,623.3748, 1,739.7653, 0.7888, 21,715.4027 and 0.3915, and 4,336.6153, 1,583.0926, 0.8585, 21,014.4980 and 0.4646 for Equations (15) and (16) of Pierce *et al.* (2010).

Best models for the 94 overall dataset are given in Table 8. The proposed model is on top concerning the five different error statistics. In addition, it is interestingly seen that the US Bureau of Reclamation (1982) methodology achieved second place three times compared to others. Similarly, Soil Conservation Service (SCS) (1981) for dams >31.4 m is also seen at the third place three times.

The scatter diagram of the best three models around a 45° perfect model line including all 94 data is presented in Figure 10. It is clearly noticeable that deviations of the proposed model are much smaller than those of the others.

The best models are also checked for relative error, which is thought to be the best criterion for the problems

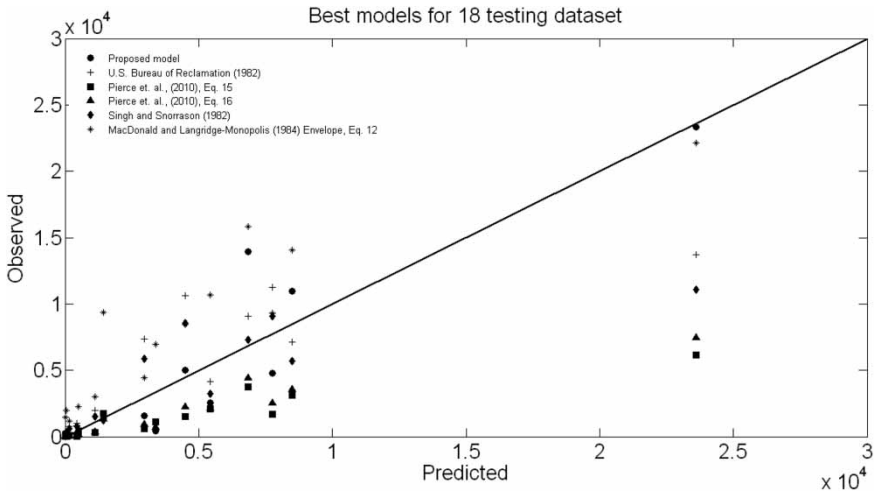

Figure 9 | Scatter diagrams around perfect line for best three models using 18 testing dataset.

Table 7 | The prediction performances for total 94 data

Models	RMSE	MEAE	CC	MAAE	COE
Costa (1985), Equation (13) in Table 1	4,851.4402	1,787.3685	0.7792	22,761.1881	0.3300
Evans (1986), Equation (8) in Table 1	4,635.3405	2,198.5408	0.6257	20,246.2663	0.3883
Froehlich (1995), Equation (14) in Table 1	4,074.8295	1,472.8915	0.8604	21,636.2897	0.5273
Hagen (1982), Equation (10) in Table 1	8,007.5054	4,521.3842	0.7862	35,140.0777	−0.8253
Kirkpatrick (1977), Equation (1) in Table 1	4,163.8169	1,547.6318	0.7791	25,263.2553	0.5064
MacDonald & Langridge-Monopolis (1984) Best-fit, Equation (11) in Table 1	4,485.7093	1,750.3435	0.7779	21,123.7315	0.4272
MacDonald & Langridge-Monopolis (1984) Envelope, Equation (12) in Table 1	5,693.0409	3,483.0735	0.7778	21,956.8529	0.0773
Pierce *et al.* (2010), Equation (15) in Table 1	4,623.3748	1,739.7653	0.7888	21,715.4027	0.3915
Pierce *et al.* (2010), Equation (16) in Table 1	4,336.6153	1,583.0926	0.8585	21,014.4980	0.4646
Singh & Snorrason (1982), Equation (4) in Table 1	3,848.9982	1,610.7772	0.7919	23,414.5304	0.5783
Singh & Snorrason (1984), Equation (7) in Table 1	4,704.9370	2,147.9114	0.6327	21,314.2779	0.3698
Soil Conservation Service (SCS) (1981) for dams >31.4 m, Equation (2) in Table 1	3,757.5886	1,657.5542	0.7920	22,844.7064	0.5981
US Bureau of Reclamation (1982), Equation (3) in Table 1	3,643.9756	1,740.5148	0.7920	21,767.1019	0.6220
Proposed model	1,179.6501	435.4958	0.9815	7,073.1954	0.9604

