

# Improving predictions made by ANN model using data quality assessment: an application to local scour around bridge piers

Sung-Uk Choi, Byungwoong Choi and Seongwook Choi

## ABSTRACT

The scour depth that develops around bridge piers is known to be related to flow intensity, particle size of bed sediment, and pier dimensions. Earlier approaches to this issue have mainly relied on empirical formulas. Even numerical simulations have not been so successful due to problems associated with interactions between water flow and streambed morphology. This necessitates the application of an artificial intelligence (AI)-type approach to understanding the effects of local scour around bridge piers. Although previous studies reported that AI-based models are better predictors, they do not predict field-scale local scour well. Motivated by this, the present study reports on the use of data quality assessment with an artificial neural network (ANN) model for predicting field-scale scour depth around bridge piers. Both univariate and multivariate methods were applied and the predicted results are compared. For the multivariate method, the Euclidean distance method and Mahalanobis distance method were used and the predicted results are compared. The ANN model was first trained and validated using laboratory data and the model was applied to data obtained in laboratory experiments. The model was then applied to field data. Quantitative descriptions are given on how much the data quality assessment improves predictions based on the use of the ANN model.

**Key words** | artificial neural network, bridge piers, data quality assessment, Euclidean distance method, local scour, Mahalanobis distance method

Sung-Uk Choi (corresponding author)  
Byungwoong Choi  
Seongwook Choi  
Department of Civil & Environmental Engineering,  
Yonsei University,  
Seoul,  
Korea  
E-mail: [schoi@yonsei.ac.kr](mailto:schoi@yonsei.ac.kr)

## INTRODUCTION

Stream flows tend to accelerate locally when they are obstructed by in-stream structures such as bridge piers. Thus, holes develop around the structure as a result of scouring. Such scour holes increase in size with increasing discharge, which threatens the safety of the structure particularly during periods of flooding. If sediment particles on the stream bed are cohesive, the scour holes are likely to be affected by a series of floods, not by a single flood (Briaud *et al.* 2001; Brandimarte *et al.* 2006). In addition, highly three-dimensional fluid dynamics phenomena are involved in the mechanism by which the scour hole is created (Melville & Coleman 2000). All of these render an

analytical approach to local scour around bridge piers nearly impossible.

The conventional approach to local scour around bridge piers is typically based on dimensional analysis using experimental observations. Several formulas have been developed (Melville & Coleman 2000), but a generally accepted formula is not available. An alternative approach is to use computational fluid dynamics techniques. Although it is true that this approach is capable of revealing detailed mechanisms regarding local acceleration around a cylinder (Ge & Sotiropoulos 2005; Ge *et al.* 2005; Kirkil *et al.* 2009; Baranya *et al.* 2012), applications to interactions between a

fluid and bed morphology are extremely limited. The other approach is to use artificial intelligence (AI) techniques such as the artificial neural network (ANN), fuzzy logic, a genetic algorithm, or an adaptive neuro fuzzy inference system (ANFIS) method. Several previous studies have shown that AI-based models are better predictors of scour depth than empirical formulas provided that the methods are properly trained (Bateni & Jeng 2007; Firat & Gungor 2009; Azamathulla *et al.* 2010; Keshavarzi *et al.* 2012; Khan *et al.* 2012; Najafzadeh & Azamathulla 2013).

Scour depth is known to be related to flow intensity, the properties of bed sediment, and pier dimensions (Melville & Coleman 2000). Flow intensity includes factors such as water depth and the approach velocity of stream flow, and the property of bed sediment includes particle size and the critical velocity for the initiation of motion. Pier dimension is reflected by the pier width. In the laboratory, all of the above variables can be controlled and can be measured more or less precisely. However, in the field, all of the above variables, except for the pier width, are difficult to measure correctly. The situation becomes more serious with regard to scour depth. That is, measurements that are made earlier or later than the equilibrium state can result in a scour depth less than the equilibrium scour depth. Moreover, if the scour hole has not been recovered from an earlier flood, then the measured scour depth may be larger than the equilibrium scour depth.

Several formulas are available for predicting local scour around bridge piers (Sheppard *et al.* 2014). Most existing formulas are based on small-scale laboratory or in the field observations. It is quite well known that these empirical formulas do not always predict well, especially when they are applied to field conditions (Dargahi 1990; Johnson 1995; Gaudio *et al.* 2010). This is partly due to the poor quality of field data. Sheppard *et al.* (2011, 2014) successfully identified outliers in data of scour depths obtained in both the laboratory and field, and proposed a new relationship. However, since the AI-based models predict the scour depth directly using the data without establishing a relationship, they are more vulnerable to the quality of data. This necessitates the introduction of data quality assessment to AI-based models.

