Implementation of artificial neural network technique in the simulation of dam breach hydrograph

Vahid Nourani, Habib Hakimzadeh and Alireza Babaeyan Amini

ABSTRACT

In the present study, two artificial neural networks were developed to simulate outflow hydrograph from earthen dam breach. The required data for the modelling were collected from literature, laboratory experiments and a physically based model (i.e. BREACH). For the laboratory modelling, five different materials were used for the construction of different dams of various sizes, and the process of the breach was recorded by two video cameras to record the breach growth as well as the output hydrograph. The genetic algorithm was also applied to divide the data into three statistically similar sub-sets for training, validation and test purposes. The obtained results demonstrate that the results of the artificial neural network (ANN) method are in good agreement with the observed values, and this method produces better results than existing classical methods. Also, the experiments show when cohesive strength is larger, the breach process becomes slower, and the peak outflow and the final width and depth of breach become smaller. Moreover, when the friction angle is larger, the breach process becomes slower, and the peak outflow and the final width and depth of breach become smaller. However, the rate of breach formation is particularly dependent upon the soil properties.

Key words | artificial intelligence, artificial neural network, dam break, earthen dam breach, genetic algorithm, outflow hydrograph

Vahid Nourani (corresponding author)
St. Anthony Falls Lab. and NCED,
Dept. of Water Eng.,
Faculty of Civil Eng.,
Univ. of Tabriz,
Iran
and
St. Anthony Falls Lab. and NCED,
Dept. of Civil Eng.,
Univ. of Minnesota,
USA
E-mail: *vnourani@yahoo.com*; *nourani@tabrizu.ac.ir*; *vnourani@umn.edu*

Habib Hakimzadeh
Alireza Babaeyan Amini
Faculty of Civil Engineering,
Sahand University of Technology,
Tabriz,
Iran

ABBREVIATIONS AND NOTATIONS

H_w	height of water at the reservoir
Q_p	peak outflow
V_w	reservoir water volume
g	acceleration of gravity
ΔH_c	erosion rate
P_o	perimeter of the breach
L	length of dam crest
L_b	length of the breach channel
P_{or}	porosity of the breach material
Δt	time step
Q_s	sediment-transport rate
Q_i	inflow to reservoir
Q_b	breach outflow
Q_{sp}	spillway flow
Q_o	crest overflow
S_a	surface area at reservoir elevation
F_w	force due to water pressure
F_{sb}	shear force along the bottom of breach
F_{ss}	shear force along the side of breach
F_{cb}	force due to cohesion along the bottom of breach
F_{cs}	force due to cohesion along the side of breach
RMSE	root mean squared error
E	Nash–Sutcliffe efficiency coefficient (or determination coefficient)
Q_i	observed peak outflow
$\hat{Q}_i$	simulated peak outflow
$\bar{Q}$	mean value of observed peak outflow
y_i	normalized parameter
x_{max}	maximum observed parameter
x_{min}	minimum observed parameter
x_i	observed parameter
c	cohesive strength
ϕ	internal friction angle
D_{50}	median diameter of soil material

doi: 10.2166/hydro.2011.114

γ unit weight of soil
z_D downstream slopes of the dam
z_u upstream slopes of the dam
B_t^b bottom width of the breach at time t
H_t^c height of the breach at time t
B_t^t top width of the breach at time t
Q_{t-1} outflow discharge at time $t-1$
w weight in the hidden layer in neural network
w_{j0} bias for the jth hidden neuron
f activation function
$\hat{y}_k$ computed output variables
N_N number of the neurons in the input layers
M_N number of the neurons in the hidden layers

INTRODUCTION

Dams have been of great importance in water resources development and hydropower generation. Earth dams are constructed with local materials and require relatively low budgets, and are therefore more widespread than other types. Dam failure, as a huge destructive disaster, can cause much damage to the environment, facilities and properties in the downstream. It can also result in the loss of human and/or animal life. Therefore, the analysis of the dam failure phenomenon should be considered in the design and construction of dams and, in this way, the estimation of dam breach outflow hydrograph is the first and most important task.

In technical literature, dam failure analysis is divided into two main parts. First, analysis of the failure and breach process and, second, flow routing and extraction of the inundation maps due to dam break. The failure of a concrete dam usually occurs instantaneously; however, on the other hand, earthen dams gradually fail due to the continuous increase of an initial breach. In this case, the output hydrograph is related to the erosion condition that takes place inside the breach. Thus, the solution needs further investigation to study safety management of the dams. Different reasons may lead to failure of an earth dam such as flow overtopping, piping discharge, land slide, war etc.; but the first two are the most common and important causes. Although some dam failures due to flow overtopping have been already monitored and/or analysed by hydraulic engineers, there is still a need to have further insights into the phenomenon and to develop more robust simulation models. As in an earth dam failure process the piping mode is changed to an overtopping mode after some time, the particular concern in this study is focused on dam failure due to flow overtopping.

Thus far, a number of models have been developed to simulate the earth dam breach process. These can be categorized into two broad lumped (or black-box) and distributed (or physically based) models, depending on the complexities involved in the process, although there are semi-distributed (semi-physically based) models that are neither fully lumped nor fully distributed (Nourani & Mano 2007). According to Wahl (1998, 2004), current analysis methods are grouped into four categories: I. physically based methods, II. predictor equations, III. comparative analysis and IV. parametric models; the second and third groups can be placed into the lumped models class, and IV may be classified as a semi-distributed (or semi-physically based) model. The lumped (or black-box) models establish relationships between input and output variables regarding the problem (e.g. dam breach) and without considering the complex physical laws governing the natural process. Some researchers have developed formulas based on black-box models and historical dam breach data that have enabled them to predict breach parameters, such as time of breach formation, breach geometry and peak outflow from breach. For instance, Johnson & Illes (1976) were the first to predict failure shapes for earth, gravity and arch concrete dams. Also, MacDonald & Monopolis (1984), Singh & Snorrason (1984), FERC (1987), Von Thun & Gillette (1990) and Froehlich (1995) developed formulas to predict the time of breach formation and breach geometry. Kirkpatrick (1977), the US Bureau of Reclamation (1982), Costa (1985), Froehlich (1995) and Webby (1996) have conducted studies to determine the peak outflow as a function of dam height and reservoir storage volume. The presented formulas in the aforementioned studies have been developed based on historical data from some failed dams. On the other hand, a physically based model is based on numerical simulation, which depends on breach formation. Such a model attempts to represent the known physical process in a simplified manner using linear/nonlinear mathematical formulations.

While distributed or semi-distributed models have proved their importance in understanding the relevant processes, their implementation and calibration present various difficulties. In the dam breach phenomenon, physical methods employ the principles of hydraulics, sediment transport and soil mechanics. Most of the physically based dam breach models consider that the dam maintains constant height through the breach and perform the dam breach analysis mostly in two dimensions. Such models often utilize the soil transportation capacity to calculate the erosion of soil on the basis of a uniform steady-state condition of flow. However, in the real world, a dam breach's outflow has a rapidly varied and non-uniform unsteady condition. Therefore, estimating the erosion rate using the soil transportation capacity is not appropriate (Wang & Bowles 2006). In the last 40 years, particularly since the 1980s, several physically based models have been developed for the simulation of breach growth in embankments, e.g. Fread (1988), Singh & Quiroga (1988), Singh & Scarlatos (1988), Visser (1998), Mohamed *et al.* (2002), Franca & Almeida (2004), Wang & Bowles (2006) and Zhu *et al.* (2006).

Semi-physical models, such as DAMBRK (Fread 1984) and FLDWAV (Fread 1993), take a more rigorous approach, though not as rigorous as physical models. Such models use historical monitored breach cases to predict the failure time and final geometry of the breach. Afterwards, they linearly interpolate the breach development over time and calculate breach outflow using the principles of hydraulics. The new Macchione's model (2008), as a new generation of DAMBRK and FLDWAV, uses the same semi-physical approach. Currently, however, sensitivity analysis indicates that the internal friction angle and compaction index are the major factors affecting nonlinear behaviour (Morris *et al.* 2008). In any case, the calibration and validation of these models remains problematic due to the lack of good empirical data. To improve the understanding of the process of embankment breaching and to collect data for the calibration and validation of a new embankment breaching mathematical model, laboratory tests should be conducted to verify mathematical analysis (Franca & Almeida 2004; Zhu *et al.* 2006; Chinnarasri *et al.* 2007).

Although the mentioned black-box and physical models represent the state-of-the-art approaches in the prediction of earth dams breach formation, the capability of the models to predict the outflow hydrographs is not sufficient, and most models can only predict the peak outflow (Faeh 2007). Dam breach is a highly nonlinear, time-varying and spatially distributed process and, in modelling such a complex process, the black-box models may be an alternative method when a suitable data set exists. In the black-box approach, models are used to map directly the inputs and outputs without detailed consideration of the internal structure of the physical process. The artificial neural network (ANN), as such a black-box approach, is capable of identifying complex nonlinear relationships between input and output data sets and is used widely to model countless engineering problems (Nourani & Kalantari 2010). ANNs can be used efficiently in situations where explicit knowledge of complex internal processes is not available. The application of ANNs in hydrology began in the early 1990s, and a state-of-the-art review of ANN applications in hydrology can be found in the ASCE Task Committee Report (2000a). Some applications of ANNs in water resources include precipitation–runoff modelling (Rajurkar *et al.* 2002); stream flow forecasting (Moradkhani *et al.* 2004); river stage forecasting (Liong *et al.* 2000); modelling of rainfall-runoff relationships (Hettiarachchi *et al.* 2005); groundwater level forecasting (Nourani *et al.* 2008); prediction of wave height (Zamani *et al.* 2009) and predicting watershed precipitation (Nourani *et al.* 2009a). Some attempts have also been made to predict the hydrograph of watershed runoff using different input parameters. Muttiah *et al.* (1997) used information on the drainage basin, elevation, average slope and average annual precipitation to predict 2-year peak discharge from a watershed. Carriere *et al.* (1996) used an ANN with a recurrent back-propagation algorithm to generate a runoff hydrograph using a virtual runoff hydrograph system. They employed rainfall intensity and duration, catchment slope and catchment cover to estimate runoff hydrographs. Smith & Eli (1995) used a back-propagation ANN to predict the peak discharge and the time of peak resulting from a single rainfall event. They used a synthetic watershed to generate runoff from stochastically generated rainfall patterns. However, the selection of appropriate input parameters that allow an ANN to successfully produce the desired output is a complex task. Good understanding of the hydrologic

system under consideration is an important prerequisite for successful application of ANNs, and a physical understanding of the process being studied leads to a better choice of input variables.