Table 8 | Best models for total 94 dataset

	RMSE	MEAE	CC	MAAE	COE
1	Proposed model	Proposed model	Proposed model	Proposed model	Proposed model
2	US Bureau of Reclamation (1982)	Froehlich (1995)	US Bureau of Reclamation (1982)	Pierce *et al.* (2010), Equation (16) in Table 1	US Bureau of Reclamation (1982)
3	Soil Conservation Service (SCS) (1981) for dams >31.4 m	Kirkpatrick (1977)	Soil Conservation Service (SCS) (1981) for dams >31.4 m	MacDonald & Langridge- Monopolis (1984) Best-fit	Soil Conservation Service (SCS) (1981) for dams >31.4 m

whose purpose is to check how well the predictions converge to measurements (Erdik & Savci 2008).

Comparative results in Figure 11 demonstrate that the proposed model is in better compliance with the observed peak discharge (Q_p) than the other model predictions. The average relative error of the proposed model is 25%, whereas these are 153, 396 and 1,574% for Pierce *et al.* (2010) Equation (16), Froehlich (1995) and US Bureau of Reclamation (1982), respectively.

The superiority of the proposed model becomes rather obvious in the prediction of Q_p, after all the aforementioned numerical and graphical comparisons. Besides, the Kriging approach is also very fast in application. Overall, these results suggest that the developed model can be utilized as an efficient tool for better Q_p prediction than others.

CONCLUSION

The application of the KM for development of the more robust and efficient predictive model for determination of peak discharge due to embankment dam breach is presented in this paper. A new database is extended and employed

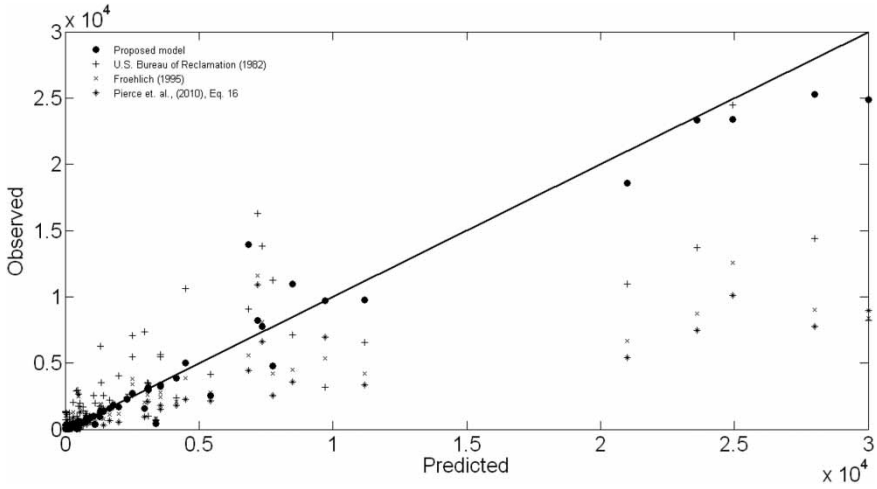

Figure 10 | Scatter diagrams around perfect line for best three models using overall 94 dataset.