The objective of this study was to improve the ANN model for predicting local scour around bridge piers using

data quality assessment. Regarding methods for assessing data quality, both univariate and multivariate methods were used to identify outliers in the dataset. The methods were then applied to data obtained in both the laboratory and in the field. It was found that the data quality assessment improved the ANN model for predicting local scours in both laboratory and field at a similar level.

## LOCAL SCOUR AROUND BRIDGE PIERS

The local scour around bridge piers is a time-dependent process, in which two characteristic scour depths can be defined (Raudkivi & Ettema 1983; Melville & Chiew 1999). Under mobile-bed conditions, the scour depth increases with the approach velocity, showing the maximum scour depth ( $d_{sm}$ ). This is the stage of the clear water scour because the shear stress away from the pier is less than its critical value. After this peak, the live-bed scour or scour with sediment transport occurs. That is, the scour depth decreases slightly with increasing velocity and the equilibrium scour depth ( $d_{se}$ ) is reached. It has been reported that the equilibrium scour depth is about 90% of the maximum scour depth (Breusers *et al.* 1977; Raudkivi & Ettema 1983). Under live-bed conditions, the scour depth tends to fluctuate with time, which is due to bed form migration (Raudkivi & Ettema 1983; Melville & Chiew 1999). The scour depth averaged over a period is called the equilibrium scour depth. Existing formulas for local scour are to predict the equilibrium scour depths.

The highly 3D complicated flow is known to develop the scour hole locally around bridge piers (Melville & Coleman 2000). Specifically, the downflow ahead of the pier, the horseshoe vortex at the base of the pier, and the wake vortex result in the scour hole. The upstream part of the scour hole is made mostly by the downflow, and the slope at this part is most likely to be steep, being close to the angle of repose of the bed sediment. The scour hole around the upstream part of the pier generates the horseshoe vortex, which transports lifted particles away from the pier. The wake vortex, formed by the separation at the side of the pier, makes the downstream part of the scour hole.

In general, scour depth is related to fluid flow, bed sediment, and the geometry of the pier (Melville & Coleman 2000; Choi & Cheong 2006). Local scour around a cylindrical pier of diameter  $b$  on a bed with non-cohesive uniform sediment particles is considered. If it is assumed that the flow is highly turbulent and the angle of attack is zero, then equilibrium scour depth  $d_{se}$  is given by

$$d_{se} = f_1(V, y, d, V_c, b) \quad (1)$$

where  $V$  is the approach velocity,  $y$  is the flow depth,  $d$  is the particle diameter,  $V_c$  is the critical mean velocity associated with the initiation of particle motion on the bed, and  $b$  is the pier width or the diameter of the pier. A dimensional analysis of Equation (1) leads to the following form:

$$\frac{d_{se}}{b} = f_2\left(\frac{V}{V_c}, \frac{y}{b}, \frac{d}{b}\right) \quad (2)$$

which is similar to the expression in Melville & Chiew (1999). Either Equation (1) or Equation (2) can be used for the ANN model to predict scour depths. However, it has been reported that the use of the dimensional form of the equation results in more accurate predictions (Bateni *et al.* 2007a, b; Muzzammil 2010).

## ARTIFICIAL NEURAL NETWORK

An ANN is a mathematical model inspired by biological neural networks of humans or animals. It is used to approximate functions that can hardly be given by an exact relationship between variables. The ANN is generally presented by a highly interconnected structure consisting of an input layer, hidden layer(s), and an output layer. The number of hidden layers, representing the complexity of the problem, can be varied. The nodes called neurons receive and process the input signals and send an output signal to other nodes. Each node sums the product of every input and its weight and passes the sum through a transfer function to yield outputs. In the learning process, values of weights that connect node to node in each layer are determined. For this process, the standard error back propagation algorithm based on the delta rule is normally used.

## DATA QUALITY ASSESSMENT

In general, two approaches are available for assessing data quality, namely univariate and multivariate methods. The univariate method assumes that variables associated with the given problem are independent of each other, while the multivariate method assumes the existence of a correlation between the variables. Identifying outliers using multivariate methods is known to be statistically more correct because the univariate method may not identify some outliers if the data are correlated (Robinson *et al.* 2005). In the present study, the  $z$ -score method was used in conjunction with the univariate method, and the Euclidean distance method (EDM) and Mahalanobis distance method (MDM) were used in the case of the multivariate method.

A general procedure for assessing data quality for identifying outliers is as follows (Robinson *et al.* 2005; Alameddine *et al.* 2010; Sheppard *et al.* 2014). First, the data are plotted in a dimensional space and statistical properties of the data, including the average, median, mode, variance, and standard deviation are calculated. Statistical parameters such as  $z$ -score, proximity parameter, and the Mahalanobis distance are determined. The statistical parameter versus experiment number is plotted and a cutoff criterion is established. Data that exceed the cutoff criterion are identified. These data are considered to be outliers and are eliminated.

