ESTIMATING PIPING POTENTIAL IN EARTH DAMS AND LEVEES USING GENERALIZED NEURAL NETWORKS

XINHUA XUE, XINGGUO YANG and XIN CHEN

about the authors

Xinhua Xue
Sichuan University,
State Key Laboratory of Hydraulics and Mountain River Eng.,
College of Water Resource and Hydropower
No.24 South Section 1, Yihuan Road, Chengdu, Sichuan
610065, P.R.China
E-mail: scuxxh@163.com

Xingguo Yang
Sichuan University,
State Key Laboratory of Hydraulics and Mountain River Eng.,
College of Water Resource and Hydropower
No.24 South Section 1, Yihuan Road, Chengdu, Sichuan
610065, P.R.China
E-mail: 89022251@163.com

Xin Chen
Sichuan University,
State Key Laboratory of Hydraulics and Mountain River Eng.,
College of Water Resource and Hydropower
No.24 South Section 1, Yihuan Road, Chengdu, Sichuan
610065, P.R.China
E-mail: chenxin_scu@163.com

Abstract

Internal erosion and piping in embankments and their foundations is the main cause of failures and accidents to embankment dams. To estimate the risks of dam failure phenomenon, it is necessary to understand this phenomenon and to develop scientifically derived analytical models that are simpler, easier to implement, and more accurate than traditional methods for evaluation of piping potential. In this study, a generalized regression neural network (GRNN) technique has been applied for the assessment of piping potential, as well, due to its ability to fit complex nonlinear models. The performance of GRNN has been cross validated using the k-fold cross validation method technique. The GRNN model is found to have very good predictive ability and is expected to be very reliable for evaluation of piping potential.

keywords

piping, generalized neural network, cross validation, BP neural network

INTRODUCTION

In recent years, dam safety has drawn increasing attention from the public. This is because floods resulting from the breaching of dams can lead to devastating disasters with tremendous loss of life and property, especially in densely populated areas [1]. Past dam-failure disasters showed that the loss of life is directly related to the evacuation time available, should failures occur. It is therefore very important to understand the process of dam breaching, and if possible, to obtain the key breaching parameters required to quantitatively describe the dam-breaching process.

Earth and rockfill dams are widely used throughout the world, and most past dam failures involved such dams. Piping is the most common cause of dam failures. According to statistics, approximately half of all the world's dam failures are attributed to piping phenomena [2-3]. Possible piping-phenomena modes include heave, internal erosion and backwards erosion. While the most common piping-failure mode is internal erosion, most often associated with conduits or other structural penetrations through dams, up to one-third of all piping failures can be attributed to backwards-erosion piping [4].

Piping phenomena have been studied by a large number of researchers and the list of publications on this topic is voluminous. For instance, in the early twentieth century, Terzaghi [5] performed experiments on heave-type piping, and developed a theoretical framework for the prediction of heave. Based on this framework, an equation for the factor of safety against heave was later adopted by practitioners for all piping-failure modes, such as backward-erosion piping and piping along internal fractures (concentrated leak erosion), although the theoretical basis for this adaptation of Terzaghi's equations has not been confirmed with laboratory experiments. Bligh [6] and later Lane [7] recognized a fundamental difference in the piping mechanisms between the seepage occurring through intergranular flow and the seepage along soil/structure boundaries. Modern practitioners define these piping mechanisms

as backwards-erosion piping (for the intergranular-flow case), and concentrated leak erosion for the case of flow along pre-existing openings (either soil-structure openings or cracks through an embankment), and so on. However, more research is still needed to fill in the gaps in our understanding of piping. Many existing structures were constructed without filters or with inadequate filters. Current methods for the evaluation of piping potential are based on theories that were developed almost 100 years ago, which have proven to be inadequate, when one considers the range of mechanisms that fall under the heading "piping"[8]. Therefore, many scholars and experts have attempted to develop scientifically derived analytical models that are simpler, easier to implement, and more accurate than traditional methods for the evaluation of piping potential.

In recent years, due to pervasive developments in computational software and hardware, several alternative computer-aided pattern-recognition approaches have emerged [9]. An artificial neural network (ANN) is a sophisticated computational approach capable of modelling a highly complex function [10]. It is frequently employed to mimic a nonlinear function with a large number of variables that cannot be modelled with a classical linear function. The theory, design and application of ANNs have been advancing a great deal during the past 20 years in order to solve complicated problems in different fields, including structural and earthquake engineering [11-12], construction engineering [13], geotechnical engineering [14], signal processing [15] and geosciences [16]. The neural network is well known as a parallel information-processing structure that consists of simple processing units with a high degree of interconnection between each unit. Each neuron performs a relatively simple job: receive impulses from input cells or neurons, carry out some types of transformation of the inputs and pass on the weighted products to the other cells or other neurons. The function of a neural network replicates that of a biological brain, which is basically composed of a very large number of massively interconnected neurons. Neural networks exhibit a mapping capability, learn from examples and possess the capability to generalize. They are robust systems that attempt to mimic the fault-tolerance and can process information in parallel with high speed. For these outstanding capabilities, neural networks are designed for pattern-recognition and classification applications.

There are different types of neural networks that can be employed for pattern recognition, e.g., generalized neural network (GRNNs), multi-layer perceptron and recurrent network, and so on. In this study, a GRNN has been chosen to predict the occurrence of a seepage piping disaster in terms of selected factors. The performance of the GRNN was cross-validated using the *k*-fold cross-validation technique. The main purpose of this paper is to determine the scope and suitability of the GRNN for the prediction of piping disasters in embankments.

2 MECHANISM OF PIPING

Piping is a process whereby internal voids are created by seepage flow. Erosion in earth structures due to water flow occurs when the erosion-resistant forces are less than the seepage forces that tend to produce it, in such a way that the soil particles are removed and carried with the water flow (Fig.1).