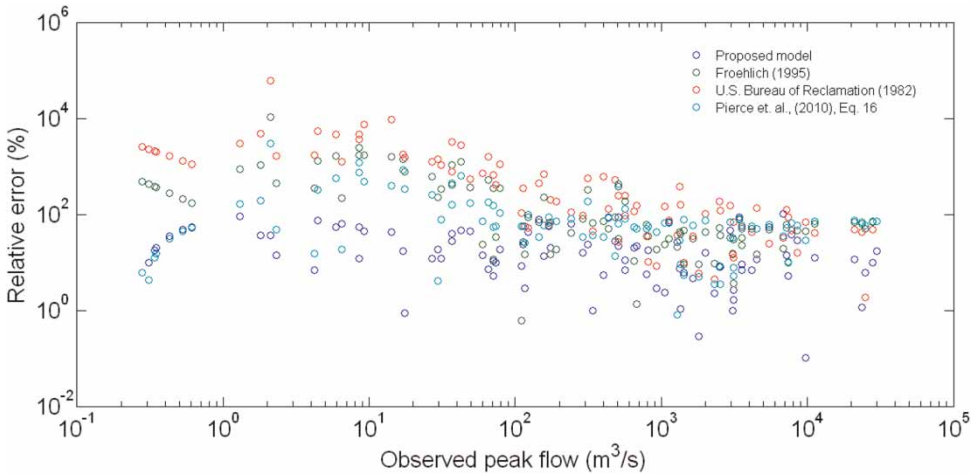

Figure 11 | Relative error of best models for total 94 data.

after performing an outlier test. The data are divided into two parts for development of the models, 76 data for constructing and 18 data for prediction. Five different error statistics are employed for determining the prediction capability of the developed model as CC, COE, RMSE, MEAE and MAAE, according to which the proposed model outnumbered all the previous models which have been developed since 1977. A contour map is also established from the constructing data. Practitioners can easily benefit from this map by inserting H and V on the axis, without fulfilling any deterministic calculations. It is also shown in this study that R^2 should not be applied solely, to reach a conclusion about which model is superior to the other. Other statistical error criteria should be employed with R^2 to get more satisfactory results. Depending on the availability of the recorded data, the proposed model can be applied to gain more reliable estimation than the previous ones.

REFERENCES

Altunkaynak, A. 2005 Significant wave height prediction by using a spatial model. *Ocean Eng.* **32** (8–9), 924–936.

Altunkaynak, A. 2009 Streamflow estimation using optimal regional dependency function. *Hydrol. Process* **23**, 3525–3533.

Altunkaynak, A. & Wang, K. H. 2010 Triple diagram models for prediction of suspended solid concentration in Lake Okeechobee, Florida. *J. Hydrol.* **387**, 165–175.

Altunkaynak, A., Özger, M. & Şen, Z. 2003 Triple diagram model of level fluctuations in Lake Van, Turkey. *Hydrol. Earth Syst. Sci.* **7** (2), 235–244.

ASCE/EWRI Task Committee on Dam/Levee Breaching 2011 Earthen embankment breaching. *J. Hydraul. Eng., ASCE* **137** (12), 1549–1564.

Costa, J. E. 1985 Floods from Dam Failures. US Geological Survey Open-File Rep. No. 85–560, US Geological Survey, Denver, Colo.

D'Eliso, C. 2007 Breaching of Sea Dikes Initiated by wave Overtopping: A Tiered and Modular Modeling Approach. PhD Dissertation, University of Braunschweig Germany and University of Florence, Italy.

Erdik, T. 2009 Fuzzy logic approach to conventional rubble mound structures design. *Expert Syst. Appl.* **36**, 4162–4170.

Erdik, T. & Savci, M. 2008 Takagi-Sugeno fuzzy approach in rock armored slopes for 2% wave runup estimation. *Coast. Eng. J.* **50** (2), 161–177.

Erdik, T., Savci, M. E. & Şen, Z. 2009 Artificial neural networks for predicting maximum wave runup on rubble mound structures. *Expert Syst. Appl.* **36**, 6403–6408.

Evans, S. G. 1986 The maximum discharge of outburst floods caused by the breaching of man-made and natural dams. *Can. Geotech. J.* **23** (3), 385–387.

Froehlich, D. C. 1995 Peak outflow from breached embankment dam. *J. Water Resour. Plann. Manage.* **121** (1), 90–97.

Hagen, V. K. 1982 Re-evaluation of design floods and dam safety. *Proc. 14th Congress of Int. Commission on Large Dams*, Q52, R29, Rio de Janeiro.