Taking inspiration from the application of ANNs in the simulation of watershed hydrographs, we present two ANN-based models to simulate the earthen dam breach process in this paper. The first model, similar to other classical black-box models, only predicts the peak outflow through the breach using some explicitly available parameters (e.g. reservoir volume, height of dam, etc.). The second model employs the distinct concept of the ANN to simulate the instant outflow hydrograph from the breach as a function of time. An ANN requires sufficient data to be trained and verified. Three data sources have been used in the current study: (1) historical dam breaches, (2) laboratory experiments conducted in the hydraulics laboratory of Sahand University of Technology, Tabriz, Iran, and (3) a physically based numerical model (i.e. BREACH).

Earth dam breach process

As mentioned above, the failure of an earth dam can be the result of various causes, such as flow overtopping and piping discharge. Dam failure due to flow overtopping has occurred frequently in the past (Tingsanchali & Chinnarasri 2001). Also, the significant differences between overtopping and piping only differ at the beginning of breaching before a breach channel is finally developed (Visser 1998). Therefore, a particular concern in this study is to focus on dam failure due to flow overtopping.

The earth dam breach process due to overtopping is divided into three stages (Wang & Bowles 2006):

1. Erosion stage 1 (Figure 1(a) and (b)) lasts from the start of erosion until the top of the dam is completely eroded. At this stage, outflows over the crest and through a breach can be approximated using the broad-crested weir equation.
2. Erosion stage 2 (Figure 1(b) and (c)) lasts from when the top of the dam is completely eroded until the entire dam is eroded. This stage involves two phenomena: gradual soil erosion due to water flow and the slope stability of the bottom and sides of the breach. Researchers have considered many different equations to calculate the erosion rate and soil transportation capacity, but most equations have been developed and calibrated for stream and river flow fields so that their use in the breach flow field may produce some errors in the results.
3. Erosion stage 3 (Figure 1(d)) lasts from when the dam is eroded until the end of the dam breach process. Stage 3 is similar to stage 2, but in stage 3 there is no instability at the non-erodible bottom of the breach. Due to the complexity of the different stages of dam breach, physically based modelling of the process is sometimes not straightforward in terms of programming and theorizing the overall breach process when sufficient field data are not available.

On the other hand, some researchers have conducted studies to develop experimental formulas using historical breach data to determine breach characteristics, especially the peak outflow, as a function of dam height and reservoir storage volume.

Costa (1985) presented a method on the basis of a regression analysis. It can be applied to both embankment and concrete dams because the 31 breach cases used to develop the model were selected from both embankment and concrete dam breaches. Costa's model estimates the peak outflow (Q_p, m^3/s) from the breach as:

$$Q_p = 0.763(V_w H_w)^{0.42} \quad (1)$$

where H_w is the height of water directly at the reservoir before breach, measured from the bottom of the final breach (m), and V_w is the reservoir water volume at the time of failure (m^3).

The Froehlich (1995) equation was derived by fitting a multiple linear regression on 22 dams in which discharge data were available:

$$Q_p = 0.607 V_w^{0.295} H_w^{1.24} \quad (2)$$

The equation shows good agreement with the measured peak flows over the entire range.

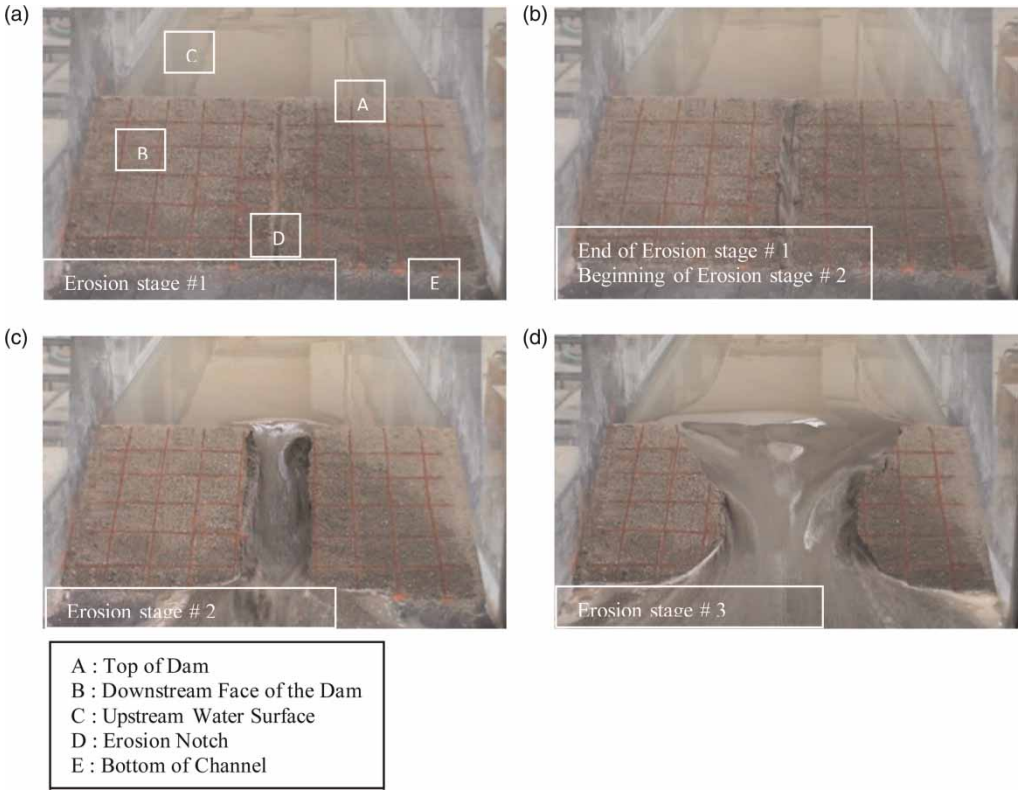

Figure 1 | Earth dam breach process.

In discussing Froehlich's study, Webby (1996) applied dimensional analysis to Froehlich's data set to develop an equation to estimate peak outflow. The equation in dimensional form is expressed as follows:

$$Q_p = 0.0443 g^{0.5} V_w^{0.367} H_w^{1.40} \quad (3)$$

where g is the acceleration of gravity (m/s^2).

The most important issue regarding empirical modelling is calibration. Calibration is often performed using a specific set of data from laboratory and/or field tests or, rarely, from historical failed dams with incomplete characteristics. There is no guarantee that the actual breach can be predicted by the models, which have not been calibrated with reliable data. The most appropriate model is calibrated with an historical failed dam, but available data for failed dams are rare and incomplete. To overcome the problem, this study used three categories of data sources for calibration and modelling, including experimental (laboratory) data, data from a physically based model and historical breach data. To gather experimental data, a number of tests were conducted in the hydraulics laboratory of Sahand University of Technology, Iran. Available historical breach data were derived from literature, and the rest of the data were obtained from the BREACH model (Fread 1988).

BREACH is a physically based model that has been developed applying hydraulics, sediment-transport formulas and soil mechanics. Input parameters are geometric and soil properties of a dam and reservoir properties (storage volume, spillway characteristics and the time-dependent reservoir inflow rate). The model predicts breach characteristics (size, shape, time of formation) and outflow hydrographs. The sequence of computation in the model is iterative, as the flow into the breach is dependent on the bottom elevation of the breach and its width; the breach properties are dependent on the sediment-transport capacity of the breach flow, and the transport capacity is dependent

on the breach size and flow. The computational algorithm of the model is briefly described below.

At each time step of the breach process, the erosion rate (ΔH_c) is computed as follows (Fread 1988):

1. ΔH is estimated.
2. Reservoir elevation is computed.
3. Q_i, Q_{sp} and Q_o are computed with calculated reservoir elevation.
4. ΔH is computed by Equation (4) using previously computed breach flow.

$$\Delta H = \frac{0.0826 \times \Delta t}{S_a}(Q_i - Q_b - Q_{sp} - Q_o) \quad (4)$$

where Q_i, Q_b, Q_{sp}, Q_o and S_a are the inflow to the reservoir, breach outflow, spillway flow, crest overflow and surface area at the reservoir elevation, respectively.

5. Q_b is computed from broad-crested weir relationship.
6. Regarding the stability of soil slopes, dimensions of breach are computed.
7. Q_s is computed using Meyer-Peter & Muller sediment-transport relation.
8. ΔH_c is computed by Equation (5).

$$\Delta H_c = \frac{3,600 \times \Delta t \times Q_s}{P_o \times L_b \times (1 - P_{or})} \quad (5)$$

in which P_o, L_b, P_{or}, Δt and Q_s are the perimeter of the breach, length of the breach channel, porosity of the breach material, increment time step and sediment-transport rate, respectively.

9. Computed ΔH is compared with estimated value. If tolerance in percent is small, the computed ΔH is accepted.
10. Breach sides and breach bottom are checked for collapse as:

$$F_w > F_{sb} + F_{ss} + F_{cb} + F_{cs} \quad (6)$$

where F_w, F_{sb}, F_{ss}, F_{cb} and F_{cs} are the force due to water pressure, shear force along the bottom of the breach, shear force along the side of breach, force due to cohesion along the bottom of the breach and force due to cohesion along the side of the breach, respectively.