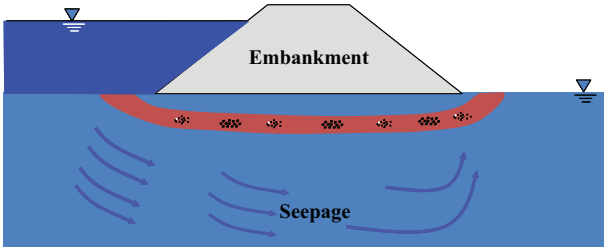

Figure 1. Illustration of piping leakage.

The resistant forces depend on the cohesion, the interlocking effect, the weight of the soil particles and the kind of protection they have downstream, if any. Since the seepage through an earth structure is not uniform, the erosion phenomenon increases where there is a concentration of seepage and water velocity; in places where this concentration emerges at the downstream side, the erosive forces on the soil particles might become very significant. This accentuates the subsequent concentration of seepage and erosive forces there. The seepage water has a velocity and force such that it is able to remove soil particles from the ground surface at the point of exit. As the soil particles are removed, the length of the seepage path decreases, resulting in a higher hydraulic gradient. The increased hydraulic gradient induces a higher seepage velocity and force, which in turn moves ever larger soil particles (Fig.2).

Piping in earth dams can be categorized based on the erosion process, i.e., concentrated leak erosion, backward erosion, suffusion and soil contact erosion [17]. This is shown in Fig.3 (a) and (b) for piping through the foundation, and from the embankment to the foundation. Concentrated leak erosion is defined as the erosion of soil in a crack or a system of interconnecting voids in an earth dam or its foundation [17]. The initial cracks generated by differential settlement, desiccation or

a) b)

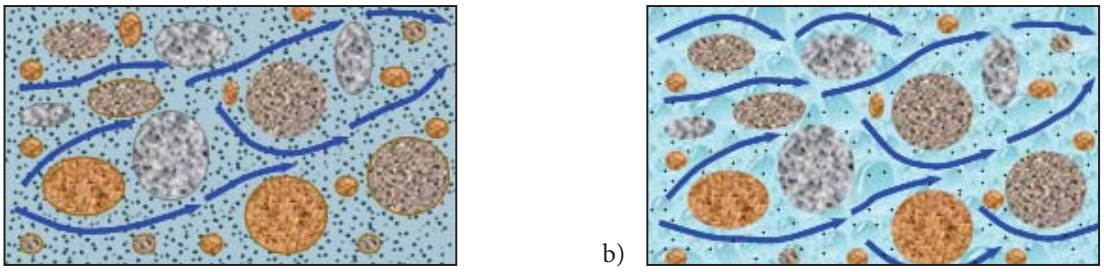

Figure 2. Illustration of sand content and seepage flow during piping leakage. a) Before piping. b) After piping.

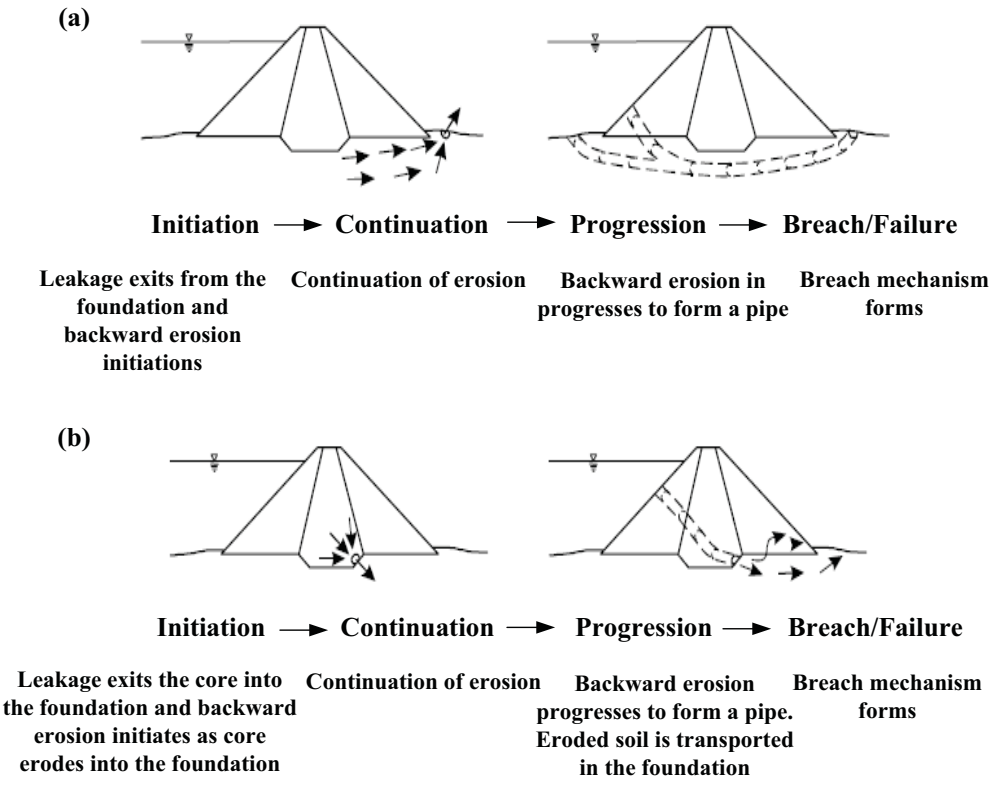

Figure 3. Models for the development of failure by piping in the foundation and from the embankment to the foundation [18]. a) Piping in the foundation initiated by backwards erosion. b) Piping from the embankment to foundation initiated by backwards erosion.

freezing may be shut by the swelling or softening of the soil, but remain to be a serious problem for dams with unfavourable internal stress conditions.