Huisman, P., Cramer, W. & van Ee, G. 1998 *Water in the Netherlands*. Netherlands Hydrological Society, Delft, Netherlands.

ICOLD 1997 *Benefits and Concerns about Dams – An Argumentaire*. Committee on Public Awareness and Education, Montreal, Canada.

Kirkpatrick, G. W. 1977 Evaluation guidelines for spillway adequacy. *Proc. Engineering Foundation Conf. ASCE*, Reston, VA, pp. 395–414.

Krause, P., Boyle, D. P. & Base, F. 2005 Comparison of different efficiency criteria for hydrological model assessment. *Adv. Geosci.* **5**, 89–97.

MacDonald, T. C. & Langridge-Monopolis, J. 1984 Breaching characteristics of dam failures. *J. Hydraul. Eng.* **110** (5), 567–586.

Matheron, G. 1963 Principles of geostatistics. *Econ. Geol.* **58**, 1246–1266.

Ozger, M. & Sen, Z. 2007 Triple diagram method for the prediction of wave height and period. *Ocean Eng.* **34** (7), 1060–1068.

Pan, J. Z. 2000 *Merits of Dams*. Tsinghua University Press, Beijing (in Chinese).

Pierce, M. W., Thornton, C. I. & Abt, S. R. 2010 Predicting peak outflow from breached embankment dams. *J. Hydrol. Eng.* **15** (5), 338–349.

Şen, Z. 1978 Autorun analysis of hydrologic time series. *J. Hydrol.* **36** (1–2), 75–85.

Şen, Z. 2009 *Spatial Modeling Principles in Earth Sciences*. Springer-Verlag, New York.

Singh, V. P. & Scarlatos, P. D. 1988 Analysis of gradual earth-dam failure. *J. Hydraul. Eng.* **114** (1), 21–42.

Singh, K. P. & Snorrason, A. 1982 Sensitivity of outflow peaks and flood stages to the selection of dam breach parameters and simulation models. State Water Survey (SWS) Contract Rep. No. 288, Illinois Dept. of Energy and Natural Resources, SWS Div., Surface Water at the Univ. of Illinois.

Singh, K. P. & Snorrason, A. 1984 Sensitivity of outflow peaks and flood stages to the selection of dam breach parameters and simulation models. *J. Hydrol.* **68**, 295–310.

Sirdaş, S. & Şen, Z. 2003 Spatio-temporal drought analysis in the Trakya region, Turkey. *Hydrol. Sci. J.* **48** (5), 809–820.

Soil Conservation Service (SCS) 1981 Simplified dam-breach routing procedure. Technical Release Rep. No. 66 (Rev. 1), US Department of Agriculture, Washington, DC.

Teegavarapu, R. S. V. 2009 Estimation of missing precipitation records integrating surface interpolation techniques and spatio-temporal association rules. *J. Hydroinf.* **11** (2), 133–146.

US Bureau of Reclamation 1982 Guidelines for defining inundated areas downstream from Bureau of Reclamation dams. Reclamation Planning Instruction Rep. No. 82–11, Assistant Commissioner-Engineering and Research, Denver, Colorado.

Uyumaz, A., Altunkaynak, A. & Özger, M. 2006 Fuzzy logic model for equilibrium scour downstream of a dam's vertical gate. *J. Hydraul. Eng.* **132** (10), 1069–1075.

Wahl, T. L. 1998 Prediction of embankment dam breach parameters: A literature review and needs assessment. Dam Safety Rep. No. DSO-98- 004, Bureau of Reclamation, US Dept of the Interior, Denver.

Wahl, T. L. 2004 Uncertainty of predictions of embankment dam breach parameters. *J. Hydraul. Eng.* **130** (5), 389–397.

Xu, Y. & Zhang, M. 2009 Breaching parameters for earth and rockfill dams. *J. Geotech. Geoenviron. Eng.* **135** (12), 1957–1970.

First received 2 November 2012; accepted in revised form 11 March 2013. Available online 10 April 2013