11. Outflow discharge is computed.

It is also assumed that the breach bottom is not eroded downward until the volume of collapsed material along the breach is removed at the rate of the sediment-transport capacity of the breach channel. The outflow hydrograph is simulated and plotted time step by time step. The BREACH model has been developed based on some well-known sediment-transport formulas (e.g. Meyer-Peter & Muller), soil mechanics principles and hydraulics relations (e.g. broad-crested weir). Also, it has been already tested on some real-world dam breach cases such as the breach of two man-made dams (Teton and Lawn Lake) and one naturally formed landslide breach in Peru (Fread 1988).

ANN model

As with natural neural networks, an ANN contains some simple elements (e.g. neurons) that are connected together and operating in parallel. A neural network is trained to perform a particular function using a set of data and by adjusting the values of the connections (weights) between neurons such that a particular input leads to a specific target output. As shown in Figure 2, a three-layer feed-forward ANN, which is usually sufficient for real-world problems (Hornik *et al.* 1989), was used in this study. In Figure 2(*i*), (*j*) and (*k*) denote input layer, hidden layer and output layer neurons, respectively, and w is the applied weight by the neuron. The term 'feed forward' means that a neuron connection only exists from a neuron in the input layer to other neurons in the hidden layer or from a neuron in the hidden layer to neurons in the output layer and the neurons within a layer are not interconnected to each other. The explicit expression for an output

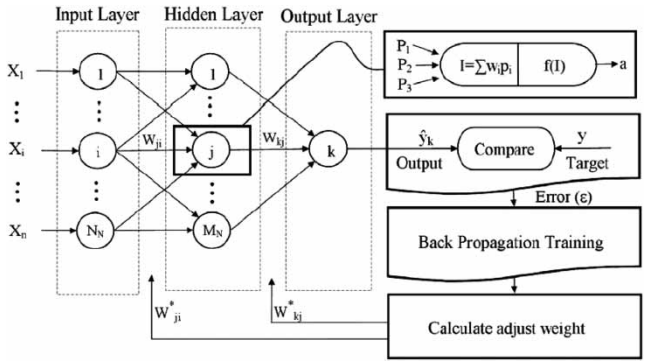

Figure 2 | A three-layered feed-forward neural network with BP training algorithm.

value of a three-layered feed-forward neural network is given by Nourani *et al*. 2011):

$$\hat{y}_k = f_0\left[\sum_{j=1}^{M_N} w_{kj} \cdot f_h\left(\sum_{i=1}^{N_N} w_{ji} \cdot x_i + w_{j0} + w_{k0}\right)\right] \quad (7)$$

where w_{ji} is a weight in the hidden layer connecting the ith neuron in the input layer and the jth neuron in the hidden layer, w_{j0} is the bias for the jth hidden neuron, f_h is the activation function of the hidden neuron, w_{kj} is a weight in the output layer connecting the jth neuron in the hidden layer and the kth neuron in the output layer, w_{k0} is the bias for the kth output neuron, f_0 is the activation function for the output neuron, x_i is the ith input variable for the input layer and $\hat{y}_k$, y are computed and observed output variables, respectively. N_N and M_N are the number of the neurons in the input and hidden layers, respectively. The weights are different in the hidden and output layers, and their values can be changed during the process of the network training.

It is very difficult to know which training algorithm will be the most reliable for a given problem. It depends on many factors, including the complexity of the problem, the number of data in the training set, the number of weights and biases in the network, the error goal and whether the network is being used for pattern recognition or function approximation (Nourani *et al*. 2009b).

There are two important issues concerning the implementation of ANNs. The first issue involves a specification of the network size (the number of layers in the network and the number of neurons in each layer). This task involves deciding the number of neurons required in the hidden layer. Generally, the more complex the mapping is, the larger the number of hidden neurons required becomes. However, an excessive amount of hidden neurons can help a network memorize the training set that will result in poor performance on hidden data. The second issue involves finding the optimal values for the connection weights. Starting with a small number of neurons and gradually increasing the network size until the desired accuracy is achieved addresses the first problem. However, this approach heavily depends on the ability to find the optimal weights. The back-propagation, as a learning algorithm for multi-layered neural networks, is widely used at different fields of engineering in order to modify the weights of network via the propagation of an error gradient backward from the output to the input. The presented ANN models in this research use the Levenberg–Marquardt back-propagation algorithm to train a feed-forward artificial neural network (ASCE 2000b). To assess the performance of the proposed models, two different criteria were used: the root mean square error (RMSE) and Nash–Sutcliffe efficiency coefficient (E) (or determination coefficient). The best fit between observed and calculated values will give RMSE and E values close to 0 and 1, respectively. These measures are defined as follows (Nourani *et al*. 2008):

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Q_i - \hat{Q}_i)^2} \quad (8)$$

$$E = 1 - \frac{\sum_{i=1}^{n}(Q_i - \hat{Q}_i)^2}{\sum_{i=1}^{n}(Q_i - \bar{Q})^2} \quad (9)$$

where Q_i is the observed peak outflow; $\hat{Q}_i$ is the simulated peak outflow; n is the number of observed data; and $\bar{Q}$ is the mean of the observed peak outflow. The Nash–Sutcliffe efficiency (E) ranges from $-\infty$ to 1. An efficiency of 1 ($E = 1$) corresponds to a perfect match of the modelled discharge to the observed data. An efficiency of 0 ($E = 0$) indicates that the model predictions are as accurate as the mean of the observed data, whereas an efficiency less than zero ($E < 0$) occurs when the observed mean is a better predictor than the model or, in other words, when the residual variance is larger than the data variance. Nash–Sutcliffe efficiency has been reported in scientific literature for model simulations of discharge and water quality constituents such as sediment, nitrogen and phosphorus loadings (Moriasi *et al*. 2007). To achieve an efficiency of 1 ($E = 1$), networks can be trained to perform classification with the function train. This function applies each vector of a set of input vectors and calculates the network weight and bias increments due to each of the inputs. Each pass through the input vectors is called an epoch. The number of training epochs is optimized to obtain precise and accurate outputs. The most commonly used transfer function for neurons in the hidden layer(s) and output layer of an ANN is the sigmoid

function, which has a bounded output range between zero and one (Rajurkar *et al*. 2002). The actual observed outputs of the network, which are outside of the bounded range of the neuron transfer function, must be normalized or rescaled such that they fall within the bounded output range. A logistic sigmoid is used here as the transfer function, and the observed discharges are normalized using the following equation:

$$y_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{10}$$

where y_i is the normalized parameter, $x_{\max}$ is the maximum observed parameter, $x_{\min}$ is the minimum observed parameter and x_i is the observed parameter. This transformation bounds the discharges in the range [0,1].

Experimental setup

An ANN model, similar to other black-box models, requires a complete data set to be calibrated. Calibration is often performed by using a specific set of data such as those from laboratory and, in some cases, field tests; however, in this research, three categories of data sources were used to produce a reliable model: experimental data, data from a physically based model and historical breach data.

To provide experimental data, some tests were conducted in a rectangular flume with a length of 12 m, a depth 0.8 m and a width 1.0 m (Figure 3(a)). The flume included a morning glory spillway upstream to provide control over the upstream reservoir water levels prior to breach initiation through an embankment. Water outflow through the developing breach channel was dropped into a stilling basin before flowing over a triangular sharp-crested weir into a second measuring-weir basin. A 90° V-notched weir was provided between two basins (Figure 3(b)).

The breach process was recorded by two video cameras (Figure 3(c)): one to monitor the breach growth and the other to record the depth of water on the weir. Forty experiments were conducted in the flume. Five different materials were used for dam construction in the experiments. The key properties of the materials are summarized in Table 1 in which c, ϕ, D_{50} and γ are the cohesive strength, internal

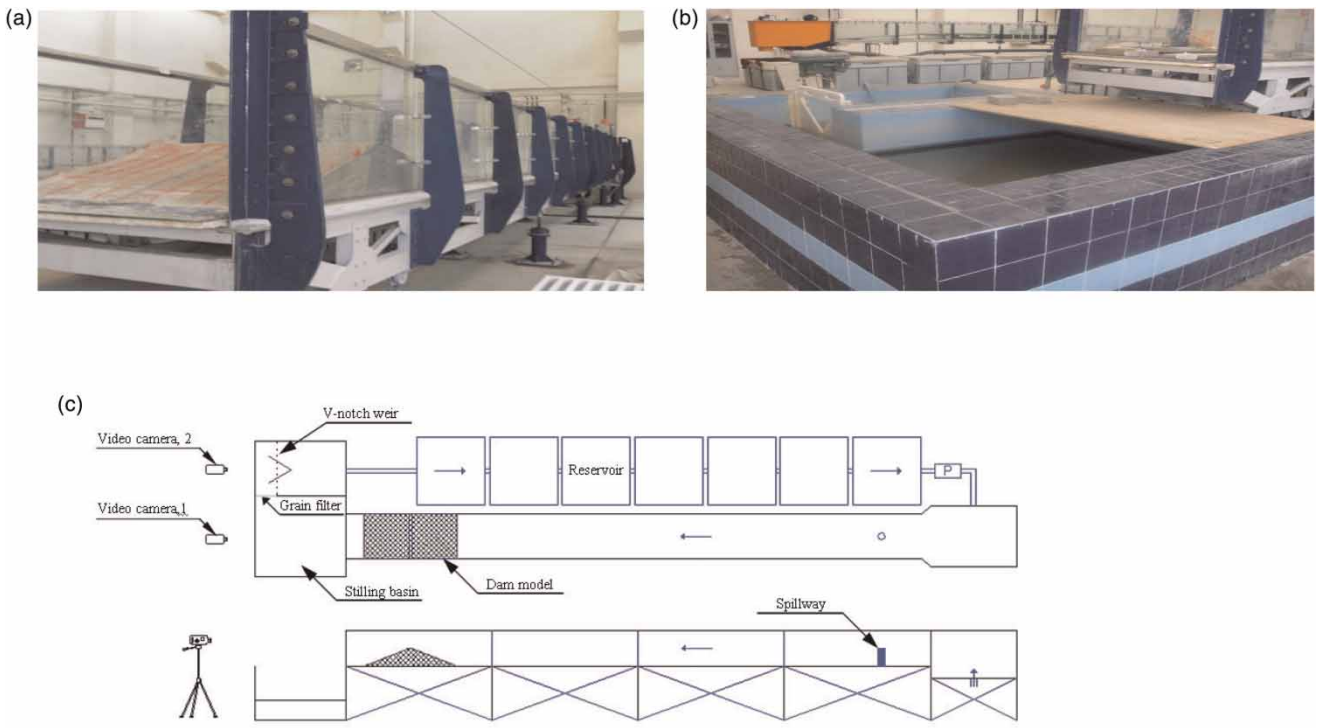

Figure 3 | Experimental setup.