The above process can be described briefly as follows [19]:

1) Initiation. Cracks or defects providing potential seepage paths in the dam or foundation may result from a number of causes: differential settlement, desiccation, backward erosion, earthquake shaking, high-permeability zones due to poor compaction or segregation of the fill during placement, suffusion (internal instability) of gap graded soils, or hydraulic fracture. Erosion begins when the hydraulic gradient in the seepage path becomes large enough to move the soil particles. Thus, the requirements for this phase are: 1) a path for concentrated flow (usually a crack of some kind) and 2) the initiation of erosion along the seepage path.

2) Continuation. The requirements for the continuation of erosion are the lack of an effective filter and the ability for the eroding particles to be removed at the downstream end of the seepage path. While a designed filter is obviously the most reliable means for stopping the removal of particles along a seepage pathway, other fill zones and native soil deposits may also provide restraint, even if they do not meet modern filter-design criteria. In such cases, some erosion of the base soil may take place before the filtering material becomes effective.

3) Progression. In the progression phase the seepage pathway enlarges to form a pipe or open eroded conduit through the embankment. This requires sufficient flow through the leak to maintain the eroding velocities, and the ability of the surrounding soil to support a pipe without collapse, as illustrated in Fig.3. The flow-limiting ability of the upstream and downstream zones can control the flow volume through the leak. In addition, crack fillers can migrate from upstream zones into defects within the seepage barriers and seal off concentrated leaks.

   Soils with the ability form and hold an arch (soils with plastic fines and cemented granular soils are likely to have the highest arching ability) are capable of supporting a pipe [20, 21]. It is also possible for a pipe to be supported by bedrock or competent soil above the eroding soil.

4) Breach. Breach mechanisms may include the gross enlargement of the pipe, crest settlement, unravelling of the downstream slope, and slope instability leading to an open flow channel through the embankment.

From the above it can be seen that the extent and severity of the piping is dependent on the embankment-soil type. A soil that allows bridging, or a "roof" to form over the void, will allow piping to develop to a large extent and remain unnoticed from above. Piping often causes sinkholes to form at isolated points along its path through the embankment. If the soil mass over the pipe is unable to bridge, or the pipe enlarges to such an extent that the overlying soil mass collapses, a sinkhole will be created. Seepage may exit the downstream face of an embankment or its foundation at any point. Therefore, piping may develop on the embankment slope, along the toe of the foundation or at some distance beyond the toe. Generally, underseepage that exits closer to the toe of the embankment is potentially more critical than seepage that initially exits at some distance away.

3 ARTIFICIAL NEURAL NETWORKS (ANNS) OVERVIEW

Artificial neural networks (ANNs) is a computational approach inspired by the human nervous system [22, 23]. Its data-processing paradigm is made up of highly interconnected nodes (neurons) that map a complex input pattern with a corresponding output pattern. They are known for their generalization ability and ease of working with highly dimensional and nonlinear data.

There are many types of neural network structures (architectures) and training algorithms, e.g., generalized regression neural networks (GRNNs), back-propagation network (BPs), etc. This section briefly describes the above two neural networks.

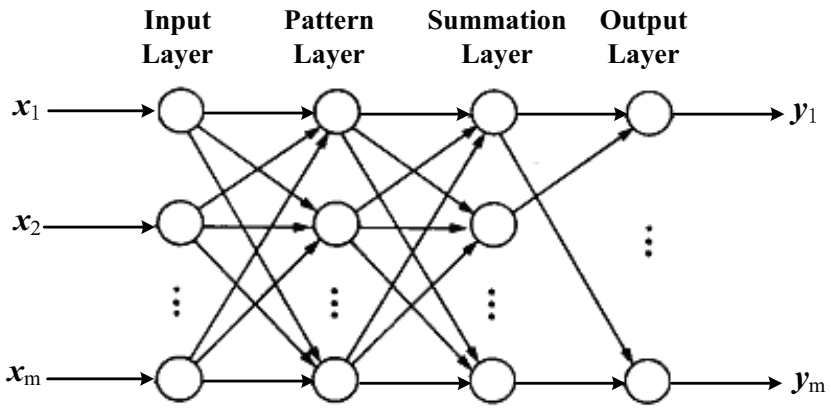

Figure 4. Structure of GRNN.

3.1 GENERALIZED REGRESSION NEURAL NETWORKS (GRNNS)

In 1991, Specht [24] proposed the generalized regression neural network (GRNN) as a variation of the radial basis function kernel (RBF) neural network, which approximates a function between the input and output vectors based on a standard statistical technique called kernel regression. The training time for the GRNN is relatively short, as the bandwidths of all the parameters are calculated, unlike the initial setting of the learning parameters. As the size of training set becomes larger, the estimation error approaches zero, while only minor restrictions are imposed on the function. The typical structure of a GRNN is shown in Fig.4.

Assume that $f(x,y)$ represents the known joint continuous probability density function of a vector random variable, x, and a scalar random variable, y. Let $\boldsymbol{X}$ be a particular measured value of the random variable x. The conditional mean of y given $\boldsymbol{X}$ (also called the regression of y on $\boldsymbol{X}$ is given by

$$E[y \mid X] = \frac{\int_{-\infty}^{+\infty} y f(X,y)\,\mathrm{d}y}{\int_{-\infty}^{+\infty} f(X,y)\,\mathrm{d}y} \qquad (1)$$

When the density $f(x,y)$ is not known, it must usually be estimated from a sample of observations of x and y. For a nonparametric estimate of $f(x,y)$, we will use the class of consistent estimators proposed by Parzen [25] and shown to be applicable to the multi-dimensional case by Cacoullos [26]. As noted in Specht [27], these estimators are a good choice for estimating the probability density function, if it can be assumed that the underlying density is continuous and that the first partial derivatives of the function evaluated at any x are small. The probability estimator $\hat{f}(X,Y)$ is based upon the sample values X_i and Y_i of the random variables x and y, where n is the number of sample observations and p is the dimension of the vector variable x:

$$\hat{f}(X,Y) = \frac{1}{(2\pi)^{\frac{p+1}{2}}\sigma^{(p+1)}n}\sum_{i=1}^{n}\exp\left[-\frac{(X-X_i)^{\mathrm{T}}(X-X_i)}{2\sigma^2}\right]\cdot\exp\left[\frac{(Y-Y_i)^2}{2\sigma^2}\right] \qquad (2)$$

A physical interpretation of the probability estimate $\hat{f}(X,Y)$ is that it assigns a sample probability of width σ for each sample X_i and Y_i, and the probability estimate is the sum of those sample probabilities. Substituting the joint-probability estimate $\hat{f}(X,Y)$ in Equation (2) into the conditional mean, (1), gives the desired conditional mean of y given $\boldsymbol{X}$. In particular, combining Equation (1) and (2) and interchanging the order of integration and summation yields the desired conditional mean, designated $\hat{Y}(X)$:

$$\hat{Y}(X) = \frac{\sum_{i=1}^{n}\exp\left[-\frac{(X-X_i)^{\mathrm{T}}(X-X_i)}{2\sigma^2}\right]\int_{-\infty}^{+\infty} y\cdot\exp\left[-\frac{(y-Y_i)^2}{2\sigma^2}\right]\mathrm{d}y}{\sum_{i=1}^{n}\exp\left[-\frac{(X-X_i)^{\mathrm{T}}(X-X_i)}{2\sigma^2}\right]\int_{-\infty}^{+\infty}\exp\left[-\frac{(y-Y_i)^2}{2\sigma^2}\right]\mathrm{d}y} \qquad (3)$$

Defining the scalar function D_i^2,

$$D_i^2 = (X-X_i)^{\mathrm{T}}(X-X_i) \qquad (4)$$

and performing the indicated integrations yields the following:

$$\hat{Y}(X) = \frac{\sum_{i=1}^{n} Y_i \exp\left(-\frac{D_i^2}{2\sigma^2}\right)}{\sum_{i=1}^{n}\exp\left(-\frac{D_i^2}{2\sigma^2}\right)} \qquad (5)$$

The resulting regression, (5), which involves summations over the observations, is directly applicable to problems involving numerical data.

The estimate $\hat{Y}(X)$ can be visualized as a weighted average of all of the observed values, Y_i, where each observed value is weighted exponentially according to its Euclidean distance from $\boldsymbol{X}$. When the smoothing parameter σ is made large, the estimated density is forced to be smooth and in the limit becomes a multivariate Gaussian with the covariance $\sigma^2\boldsymbol{I}$. On the other hand, a smaller value of σ allows the estimated density to assume non-Gaussian shapes, but with the hazard that wild points may have too great an effect on the estimate. As σ becomes very large, $\hat{Y}(X)$ assumes the value of the sample mean of the observed Y_i, and as σ goes to 0, $\hat{Y}(X)$ assumes the value of the Y_i associated with the observation closest to $\boldsymbol{X}$. For intermediate values of σ, all the values of Y_i are taken into account, but those corresponding to points closer to $\boldsymbol{X}$ are given greater weight.

When the underlying parent distribution is not known, it is not possible to compute an optimum σ for a given number of observations n. It is therefore necessary to find σ on an empirical basis. This can be done easily when the density estimate is being used in a regression equation because there is a natural criterion that can be used for evaluating each value of σ, i.e., the mean-squared error between Y_i and the estimate $\hat{Y}(X_i)$. This procedure is used to avoid an artificial minimum error as $\sigma \to 0$ that results when the estimated density is allowed to fit the observed data points, that is

$$E=\frac{1}{n}\sum_{i=1}^{n}\left[\hat{Y}\left(X_i\right)-Y_i\right]^2 \qquad (6)$$

The principal advantages of the GRNN are the rapid learning and the convergence to the optimum regression surface as the number of samples becomes very large. The GRNN is particularly advantageous with sparse data in a real-time environment, because the regression surface is instantly defined everywhere, even with just one sample.

3.2 BACK-PROPAGATION NETWORK (BP)

The back-propagation neural is a multilayered, feed-forward neural network and is by far the most extensively used [28]. Back-propagation works by approximating the non-linear relationship between the input and the output by adjusting the weight values internally. A supervised learning algorithm of back-propagation is utilized to establish the neural network modelling. A normal back-propagation (BP) neural model consists of an input layer, one or more hidden layers, and an output layer (Fig.5).

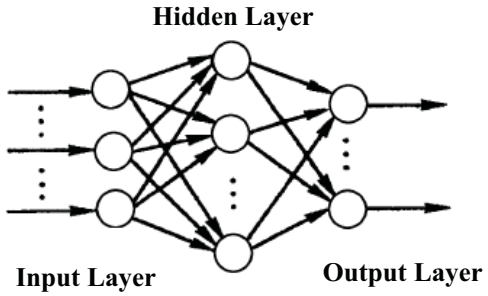

Figure 5. The back-propagation neural network model.

The BP technique is only outlined in this section, and for more details, the reader is referred, for example, to Freeman and Skapura [29].

1) Apply the input vector, $\mathbf{X}_p=\left(x_{p1},x_{p2},\cdots,x_{pN}\right)^{\mathrm{T}}$ to the input units.