Table 1 | Material properties

Type of material	c (KN/m^2)	ϕ	D_{50} (mm)	γ (KN/m^3)
1	0	45°	2	19.00
2	0	30°	0.25	16.25
3	33	30°	0.1	18.00
4	15	30°	0.15	17.60
5	5	30°	0.175	17.40

friction angle, median diameter of the soil material and unit weight of the soil, respectively.

The earth dams constructed for the present study were 30 and 40 cm in height, $z_D = 2.5$ and $z_u = 2.5$, 3.5 in which z_D and z_u, respectively, are the downstream and upstream slopes of the dam as given by ratio 1 (Vertical): z (Horizontal). Sixteen dams were constructed with two different types of sand, and 24 dams were constructed with different mixtures of fine sand and clay. To control the compaction of the dam, the material was water-sprayed and compacted in layers 5 cm thick, and a hand-operated compaction roller was used to compact every loose layer. During the test, the pilot channel initially eroded through a small breach in the dam downstream of the crest to the toe of the dam and continued until the breach reached the bottom flume. To draw the outflow hydrograph, the discharges were measured as a function of time according to the monitored depth of the water over the weir.

To validate the tests, the obtained results were compared with the results of other laboratory tests (Wishart 2007) and the data from four historical earthen dam failures. The geometry of the dam constructed by Wishart (2007) is shown in Figure 4(a). Wishart (2007) recognized the symmetry of the phenomenon and constructed only half of the dam. However, in the current study, the entire dam was constructed and breached (test No. = $t1$; Figure 4(b)).

Each breach process was initiated by cutting a rectangular pilot channel at the dam crest. As indicated in Figure 4(a), the data provided by Wishart (2007) were measured for a half-breach section in which the initial pilot channel was adjacent to the side of the flume. In the overtopping mode of dam failure, when water overflows from the crest, erosion begins at the toe of the downstream slope and creates a rivulet that migrates towards the embankment crest. When the rivulet reaches the crest, the dam breach has begun, and water passing through the breach leads to the enlargement of the breach until dam failure is completed. To simulate the overtopping mode on the laboratory scale, an initial notch should be created because it is preferred that the breach be initiated and take place in the middle of crest (arbitrary position). This initial notch was very small and did not have any significant effect on the breach outflow hydrograph. Figure 5 shows the result reported by

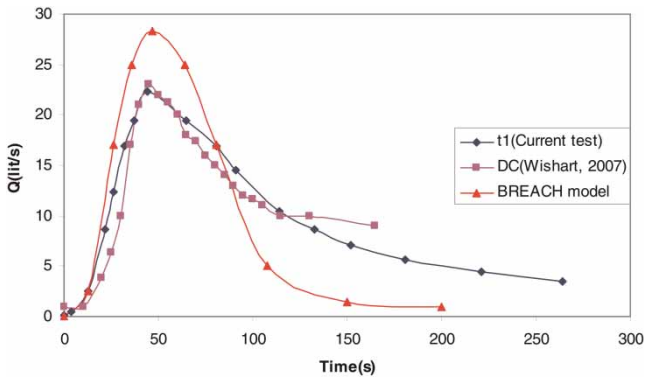

Figure 5 | Outflow hydrograph of two tests and BREACH model.

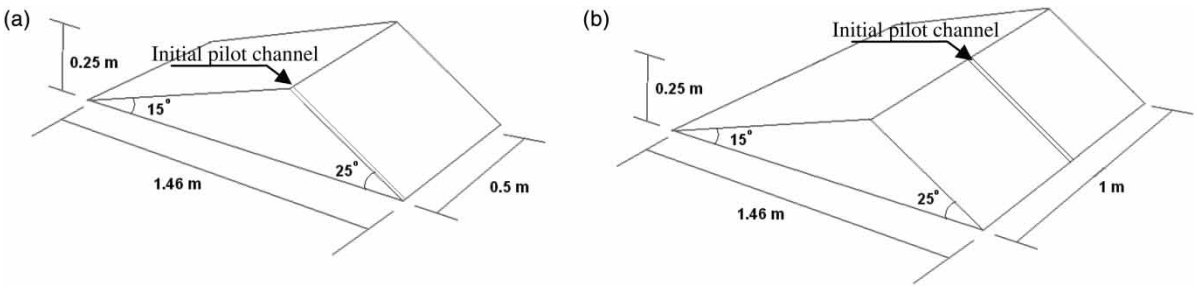

Figure 4 | Dams geometry: (a) Wishart (2007) test, (b) $t1$.

Wishart (2007), the observed data in the current study and the result obtained by the BREACH model.

As shown in Figure 5, the measured peak outflows are 22.3 and 22 l/s in Wishart (2007) and the current study ($t1$), respectively, but the calculated peak outflow by BREACH is 28 l/s. The breach area determined in Wishart's test (Figure 6(a)) has also been compared with the $t1$ test in Figure 6(b), which shows approximately the same breach in terms of shape and size. Regarding Figure 5, the time to the peak outflow in Wishart's test, the current test and the BREACH model are 45, 44 and 47 s, respectively.

For further validation, some experiments were also conducted by simulating selected historical dam breach cases: Apishapa, Mammoth, Otto Run and Teton dam (Wahl 1998). The characteristics of the dams are presented in Table 2.

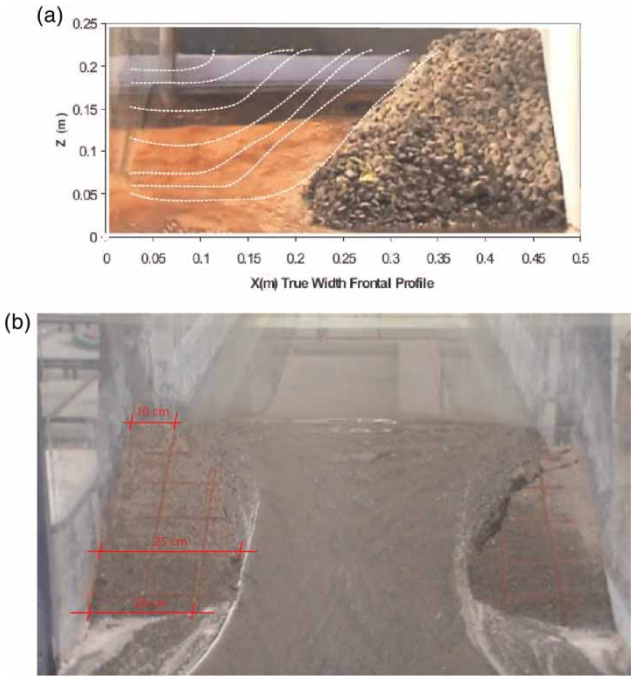

Figure 6 | Breach area (a) in Wishart (2007) model, (b) in $t1$ test.

Because in the overtopping failure we deal with free surface flow, the Froudian scaling law was applied such that $\lambda_F = 1$ to scale and set up the models in the laboratory, where λ_F is the scale ratio of Froude number $[F = V/(gL)^{0.5}]$, V is flow velocity, g is gravitational acceleration and L is the governing length. Therefore $\lambda_F = 1$, $\lambda_g = 1$ implies the following scaling ratios for the velocity, time and flow rate, respectively: $\lambda_V = \lambda_L^{0.5}$, $\lambda_t = \lambda_L/\lambda_V = \lambda_L^{0.5}$ and $\lambda_Q = \lambda_V\lambda_A = \lambda_L^{2.5}$. The scaling results of the dams are shown in Table 3. Two models were assumed to scale the dams. Model 1 used a 2-m length, considering the symmetry property used to scale the Apishapa and Mammoth dams; model 2 was used to scale the Otto Run and Teton dams with a 1-m length.

The measured peak outflow discharges obtained by the physical models are shown in Table 4. The test results are in good agreement with the corresponding historical breaches and, considering the presented results,

Table 3 | Froude scaling ratios

Parameter	Ratio	Apishapa	Mammoth	Otto Run	Teton
Length	λ_L	104	70	20	304.8
Time	$\lambda_t = \lambda_L^{0.5}$	10.2	8.4	4.4	17.5
Flow	$\lambda_Q = \lambda_L^{2.5}$	110,332	41,200	1,760	1.63×10^6

Table 4 | Physical models and historical dams peak outflows

Dam	Model	Q_p (lit/s) (measured)	Q_p (m^3/s) (scaled)	Q_p (m^3/s) (observed)	Error (%)
Apishapa	Model 1	56	6,180	6,850	9
Mammoth	Model 1	56	2,310	2,520	6
Otto Run	Model 2	36	63	60	5
Teton	Model 2	36	58,400	65,120	8

Table 2 | Characteristics of the dams

Name	Type	Height (m)	Length (m)	Storage (m^3)	Q_p (m^3/s)
Apishapa	Homogeneous	31.14	208	2.25×10^7	6,850
Mammoth	Homogeneous	21.3	140	1.36×10^7	2,520
Otto Run	Homogeneous	5.8	20	n/a	60
Teton	Zoned earth dam	92.96	304.8	3.56×10^8	65,120

it can be concluded that the experimental setup and measured data are reliable and can be used for the ANN modelling.

RESULTS AND DISCUSSION

In this research, two ANN models were developed to estimate the outflow discharge from an earth dam breach. The first ANN model, similar to some formerly presented black-box models (e.g. Equations (1)–(3)), used the physical parameters of the dam's geometry (height of water before breach H_w, reservoir storage volume V_w, length of crest L) and its material characteristics (cohesive coefficient c, internal friction angle ϕ and average grain of the soil D_{50}) as inputs to estimate the peak value of the outflow hydrograph. In the second model, in addition to the first models' inputs, some temporal variables, such as B_t^b, H_t^c, B_t^t and Q_{t-1} (which are the bottom width of the breach, height of the breach, top width of the breach at time t and the outflow discharge at time $t-1$, respectively), were also imposed on the input layer to estimate the outflow discharge as a function of time (i.e. Q_t).

The properties of the material used in the experiments, which also apply to real-world dams (prototypes), are presented in Table 1. Overall, to obtain the output hydrographs under different conditions, 40 dams were constructed, breached and monitored in the laboratory. The results of four tests are presented in Figure 7 for dams with height of 30 cm, downstream and upstream slopes of 2.5 and different soil properties.

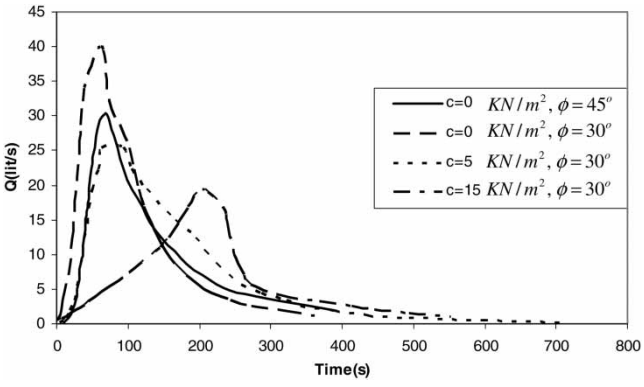

Figure 7 | Outflow hydrographs of tests.

Furthermore, the BREACH model was also used to determine the outflow hydrograph and breach shape of some hypothetical dams to produce the required data for the ANN training as the second category of the data source. The input parameters are given in Table 5.

As examples, Figure 8(a) and (b) illustrate dam breach hydrographs obtained via the BREACH model for dams with heights of 60 and 30 m, respectively, and different soil properties (i.e. $\phi = 30^\circ$ and $c = 0$, 15, 33 kN/m^2). It is clear that by increasing the cohesion of the material, the time of the breach is also delayed and, consequently, the peak of the outflow notably decreases. In the laboratory tests, it was observed that the shape of breach remained approximately rectangular until t_p (time to Q_p); at t_p, however, the collapse of the sides leads to a trapezoidal shape. The rate of breach enlargement in the dams with cohesive material is lower than that in non-cohesive dams and, consequently, in the rising limbs of the outflow hydrographs, the discharge values (including Q_p) are small compared with the non-cohesive breach cases; therefore, the time to Q_p increases.

Although there is a discrepancy between the peak value of the outflow from a laboratory breach and the value estimated by the BREACH (Figure 5) due to the assumed simplifications in the physically based model, it is possible to simulate the model several times for different conditions by the BREACH. Therefore, the simulated hydrographs may help the ANN model to estimate the rising and recession limbs of the breach hydrograph.

Numerous dam failures have occurred in the past, but only a few have been monitored. A database containing data from 108 embankment failures around the world collected by Wahl (1998) shows the details of embankment types and estimations of the concerned data. Some outflow hydrographs from historical breaches were used in this study

Table 5 | Input parameters of hypothetical dams

Length (m)	300			
Height (m)	30	40	50	60
Material property	$\phi = 0^\circ$ $c = 0$ KN/m^2	$\phi = 30^\circ$ $c = 15$ KN/m^2		$\phi = 45^\circ$ $c = 33$ KN/m^2
Slope	$z_D = z_u = 2.5$			

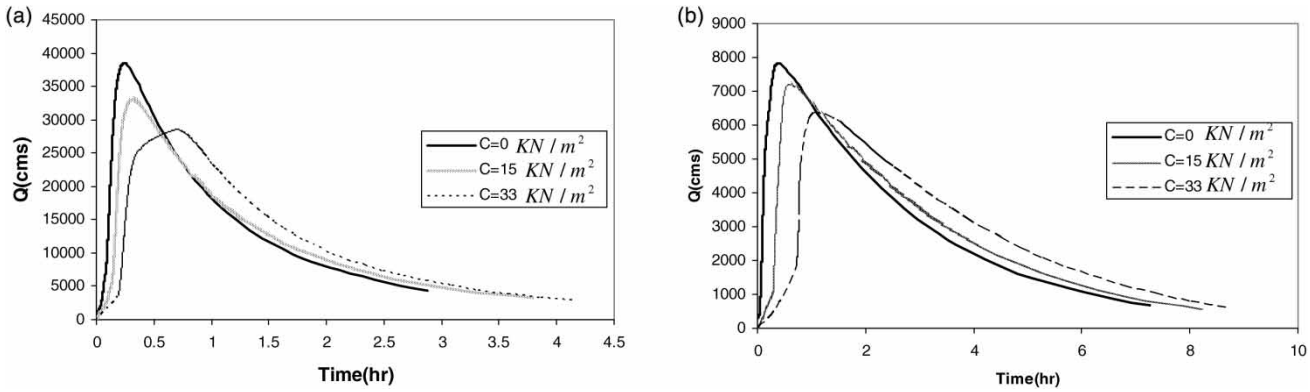

Figure 8 | Outflow hydrographs of hypothetical dams.

as the third category of data sources, two of which are shown in Figure 9.

Table 6 presents the statistical analysis of the data gathered from three sources, which are divided into three subsets of training, verification and test by the genetic algorithm (GA) technique, which includes the minimum, maximum, mean and standard deviation (S_d) of the data. The data were employed to develop two multilayer feed-forward ANN models, each with only one hidden layer.

In fact, a feed-forward network with only one hidden layer has been found capable of approximating any measurable function to any desired degree of accuracy (Hornik *et al.* 1989). This kind of ANN model accompanied by a back-propagation training algorithm is heavily used in hydraulic and hydrologic modelling (ASCE 2000b). Important issues in ANN modelling are architecture (i.e. number of neurons in input and hidden layers) and training epoch number, the appropriate selection of which can progress the model efficiency at both calibration (training) and verification steps. It also prevents the ANN model from being over-trained. To protect the ANN from overtraining, data were divided into three training, verifying and test sets. The network was trained by training data and in each training epoch the network was also verified by the validation data. The training process was continued until a critical epoch number where, in spite of a decrease in RMSE of the network by the training data, the RMSE of the network with the verification data was going to be increased. At this point, the training was stopped and the network was checked with the test data set. A network with the highest values of E (or lowest RMSE) in all three steps was selected as the best network. There are two new methods for the optimal division of data. The first method employs a GA to divide the data so as to minimize the statistical difference (measured by the mean and standard deviation)

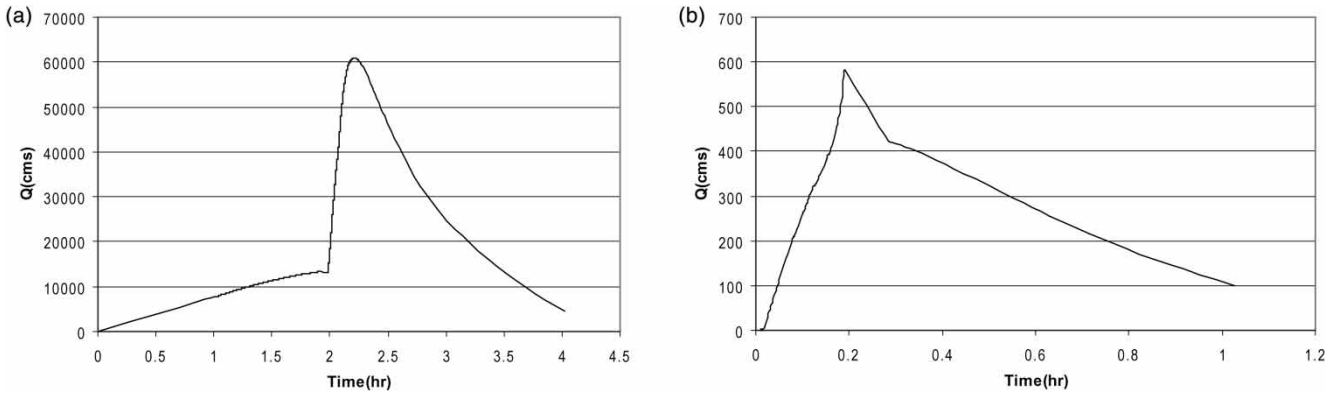

Figure 9 | Historical dam breach hydrographs: (a) Teton dam, (b) Lawn dam.