2) Calculate the net-input values to the hidden layer units:
$$\mathrm{net}_{pj}^{h}=\sum_{i=1}^{N}w_{ji}^{h}x_{pi}+\theta_j^h \qquad (7)$$

3) Calculate the outputs from the hidden layer:
$$i_{pj}=f_j^h\left(\mathrm{net}_{pj}^{h}\right) \qquad (8)$$

4) Move to the output layer. Calculate the net-input values to each unit:
$$\mathrm{net}_{pk}^{o}=\sum_{j=1}^{L}w_{kj}^{o}i_{pj}+\theta_k^o \qquad (9)$$

5) Calculate the outputs:
$$o_{pk}=f_k^o\left(\mathrm{net}_{pk}^{o}\right) \qquad (10)$$

6) Calculate the error terms for the output units:
$$\delta_{pk}^{o}=\left(y_{pk}-o_{pk}\right)f_k^{o'}\left(\mathrm{net}_{pk}^{o}\right) \qquad (11)$$

7) Calculate the error terms for the hidden units:
$$\delta_{pj}^{h}=f_j^{h'}\left(\mathrm{net}_{pj}^{h}\right)\sum_k\delta_{pk}^{o}w_{kj}^{o} \qquad (12)$$
Notice that the error terms on the hidden units are calculated before the connection weights to the output layer units have been updated.

8) Update weights on the output layer:
$$w_{kj}^{o}\left(t+1\right)=w_{kj}^{o}\left(t\right)+\eta\delta_{pk}^{o}i_{pj} \qquad (13)$$

9) Update weights on the hidden layer:
$$w_{ji}^{h}\left(t+1\right)=w_{ji}^{h}\left(t\right)+\eta\delta_{pj}^{h}x_i \qquad (14)$$

where i is the unit node in the input layer, j is the unit node in the hidden layer, p is the pattern and k is related to the output layer, η is the learning rate, t is the t^{th} iteration, δ_{pk}^{o} and δ_{pj}^{h} are the error terms (which encompass a derivative part) for the output units and hidden units, respectively. w_{ji}^{h} and w_{kj}^{o} are the weights for the hidden units and output units, respectively.

4 CROSS-VALIDATION

Cross-validation is a technique for assessing how the results of statistical analyses can be generalized to an independent dataset. This technique is mainly used in situations where the goal is prediction, and one wants to estimate how accurately a predictive model will perform in practice [30].

This study adopts a *k*-fold cross-validation technique that randomly partitions the original sample into *k* subsamples. A single subsample is retained as validation data for testing the model, and the remaining *k*—1 subsamples are used as training data. The cross-validation process is repeated *k* times (the folds), with each of the *k* subsamples used exactly once as the validation data. The *k* results from the folds can be averaged (or otherwise combined) to produce a single estimation. Fig.6 provides an example of a *k*-fold cross-validation procedure.

Fig.7 shows the *k*-fold cross-validation error versus *k* for a large dataset, and indicates that a *k* value between 4

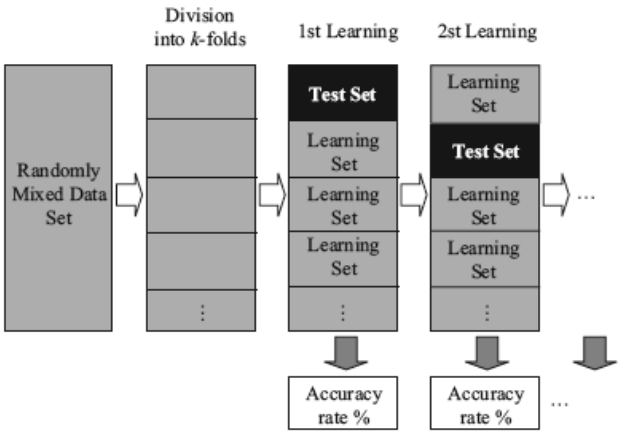

Figure 6. A *k*-fold cross-validation procedure [30].

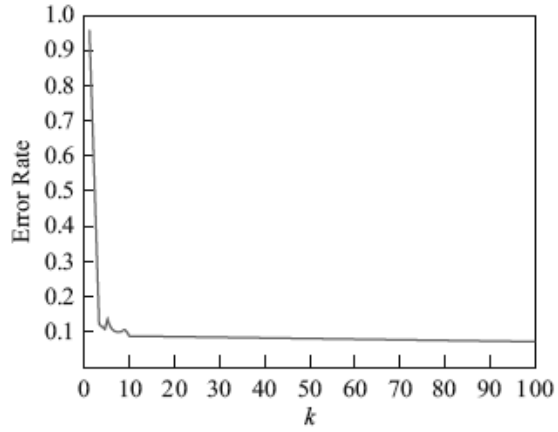

Figure 7. A plot of *k*-fold cross-validation errors vs. *k* [31].

and 10 is a good trade-off: increasing this value significantly increases the computation time and does not significantly improve the results [31]. Thus, this study adopts a 10-fold cross-validation. This approach may not be useful in achieving high training accuracy, but it can prevent the over-fitting problem.

In order to select the optimum value of σ, Specht [24] suggested the holdout method. The holdout method is the simplest kind of cross-validation. The data set is separated into two sets, called the training set and the testing set. The function approximator fits a function using the training set only. Then the function approximator is asked to predict the output values for the data in the testing set (it has never seen these output values before). The errors it makes are accumulated as before to give the mean absolute test-set error, which is used to evaluate the model. The advantage of this method is that it is usually preferable to the residual method and takes no longer to compute. However, its evaluation can have a high variance. The evaluation may depend heavily on which data points end up in the training set and which end up in the test set, and thus the evaluation may be significantly different depending on how the division is made. As a step further, in this paper the 10-fold cross-validation method was used. After the selection of a fixed value of σ, the model was trained using the training dataset. Then the model was evaluated using the test dataset. This process was repeated for each k=10 folds (90% for training and 10% for testing) and then many times using different values of the smoothing coefficient. Finally, the best value of σ that should be used was selected as the value that minimized the error that was estimated as the average error rate on these cross-validation examples. The advantage of this method over repeated random subsampling (which is also known as the Monte Carlo variation, meaning that the results will vary if the analysis is repeated with different random splits) is that all the observations are used for both training and validation, and each observation is used for the validation exactly once. The main drawback of this method is that it requires intense computation.