Table 6 | Data statistics

		H_d (m)	B_t^s (m)	H_t^c (m)	B_t^b (m)	D_{50} (mm)	c (KN/m^2)	ϕ	Q (cm)
Training set	Max	95	120	58.65	28.65	2	33	45	70,000
	Min	0.3	0.03	0.02	0.03	0.1	0	0	0.006
	Mean	45.48	58.45	13.53	14.62	1.4	25	7.99	10,651
	S_d	10.86	29.14	9.25	9.50	0.71	17	2.82	8,468
Validation set	Max	95	119.8	58.55	28.55	2	33	45	69,800
	Min	0.3	0.03	0.02	0.03	0.1	0	0	0.0065
	Mean	45.55	58.60	13.26	14.84	1.4	25	7.90	10,897
	S_d	10.80	29.60	9.13	9.39	0.70	17	2.79	8,590
Testing set	Max	95	119.6	58.45	28.45	2	33	45	69,550
	Min	0.3	0.03	0.02	0.03	0.1	0	0	0.0067
	Mean	45.24	58.35	13.19	14.57	1.4	25	7.99	10,588
	S_d	10.56	29.07	8.79	9.55	0.71	17	2.82	8,145

between training, testing and validation data sets (Goldberg 1989; Nourani 2010). The second data division method employs a self-organizing map (SOM) to cluster similar data records together (Bowden *et al.* 2002). In this study, a GA was applied to the problem of dividing the data into three statistically similar sub-sets. Sixty percent of the available data were used for training, 20% for validation and the remaining 20% for testing. At first, the data values were assigned to the sub-sets randomly and transformed to some GA chromosomes (strings) containing different genes and the fitness function was computed. In order to determine the fitness of each solution, an objective function was required. In this application, a suitable objective function to minimize was the sum of the absolute difference in the mean values between each pair of the three sub-sets. Penalty constraints were also added to the fitness to ensure that the maximum and minimum values of each input and output variables were included in the training set, rather than in the testing or validation set. To find the optimum solution, the gradually evolving concept of GA was used through different generations until a stopping criterion (an appropriate fitness) was reached. For this purpose at each generation, mutation and crossover operations were employed to guarantee the diversity and pressure of the answer's domain, respectively. The selection of the elite answers at each generation was also done by ranking and Roulette Wheel technique. For details regarding the GA approach the reader may refer to Nourani (2010).

The input and target data were normalized in the range of 0 to +1; finally, the log-sigmoid transfer function (Hornik *et al.* 1989), in the form of Equation (11), was applied:

$$\log sig(x_i) = \frac{1}{1 + e^{-x_i}} \tag{11}$$

Such a nonlinear function can map the weighted inputs of the ANN to the output of the model in order to capture Q_p.

The details and results of the proposed ANN models are presented in the following sections.

ANN-Q_p model

The ANN-Q_p model was developed to estimate the peak outflow discharge from a dam breach and also to investigate the sensitivity of the model (and process) to each input parameter. H_w, V_w, L, c, ϕ, D_{50} were used as input data, and Q_p was considered as the output. Eight combinations of the input variables were tested, and the models' architectures (number of neurons in input and hidden layers) and results are shown in Table 7. The result is reasonable, because the values of E and RMSE obtained in training, validation and test steps are close to 1 and 0, respectively,

Table 7 | Input variables in different structures for ANN-Q_p model

Structure no.	Architecture[a]	Input data	Epoch no.	*E* (Train)	*E* (Val.)	*E* (Test)	RMSE (Train)	RMSE (Val.)	RMSE (Test)
1	(6,11,1)	H_w, V_w, L, c, ϕ, D_{50}	100	0.9839	0.9484	0.9258	0.0014	0.0188	0.0275
2	(5,10,1)	H_w, V_w, c, ϕ, D_{50}	150	0.9632	0.9332	0.9069	0.0101	0.0193	0.0293
3	(3,7,1)	H_w, V_w, L	120	0.9124	0.8934	0.8764	0.0270	0.0495	0.0538
4	(2,5,1)	H_w, V_w	70	0.9081	0.8856	0.8701	0.0311	0.0498	0.0588
5	(1,3,1)	H_w	100	0.5962	0.5656	0.4026	0.0939	0.1145	0.1659
6	(1,3,1)	V_w	120	0.5009	0.4074	0.3601	0.1264	0.1178	0.1604
7	(3,6,1)	c, ϕ, D_{50}	90	0.4871	0.4125	0.3578	0.1301	0.1181	0.1655
8	(1,4,1)	L	80	0.4651	0.4277	0.3022	0.1312	0.1180	0.1763

[a]Number of neurons in (input layer, hidden layer, output layer).

which indicate that the overtraining problem has not taken place in the ANNs.

Figure 10 compares these eight structures in terms of the *E* and RMSE (for normalized data), which were calculated for the training, validation and testing data sets. The optimum numbers of hidden neurons and training epochs were determined through a trial–error process. Although the first structure gave the largest *E* and smallest RMSE, its architecture is complex and contains more parameters. Figure 10 indicates that the model leads to less accuracy for structures 5 to 8. Comparing the values of *E* and RMSE for structures 3 and 4 shows that H_w, V_w are the most dominant parameters in the dam breach analysis. This finding is in agreement with the formerly presented empirical models (e.g. Equations (1)–(3)), which use only H_w, V_w as the input parameters in the form of some exponential functions. Furthermore, a comparison of structures 5 and 6 reconfirms the appearance of a higher power of H_w in Equations (2) and (3).

On the other hand, in the cases that material properties of dam are not available prior to prediction (even for first estimation) these structures are useful.

The scatter plots of the computed versus observed values of Q_p (normalized) for structure 4 have been plotted in Figure 11(a) and (b) for training and test data sets, respectively.

To achieve an overall comparison for the developed ANN model and empirical relationships (Equations (1)–(3)), the obtained results were summarized, as shown in

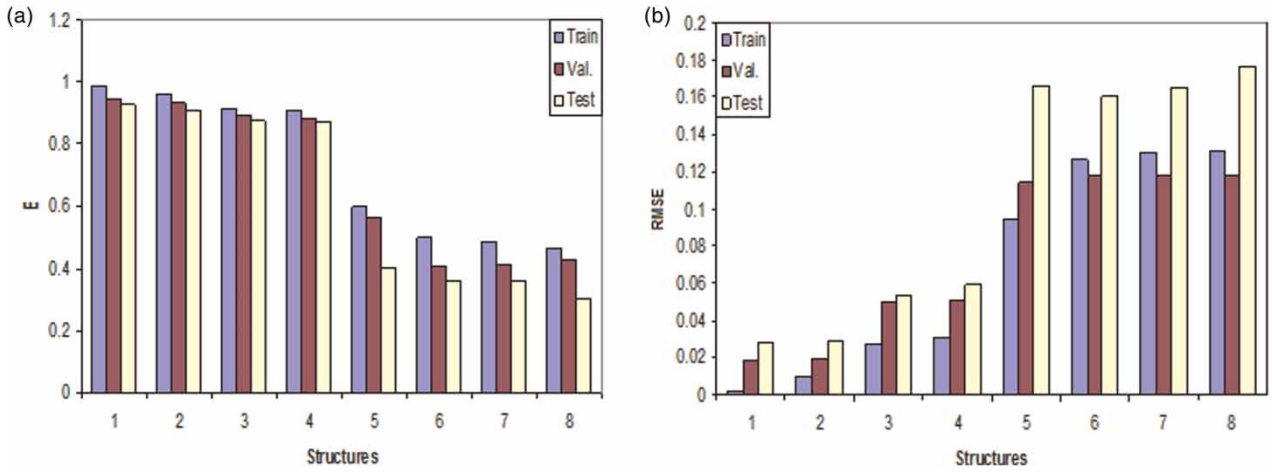

Figure 10 | (a) *E* and (b) RMSE of the models.

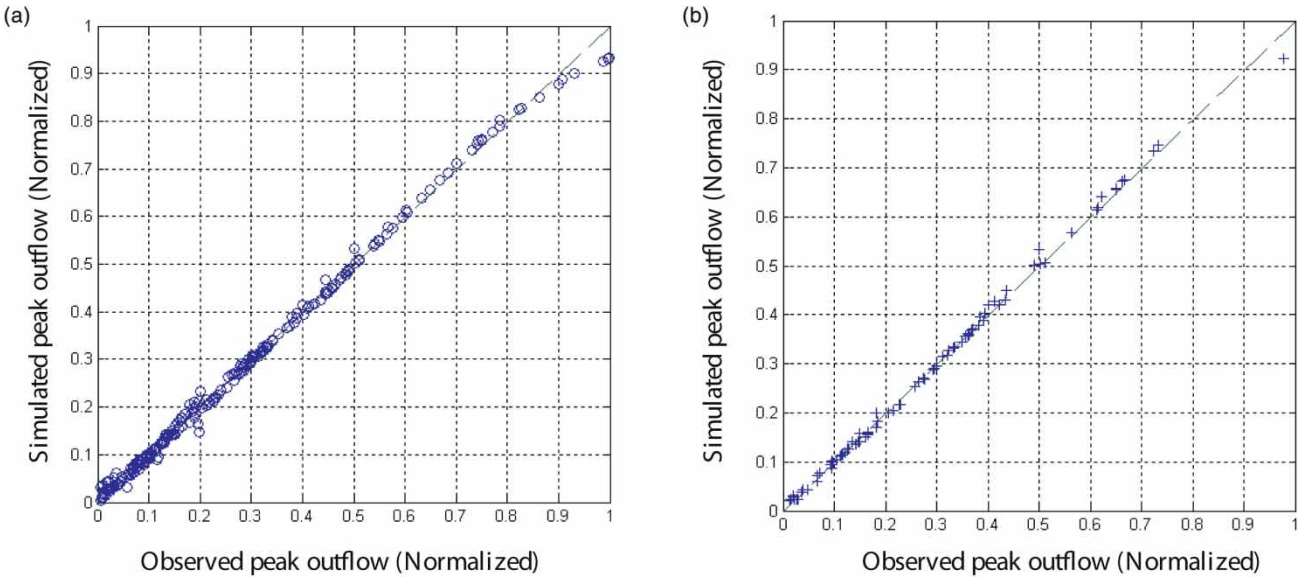

Figure 11 | Scatter plot for Q_p: (a) train, (b) test.

Table 8, for the Wahl database. The results indicate the superiority of the ANN model to the empirical models, which consider only an exponential relationship among the input and output parameters.

The appropriate model among the presented models may be selected in terms of both expected accuracy and data availability. Structures 1 and 2 lead to better results but also they need more field data. On the other hand, in the cases that material properties of the dam are not available prior to prediction (even for first estimation) structures 3 and/or 4 can be a reliable choice.

ANN-Q_t model

The ANN-Q_t model was developed to estimate the breach outflow hydrograph and, in this way, two different structures were examined.

At each time step (t), the outflow discharge (Q_t) through a breach can be expressed as a nonlinear function of the breach's size at that time step:

$$Q_t = f_1(B_t^b, H_c^t, B_t^t) \quad \text{in which } 0 \le Q_t \le Q_p \tag{12}$$

or

$$Q_t = f_2(Q_p, B_t^b, H_c^t, B_t^t) \tag{13}$$

such that f_2 interpolates the discharge values between 0 and Q_p, temporally.

In the previous section, it was verified that Q_p can be estimated by H_w, V_w, L, c, ϕ, D_{50} via a nonlinear ANN model (i.e. ANN-Q_p). On the other hand, although during most time steps Q_t is not affected by the initial size of the breach (i.e. initial conditions as B_0^b, H_0^t, B_0^t), it is possible that during some first time steps, Q_t is affected by the initial conditions of the breach. It has already been proved that the initial size of the breach may be represented by the physical properties of the dam (Fread 1988). Thus, Equation (12) can be rewritten as follows:

$$Q_t = f(H_w, V_w, L, c, \phi, D_{50}, B_t^b, H_c^t, B_t^t) \tag{14}$$

However, in real-world applications, it is usually difficult to monitor breach growth and the size of a breach,

Table 8 | Verification results for three empirical models and ANN

Model	Costa	Froehlich	Webby	ANN (structure 4)
E	0.5415	0.7812	0.8466	0.9081
RMSE	0.1269	0.0758	0.0421	0.0278

time step by time step. Therefore, assuming the breach phenomenon is a Markovian process (i.e. auto-regressive process of order 1, AR(1)) as:

$$Q_t = g_1(Q_p, Q_{t-1}) \quad (15)$$

or

$$Q_t = g(H_w, V_w, L, c, \phi, D_{50}, Q_{t-1}) \quad (16)$$

the second structure of the ANN-Q_t model is presented. In this model, the network uses the output of the previous time step simulation as the model's input at the current time step. It has been shown that for the simulation of a dam breach, H_w and V_w can be reliable representatives of the physical properties of the dam (Table 7). Therefore, two simplified alternatives for structures 1 and 2 were also considered in modelling structures 3 and 4, which employ four fewer parameters than structures 1 and 2 in the input layer. The results of the modelling are shown in Table 9 for the proposed ANN-Q_t structures.

Although the results show that structures 1 and 3 performed better than structures 2 and 4, respectively, it is clear that structures 2 and 4 are more applicable to real-world problems. For example, the breach hydrographs of the Teton dam computed by the presented ANN-Q_t (structure 1) and BREACH models were drawn versus the observed hydrograph, which shows that the presented ANN-Q_t, which is derived from different kinds of data sources, can be considered an adequate tool to simulate the breach process (see Figure 12).

According to the current research, Figure 13 briefly shows a step-by-step algorithm for application of ANN technique to predict the dam breach output discharge. At the preparation phase, as mentioned in this paper, different

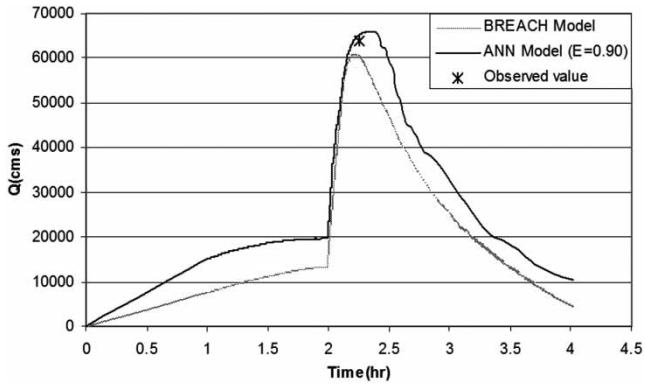

Figure 12 | Calculated and observed hydrographs for Teton dam breach.

data from several sources are employed to train and test different ANNs with different abilities; thereafter and according to the presented steps at simulation phase, the prepared ANNs as ready packages can be used for any other study dam that is in the design and/or operation phase.

CONCLUSIONS

Prediction of the outflow hydrograph resulting from a gradually failed earth dam is the basic step for estimation of the downstream flood levels and extraction of the inundation maps. For this purpose, there are different methods to determine outflow hydrographs which have their own advantages and shortages. The dam breach problem is a complex phenomenon, and it is difficult to model through a fully physically based model. Currently, to model complex phenomena, artificial intelligence methods such as ANNs are widely used. All methods require data for training. Therefore, to have an adequate model, three data categories were used to cover a wide range of dam failures. In category 1, to gather experimental data which are more similar to real cases but on a small

Table 9 | Input variables in different structures for ANN-Q_t model

Structure no.	Architecture[a]	Input data	Epoch no.	*E* (Train)	*E* (Val.)	*E* (Test)	RMSE (Train)	RMSE (Val.)	RMSE (Test)
1	(9,13,1)	H_w, V_w, L, c, ϕ, D_{50}, B_b^t, H_c^t, B_t^t	130	0.9526	0.9145	0.8941	0.0154	0.0280	0.0344
2	(7,10,1)	H_w, V_w, L, c, ϕ, D_{50}, Q_{t-1}	120	0.9307	0.9003	0.8765	0.0211	0.0321	0.0412
3	(5,9,1)	H_w, V_w, B_b^t, H_c^t, B_t^t	110	0.9218	0.8966	0.8601	0.0357	0.0577	0.0711
4	(3,7,1)	H_w, V_w, Q_{t-1}	90	0.9058	0.8845	0.8453	0.0489	0.0685	0.0901

[a]Number of neurons in (input layer, hidden layer, output layer).

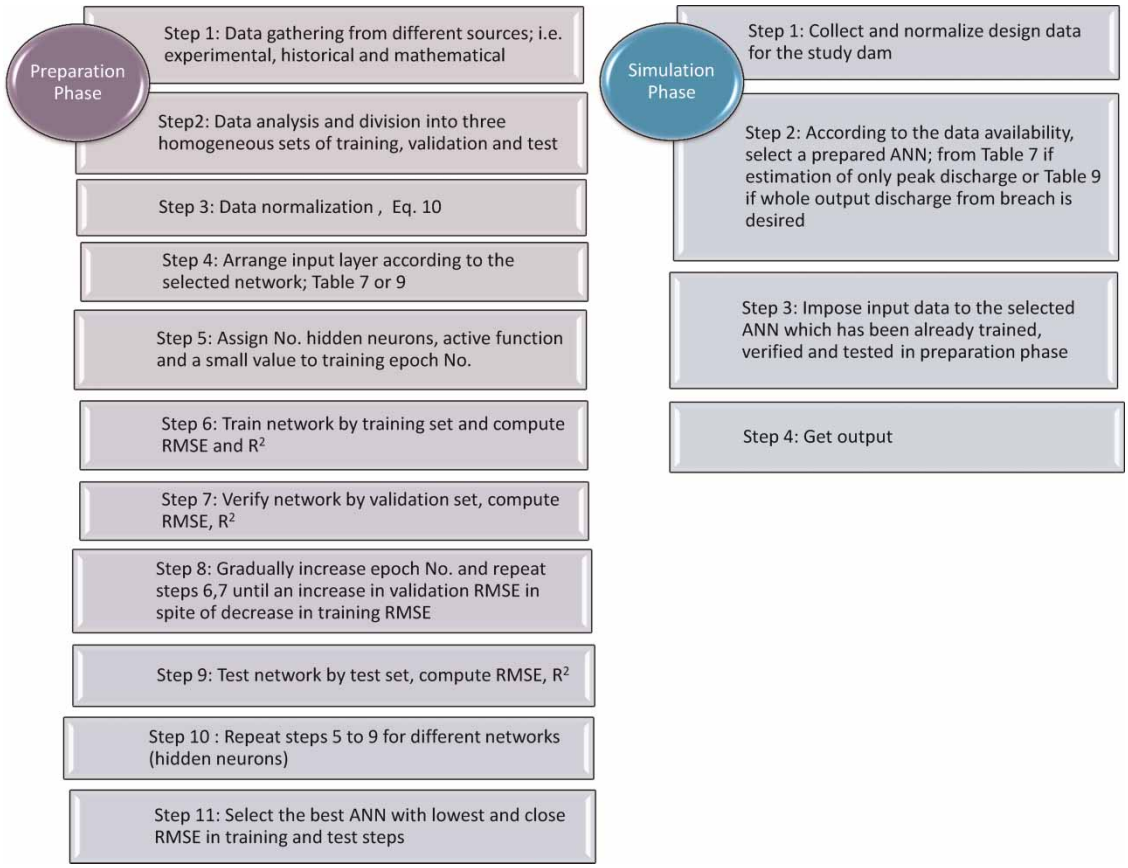

Figure 13 | Step-by-step algorithm for application of ANN to predict dam breach outflow.

scale, homogeneous small dams were constructed in a flume using a range of uniform non-cohesive and cohesive materials and were breached by overtopping flows. In category 2, a physically based numerical model (BREACH) was used to produce large-scale results, but the model included some simplifying assumptions. Alternative approaches to numerical modelling of dam breaching also have significant associated uncertainties for two reasons. First, they do not simulate erosion mechanisms that are relevant to the dam breach and, second, they typically rely on sediment-transport relations that are not applicable to the regime of flow conditions occurring in dam breaching. Category 3 includes historical breach data, which are scarce, and predictions of breach flows and breach parameters based on statistical analyses of historical data tend to handle significant associated uncertainties. This finding is due to the rare, highly variable and often subjective nature of the available historical data; these data represent, at best, only a limited range of earth dams.