5 CASE STUDIES

5.1 MODEL VALIDATION WITH EXPERIMENTAL DATA

Model development (i.e., model "training") should be followed by a model validation to assess the predictive capability [32]. To evaluate the overall performance of the GRNN model, we compared the predictions against the experimental data. The case records listed in Table 1 are evaluated using the GRNN model. The database includes 25 experimental data from ref. [33]. And then, seven selected factors, which include various characteristic particle sizes, such as d_{85}, d_{15}, d_5, d_{60}, d_{30}, d_{10}, and the porosity Φ of non-cohesive soils, were taken into account as the input parameters of the GRNN model. Moreover, the seepage failure types of non-cohesive soils are also listed in Table 1; it is 1 for piping and 0 for soil flow.

As can be seen from Table 1, the predictions from the GRNN model agree well with the experimental data, and it has a high success rate (100%). These results indicate that our GRNN model seems to be a good tool for predicting the piping potential.

Table 1. Experimental data.

	No.	d_{85}/mm	d_{15}/mm	d_5/mm	d_{60}/mm	d_{30}/mm	d_{10}/mm	Φ	Test results	Predicted results
Calibration	1	7.00	0.37	0.20	3.50	0.60	0.29	0.44	0	0.0124
	2	5.90	0.29	0.15	2.10	0.50	0.20	0.28	0	0.0235
	3	4.40	0.15	0.11	2.70	1.10	0.13	0.28	1	0.9891
	4	5.10	0.22	0.12	2.60	1.13	0.14	0.28	1	0.9867
	5	7.00	0.21	0.12	4.10	2.60	0.16	0.28	1	0.9895
	6	6.68	0.21	0.11	4.00	2.10	0.16	0.28	1	0.9989
	7	5.00	0.20	0.11	2.20	1.05	0.13	0.34	1	0.9856
	8	5.60	0.25	0.12	3.30	2.00	0.20	0.31	1	0.9687
	9	4.30	0.19	0.12	3.70	2.30	0.15	0.35	1	0.9888
	10	5.00	0.26	0.16	3.50	2.10	0.18	0.29	1	0.9886
	11	5.00	0.30	0.11	3.20	2.10	0.18	0.33	1	0.9883
	12	5.00	0.23	0.15	3.10	2.10	0.17	0.34	1	0.9889
	13	7.00	0.25	0.17	4.10	2.30	0.19	0.29	1	0.9888
	14	6.00	0.20	0.15	3.40	2.00	0.17	0.30	1	0.9885
	15	5.00	0.21	0.15	3.00	1.70	0.18	0.30	1	0.9887
	16	7.50	0.23	0.16	2.00	0.40	0.19	0.28	0	0.0104
	17	5.30	0.23	0.16	2.30	0.50	0.19	0.30	0	0.0125
	18	0.42	0.14	0.04	0.37	0.24	0.12	0.40	0	0.0061
	19	0.30	0.02	0.01	0.12	0.06	0.02	0.43	0	0.0053
	20	0.19	0.02	0.01	0.10	0.07	0.02	0.39	0	0.0105
Validation	21	0.22	0.05	0.01	0.16	0.09	0.03	0.38	0	0.0054
	22	0.18	0.02	0.01	0.10	0.07	0.02	0.42	0	0.0015
	23	0.22	0.05	0.01	0.19	0.11	0.05	0.41	0	0.0031
	24	0.21	0.05	0.01	0.17	0.09	0.04	0.40	0	0.0073
	25	0.22	0.07	0.01	0.20	0.13	0.04	0.38	0	0.0112

5.2 PRACTICAL PROJECTS

In this section, an attempt has been made to predict the seepage piping of embankments with the help of the GRNN and BP. With chosen network parameters, 23 case studies containing the most necessary information of embankment structures, materials and states of seepage piping are filtered and selected from ref. [34, 35]. The 23 case records are shown in Table 2. A total of nine selected factors, which include the height of the

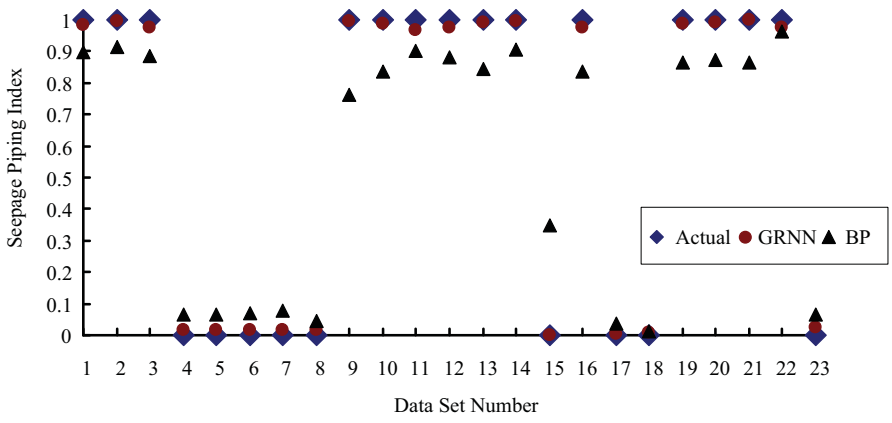

Figure 8. Comparison between forecasted and actual results.

Table 2. Practical projects data.