The results show that the ANN method is more flexible than other black-box or physical based models. Also, ANN is a flexible model for sensitivity analysis. Sensitivity analysis shows that H_w and V_w are more important physical parameters than L, c, ϕ and D_{50} in dealing with the breach process. In this research, an ANN model was also developed to derive outflow hydrographs, considering the size of the breach and instantaneous Q_{t-1} as inputs. It should be emphasized that for many existing embankments (both dikes and earth dams) c and D_{50} are not exactly known. Therefore as a preliminary phase of the study and/or when material properties of the dam are not available, a prediction can be done by using just H_w and V_w. To predict Q_p with much higher accuracy, we will need internal material properties of the dam.

Due to the complexity of the process on one hand and the lack of sufficient data on the other, we cannot claim that the presented ANN model is a perfect approach to simulate the breach phenomenon. However, according to the flexibility of the proposed method, we suggest that the current database be extended by more field and experimental data to improve the efficiency of the ANNs. For instance, some non-homogenous zoned dams can be scaled and breached in the laboratory to investigate the effect of material heterogeneity on the breach process.

The database formed in this research can be used by other researchers to develop and verify other new models and methods in future studies.

ACKNOWLEDGEMENT

The authors gratefully acknowledge Sahand University of Technology for laboratory facilities provided for this study.

REFERENCES

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology 2000a Artificial neural networks in hydrology 2: hydrologic applications. *J. Hydrol. Eng.* **5** (2), 124–137.

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology 2000b Artificial neural networks in hydrology 1: preliminary concepts. *J. Hydrol. Eng.* **5** (2), 115–123.

Bowden, G. J., Maier, H. R. & Dany, C. 2002 Optimal division of data for neural network models in water resources applications. *Water Res. Resour.* **38** (2), 1–11.

Carriere, P., Mohaghegh, S. & Gaskari, R. 1996 Performance of a virtual runoff hydrograph system. *J. Water Res. Pl.* **122** (6), 421–427.

Chinnarasri, C., Jirakitlerd, S. & Wongwises, S. 2007 Embankment dam breach and outflow characteristics. *Civil Eng. Environ. Syst.* **21** (4), 247–264.

Costa, J. E. 1985 Floods from dam failures. *U.S Geological Survey Open-File Report 85–560*, Denver, Colorado, 54 pp.

Faeh, R. 2007 Numerical modelling of breach erosion of river embankments. *J. Hydraul. Eng.* **133** (9), 1000–1009.

Federal Energy Regulatory Commission, FERC 1987 *Engineering Guidelines for the Evaluation of Hydropower Projects*, FERC 0119-1. Office of Hydropower Licensing, Washington, DC, USA. 9 pp.

Franca, M. J. & Almeida, A. B. 2004 A computational model of rockfill dam breaching caused by overtopping. *J. Hydraul. Res.* **42** (2), 197–206.

Fread, D. L. 1984 DAMBRK: *The NWS Dam-Break Flood Forecasting Model, National Weather Service*. Office of Hydrology, Silver Spring, Maryland.

Fread, D. L. 1988 BREACH: An erosion model for earthen dam failures. *National Weather Service, National Oceanic and Atmospheric Administration*, Silver Spring, MD, USA.

Fread, D. L. 1993 NWS FLDWAV Model: The Replacement of DAMBRK for Dam-Break Flood Prediction. In *Dam Safety'93, Proceedings of the 10th Annual ASDSO Conference*. Kansas City, Missouri, pp. 177–184.

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

Goldberg, D. E. 1989 *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Longman Publishing Co., Inc., MA, USA.

Hettiarachchi, P., Hall, M. J. & Minns, A. W. 2005 The exploration of artificial neural networks for the modelling of rainfall–runoff relationships. *J. Hydroinform.* **7** (4), 291–296.

Hornik, K., Stinchcombe, M. & White, H. 1989 Multilayer feedforward networks are universal approximators. *Neural Net.* **2** (5), 359–366.

Johnson, F. A. & Illes, P. 1976 A classification of dam failures. *Int. Water Power and Dam Const.* **28**, 43–45.

Kirkpatrick, G. W. 1977 *Evaluation Guidelines for Spillway Adequacy. The Evaluation of Dam Safety*. Engineering Foundation Conference, ASCE, Pacific Grove, California, pp. 395–414.

Liong, S. Y., Lim, W. H. & Paudyal, G. N. 2000 River stage forecasting in Bangladesh: neural network approach. *J. Comput. Civil Eng.* **14** (1), 1–8.

Macchione, F. 2008 Model for predicting floods due to earthen dam breaching. I: formulation and evaluation. *J. Hydraul. Eng.* **134** (12), 1688–1696.

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

Mohamed, M. A., Samuels, P. G., Morris, M. W. & Ghataora, G. S. 2002 *Modelling Breach Formation through Embankments*. HR Wallingford Howbery Park, Wallingford, Oxon, OX10 8BA, UK.

Moradkhani, H., Hsu, K. L., Gupta, H. V. & Sorooshian, S. 2004 Improved streamflow forecasting using self-organizing radial basis function artificial neural networks. *J. Hydrol.* **295** (1–4), 246–262.

Moriasi, D. N., Arnold, J. G., Vanliew, M. W., Bingner, R. L., Harmel, R. D. & Veith, T. L. 2007 Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Tran. ASABE* **50** (3), 885–900.

Morris, M., Hanson, G. & Hassan, M. 2008 Improving the accuracy of breach modelling: why are we not progressing faster? *J. Flood Risk Manage.* **1**, 150–161.

Muttiah, R. S., Srinivasan, R. & Allen, P. M. 1997 Prediction of two year peak stream discharges using neural networks. *J. Am. Water Resour. Assoc.* **33** (3), 625–630.

Nourani, V. 2010 Application of Meta-Heuristic optimization approaches to investigate velocity profile effect on optimal

design of open channels. In: *Hydraulic Engineering: Structural Applications, Numerical Modelling and Environmental Impacts* (G. Hirsch & B. Kappel, eds.). NOVA Pub., New York, NY, USA. pp. 157–180.

Nourani, V. & Kalantari, O. 2010 Integrated Artificial Neural Network for spatiotemporal modelling of rainfall–runoff–sediment processes. *Environ. Eng. Sci.* **27** (5), 411–422.

Nourani, V. & Mano, A. 2007 Semi-distributed flood runoff model at the sub-continental scale for south-western Iran. *Hydrol. Process.* **21**, 3173–3180.

Nourani, V., Asghari Mogaddam, A. & Nadiri, A. 2008 An ANN-based model for spatiotemporal groundwater level forecasting. *Hydrol. Process.* **22**, 5054–5066.

Nourani, V., Alami, M. T. & Aminfar, H. 2009a A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation. *Eng. Appl. Artif. Intel.* **22**, 466–472.

Nourani, V., Komasi, M. & Mano, A. 2009b A multivariate ANN-Wavelet approach for rainfall-runoff modelling. *Water. Resour. Manag.* **23** (14), 2877–2894.

Nourani, V., Kisi, O. & Komasi, M. 2011 Two hybrid artificial intelligence approaches for modeling rainfall–runoff process. *J. Hydrol.* **402**, 41–59.

Rajurkar, M. P., Kothyari, U. C. & Chaube, U. C. 2002 Artificial neural networks for daily rainfall-runoff modelling. *Hydrol. Sci. J.* **47** (6), 865–877.

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.

Singh, V. P. & Quiroga, C. A. 1988 Dimensionless analytical solutions for dam-breach erosion. *J. Hydraul. Res.* **26** (2), 179–197.

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

Smith, J. & Eli, R. N. 1995 Neural-network models of rainfall-runoff process. *J. Water Res. Pl.* **121** (6), 499–508.

Tingsanchali, T. & Chinnarasri, C. 2001 Numerical modelling of dam failure due to flow overtopping. *Hydrol. Sci. J.* **46** (1), 113–130.

US Bureau of Reclamation 1982 Guidelines for Defining Inundated Areas Downstream from Bureau of Reclamation Dams. *Reclamation Planning Instruction* **8**, 2–11.

Visser, P. J. 1998 *Breach Growth in Sand-Dikes*. PhD Thesis, Delft University of Technology, Delft, The Netherlands.

Von Thun, J. L. & Gillette, D. R. 1990 *Guidance on Breach Parameters, unpublished internal document*. US Bureau of Reclamation, Denver, Colorado, 17 pp.

Wahl, T. L. 1998 Prediction of embankment dam breach parameters. *Dam Safety Research Report, DSO-98–004*. Available from: http://www.usbr.gov/pmts/hydraulics_lab/twahl/index.cfm

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

Wang, Z. & Bowles, D. S. 2006 Three-dimensional non-cohesive earthen dam breach model. *Adv. Water Res.* **29**, 1528–1545.

Webby, M. G. 1996 Discussion of peak outflow from breached embankment dam. *J. Water Res. Pl.* **122** (4), 316–317.

Wishart, J. S. 2007 *Overtopping breaching of rock-avalanche dams*. MSc Thesis, University of Canterbury.

Zamani, A., Azimian, A., Heemink, A. & Solomatine, D. 2009 Wave height prediction at the Caspian Sea using a data-driven model and ensemble-based data assimilation methods. *J. Hydroinform.* **11** (2), 154–164.

Zhu, Y., Visser, P. J. & Vrijling, J. K. 2006 Laboratory observations embankment breaching. In *The 7th Int. Conf on Hydroscience and Engineering*, Philadelphia, USA.

First received 30 January 2011; accepted in revised form 11 July 2011. Available online 7 November 2011