	No.	H/m	H_w/m	λ	c /kPa	φ/°	γ_{sat}/kN/m^3	k/cm/s	d_b/cm	δ/°	Actual
Calibration	1	133	123	0.455	40	27.0	21.3	3.0×10^{-7}	0.009	79.0	1
	2	87.5	80	0.4	12	28.9	21.2	3.5×10^{-7}	0.008	75.0	1
	3	35.5	31	0.295	60	15.6	20.5	1.0×10^{-7}	0.008	43.0	1
	4	31	29	0.249	20	26.7	20.8	4.0×10^{-4}	0.01	14.0	0
	5	31	29	0.249	15	26.7	20.1	7.8×10^{-5}	0.009	14.0	0
	6	29	25	0.435	30	31.6	20.8	2.0×10^{-5}	0.01	23.5	0
	7	39	35.5	0.466	109	21.2	20.7	5.1×10^{-6}	0.04	29.0	0
	8	39	35.5	0.466	76	13.8	20.7	5.1×10^{-6}	0.04	25.0	0
	9	28	25	0.286	157	13.2	20.3	3.6×10^{-7}	0.009	60.0	1
	10	28	25	0.286	153	24.8	21.2	4.8×10^{-8}	0.009	60.0	1
	11	96	90	0.417	20	26.0	21.0	3.5×10^{-8}	0.012	65.0	1
	12	56	49	0.364	30	29.0	19.5	2.0×10^{-7}	0.013	59.0	1
	13	51	47	0.308	42	34.5	21.2	2.2×10^{-8}	0.012	66.0	1
	14	133.1	126	0.476	41	32.0	21.7	1.3×10^{-6}	0.004	76.0	1
	15	13.0	10.5	0.364	44	38.4	22.9	1.0×10^{-2}	0.02	26.6	0
	16	6.7	5.5	0.4	109	21.2	20.7	5.1×10^{-6}	0.004	67.5	1
	17	6.0	4.75	0.5	51	38.5	20.9	7.0×10^{-3}	0.024	26.6	0
	18	87.5	80	0.256	14	27.0	21.2	4.0×10^{-6}	0.008	28.0	0
	19	51.5	46	0.455	100	19.3	21.0	2.9×10^{-6}	0.007	45.0	1
Validation	20	39.5	33	0.347	32	27.2	20.2	5.5×10^{-8}	0.01	67.0	1
	21	29.0	26	0.315	26	27.8	20.8	2.0×10^{-5}	0.01	65.0	1
	22	42.5	39	0.361	84	32.2	20.6	1.3×10^{-8}	0.004	53.1	1
	23	7.0	5.6	0.4	20	30.0	19.5	3.8×10^{-5}	0.017	26.6	0

Table 3. Comparison between forecasted and actual results.

No.	Actual	GRNN	BP
1	1	0.9845	0.8967
2	1	0.9971	0.9125
3	1	0.9764	0.8834
4	0	0.0164	0.0657
5	0	0.0183	0.0637
6	0	0.0171	0.0681
7	0	0.0157	0.0764
8	0	0.0162	0.0452
9	1	0.9967	0.7634
10	1	0.9863	0.8346
11	1	0.9657	0.9023
12	1	0.9753	0.8827
13	1	0.9912	0.8424
14	1	0.9961	0.9068
15	0	0.0013	0.3467
16	1	0.9735	0.8359
17	0	0.0054	0.0357
18	0	0.0067	0.0139
19	1	0.9893	0.8641
20	1	0.9931	0.8744
21	1	0.9987	0.8655
22	1	0.9734	0.9615
23	0	0.0258	0.0637

dam H, the water level H_w, the downstream slope ratio λ, the effective internal frictional angle φ, the effective cohesion c, the saturated unit weight γ_{sat}, the coefficient of permeability k, the maximum effective particle size d_b and the inclination angle of the seepage filter δ, were taken into account as the input parameters for the artificial neural network systems. The neural network output consisted of a single neuron representing the piping potential. The neuron was given a binary value of 1 for seepage piping occurrence. A comparison between the forecasted and actual results is shown in Table 3 and Fig.8. From the results obtained here, it can be said that the GRNN demonstrates superiority over the BP technology. These results suggest that the GRNN is potentially useful for predicting seepage-piping occurrence.

6 CONCLUSIONS

Piping is known to cause the catastrophic failures of levees and earthen dams and has been studied for nearly a century. The factors that influence the piping occurring in embankments are uncertain and random, and in the process of piping they turn out to be a

nonlinear behaviour. In order to estimate the risks of dam failure phenomena, it is necessary to develop scientifically derived analytical models that are simpler, easier to implement, and more accurate than traditional methods for the evaluation of piping potential. In this paper, the GRNN technique was chosen to predict the seepage-piping potential. The performance of the GRNN was cross-validated using the *k*-fold cross-validation technique. From these studies, the following conclusions and recommendations may be drawn:

(1) The above results show that the GRNN model is efficient for predicting the occurrence of seepage piping with high accuracy and it is expected to be very reliable for the evaluation of the piping potential. Thus, the GRNN is a powerful piping-assessment tool worthy of promotion and support.

(2) The main factors that affect the erosion phenomenon are: 1) the erodibility of the soil; 2) the water velocity inside the soil mass; 3) the geometry of the earth structure, and so on. As can be seen, the mechanism of piping is very complicated and needs to be further investigated through laboratory and field tests. Meanwhile, the advanced digital photographic video tracing technique, image processing technology, etc., should also be introduced into the mechanism research of piping phenomenon.

(3) The performance of the GRNN is dependent on the smoothing parameter, and the appropriate smoothing parameter can minimize the misclassification rate or the error of the final network. Thus, to improve the prediction performance of GRNN models, it is necessary to study the optimization algorithms in future investigations.

REFERENCES

[1] Xu, Y., Zhang, L.M.(2009). Breaching parameters of earth and rockfill dams, J. Geotech. Geoenviron. Eng., ASCE 135, 12, 1957-1970.

[2] Foster, M., Fell, R., Spannagle, M. (2000). The statistics of embankment dam failures and accidents, Can. Geotech. Journal 37, 5, 1000-1024.

[3] Zhang, L.M., Xu, Y., Jia, J.S. (2009). Analysis of earth dam failures-A database approach, Georisk 3, 3, 184 -189.

[4] Richards, K.S., Reddy, K.R. (2012). Experimental investigation of initiation of backward erosion piping in soils, Géotechnique 62, 10, 933-942.

[5] Terzaghi, K., Der. (1922). Grundbruch an Stauwerken und seine Verhütung, In From theory to practice in soil mechanics (1960), New York, NY, USA: John Wiley & Sons (in German), 445-449.

[6] Bligh, W.G. (1910).Dams, barrages and weirs on porous foundations, Eng. News 64, 26, 708-710.

[7] Lane, E.W. (1934). Security from under-seepage masonry dams on earth foundations, Trans. ASCE, 100, 1, 1235-1272.

[8] Kevin, S.R., Krishna, R.R. (2007). Critical appraisal of piping phenomena in earth dams, Bull. Eng. Geol. Environ. 66, 381-402.

[9] Xue, X.H.,Yang, X.G. (2013). Application of the adaptive neuro-fuzzy inference system for prediction of soil liquefaction, Nat. Hazards 67, 901-917.

[10] Saman, Y.S., Hing, H.T. (2011). A new site classification approach based on neural networks, Soil Dyn. Earthq. Eng., 31, 974-981.

[11] Kamali, M., Ataei, M. (2011). Prediction of blast induced vibrations in the structures of Karoun III power plant and dam, J. Vib. Control 17, 4, 541-548.

[12] Tesfamariam, S., Liu, Z. (2010). Earthquake induced damage classification for reinforced concrete buildings, Struct. Saf. 32, 2, 154-164.

[13] Senouci, A.B., Adeli, H. (2001). Resource scheduling using neural dynamics model of Adeli and Park, J. Constr. Eng. M., ASCE 127, 1, 28-34.

[14] Yang, Y., Rosenbaum, M.S. (2002). The artificial neural network as a tool for assessing geotechnical properties, Geotech. Geol. Eng. 20, 149-168.

[15] Cichocki, A., Zdunek, R. (2007). Multilayer nonnegative matrix factorization using projected gradient approaches, Int. J. Neural Syst. 17, 6, 431-446.

[16] Lakshmi, S.S., Tiwari, R.K. (2009). Model dissection from earthquake time series: a comparative analysis using modern non-linear forecasting and artificial neural network approaches, Comput. Geosci. 35, 2, 191-204.

[17] Fell, R., Fry. J.J. (2007). The state of art of assessing the likelihood of internal erosion of embankment dams, water retaining structures and their foundations, Internal Erosion of Dams and their Foundations, Taylor & Francis, London.

[18] Foster, M.A., Fell, R. (1999).A framework for estimating the probability of failure of embankment dams by piping using event tree methods. UNICIV Report No.R-377. The University of New South Wales, Sydney, Australia.

[19] Rice, J.D., Duncan, J.M. (2007).Davidson, R.R.: Identification of failure mechanisms associated with seepage barriers in dams. GSP 161 Embankments, Dams, and Slopes Geo-Denver 2007: New Peaks in Geotechnics, 1-11.

[20] Rice, J.D., Duncan, J.M. (2008).A study on the long-term performance of seepage barriers in dams. Virginia Tech Center for Geotechnical Practice and Research Publication 49.
[21] Rice, J.D., Duncan, J.M. (2010). Findings of case histories on the long-term performance of seepage barriers in dams. Journal of Geotechnical and Geoenvironmental Engineering, ASCE 136, 1, 2-15.
[22] Gholam, R.R., Mohammad, V., Mehdi, A.A. (2012). Forecasting groundwater level in Shiraz Plain using artificial neural networks, Arab.J.Sci.Eng. 37, 1871-1883.
[23] Seyed, S.E., Seyed, A.G.,Mohammad, J.Z., Alireza, F. (2012). Estimating Penman-Monteith reference evapotranspiration using artificial neural networks and genetic algorithm: a case study, Arab.J.Sci.Eng. 37, 935-944.
[24] Specht, D.F. (1991). A general regression neural network, IEEE T. Neural Network 2, 6, 568 -576.
[25] Parzen, E. (1962).On estimation of a probability density function and mode. Annals of Mathematical Statistics 33,1065-1076.
[26] Cacoullos, T. (1966). Estimation of a multivariate denstiy. Annals of the Institute of Statistical Mathematics 18, 2,179-189.
[27] Specht, D. F. (1967). Generation of polynomial discriminant functions for pattern recognition. IEEE Transactions on Electronic Computers (EC-16),308-319.
[28] Lee, M.C., To Chang (2010). Comparison of support vector machine and back propagation neural network in evaluating the enterprise financial distress, International Journal of Artificial Intelligence & Applications 1, 3, 31-43.
[29] Freeman, A.J., Skapura, M.D. (1991). Neural networks: Algorithms, Application and Programming Techniques. Addition Wesley Publications Company, Inc.
[30] Lee, C.Y., Chern, S.G. (2013). Application of a support vector machine for liquefaction assessment, J.Mar.Sci. Tech. 21, 3, 318-324.
[31] Sterkubm, P. (2007). Overfitting prevention with cross-validation, Paris.
[32] Oommen, T., Baise, L.G.,Vogel, R. (2010). Validation and application of empirical liquefaction models, J.Geotech.Geoenviron.Eng. 136, 12, 1618-1633.
[33] Zhao, Z.X., Chen, J.S., Chen L. (2008). Application of BP neural network to assessment of noncohesive piping-typed soils, Chinese J. Geot. Eng., 30, 4, 536-540. (in Chinese)
[34] Liu, J. (1992). Soil seepage stability and seepage control, China WaterPower Press, Beijing. (in Chinese)
[35] Zheng, W.L., Shen, M.S. (1995).A study of piping in earth dams, J. of Chinese Institute Civil and Hydraulic Engineering 7, 3, 283- 295. (in Chinese)