The  $z$ -score denotes the distance a data point is from the mean. The  $z$ -score is simply calculated by  $(x - \mu)/\sigma$  (here,  $x$  is the value of the data,  $\mu$  is the mean, and  $\sigma$  is the standard deviation). To transform non-normal data to normal data, a power series transform such as the Box–Cox method is used (Johnson & Wichern 1998; Robinson *et al.* 2005). The EDM calculates the variance of a dependent variable (measured scour depth in the present study) for the closest five data points (scour variance). The variance of independent variables from the target point to its neighboring four points is determined (distance variance). The ratio of scour variance to distance variance is then computed and is referred to as the proximity parameter. The Mahalanobis distance is a multi-dimensional generalization of how far a point is located from a distribution. The Mahalanobis distance is

calculated by  $(x - \mu)^T \cdot \Sigma^{-1} \cdot (x - \mu)$  (here,  $\Sigma$  denotes the covariance matrix). The Mahalanobis distance denotes the number of standard deviations a point is from the mean of the distribution.

## APPLICATION TO LABORATORY DATA

### Assessment of data quality

Figure 1 shows plots of dimensional variables associated with scour depth measured in the laboratory. In fact, the variables shown in the figure are from three datasets, namely a training dataset, a validation dataset, and an application dataset from Chiew's (1984) experiment. The training data were taken from four different sources of laboratory experiments. They include the experiments of Chabert & Engeldinger (1956), Shen *et al.* (1969), Jain & Fischer (1979), and Dey *et al.* (1995). The total number of data used for training is 64. For the validation, the data from Yanmaz & Altinbilek's (1991) experiments are used. The data from Chiew's (1984) experiments were used for the application of the ANN model. The total number of data for validation and application are 33 and 248, respectively. The ranges of variables for each dataset are listed in Tables 1–3. In the figure, the variables are plotted against other variables. The histograms located diagonally in the figure indicate the distribution of data, and the off-diagonal plots provide the ranges of two independent variables. For example, the second row of the figure shows the velocity with four other variables. It can be seen that, with a few exceptions, most of the velocity data range between 0.2 and 1.8 m/s at water depths less than 0.4 m. It also appears that the critical velocity ranges between 0.2 and 0.4 m/s and the pier width is less than 0.35 m. The diameters of bed sediment particles are less than 5 mm.

The  $z$ -score versus experiment number is plotted in Figure 2 for the univariate method. It can be seen in the figure that most of the data are distributed within a short distance from the origin. To determine the cutoff value, the number of eliminated data versus  $z$ -score is plotted and the value of the cutoff beyond which the number of eliminated data does not change greatly is selected. For the present laboratory data, a cutoff value of 5.0 is selected. This cutoff value eliminated 15 data, 11 from the training dataset and

four from the application dataset. A sensitivity study was carried out for the cutoff value of the data quality assessment. It was found that the selection of the cutoff value does not affect seriously the accuracy of the prediction.

EDM was applied to laboratory data, and the results are presented in Figure 3. In general, it can be seen that most laboratory data are distributed within a short distance from the origin while only a limited number of data are distributed far away from the origin. Specifically, the proximity parameters of the training data are distributed dominantly within 1.0, and those of the application data are also concentrated around 1.0. However, the proximity parameters of the validation data are distributed very uniformly within 0.2. Similarly, the number of eliminated data versus the proximity parameter is plotted, and a cutoff value of 2.0 is selected. This cutoff value eliminated 7 data from the training dataset and 30 from Chiew's dataset.

MDM is now applied to laboratory data, and the resulting Mahalanobis distance versus experiment number is plotted in Figure 4. It can be seen that outliers are more clearly identified by this method compared with the EDM. A cutoff value of 10 for the Mahalanobis distance is selected similarly. This cutoff value eliminated 14 data from the training dataset and 21 data from the application dataset. For every method of data quality assessment, outliers have values of  $V/V_c$  larger than 0.85, indicating high velocities and large scour holes. This is consistent with the results of the analysis in Sheppard *et al.* (2011).

### Training and validation

Training the ANN model using data that cover an appropriate range is important for successful predictions using the ANN model. Outliers in the laboratory data were identified by three different methods, namely one univariate method and two multivariate methods. After eliminating outliers from the data, the ranges of variables of the training dataset changed slightly but not noticeably. Test runs of the ANN model result in the number of nodes in the hidden layer of 15, a learning rate of 0.7, a momentum constant of 0.95, and a system error of 0.001.

The trained ANN model was then applied to Yanmaz & Altinbilek's (1991) dataset for validation. The dataset for training is independent of that for validation, and the

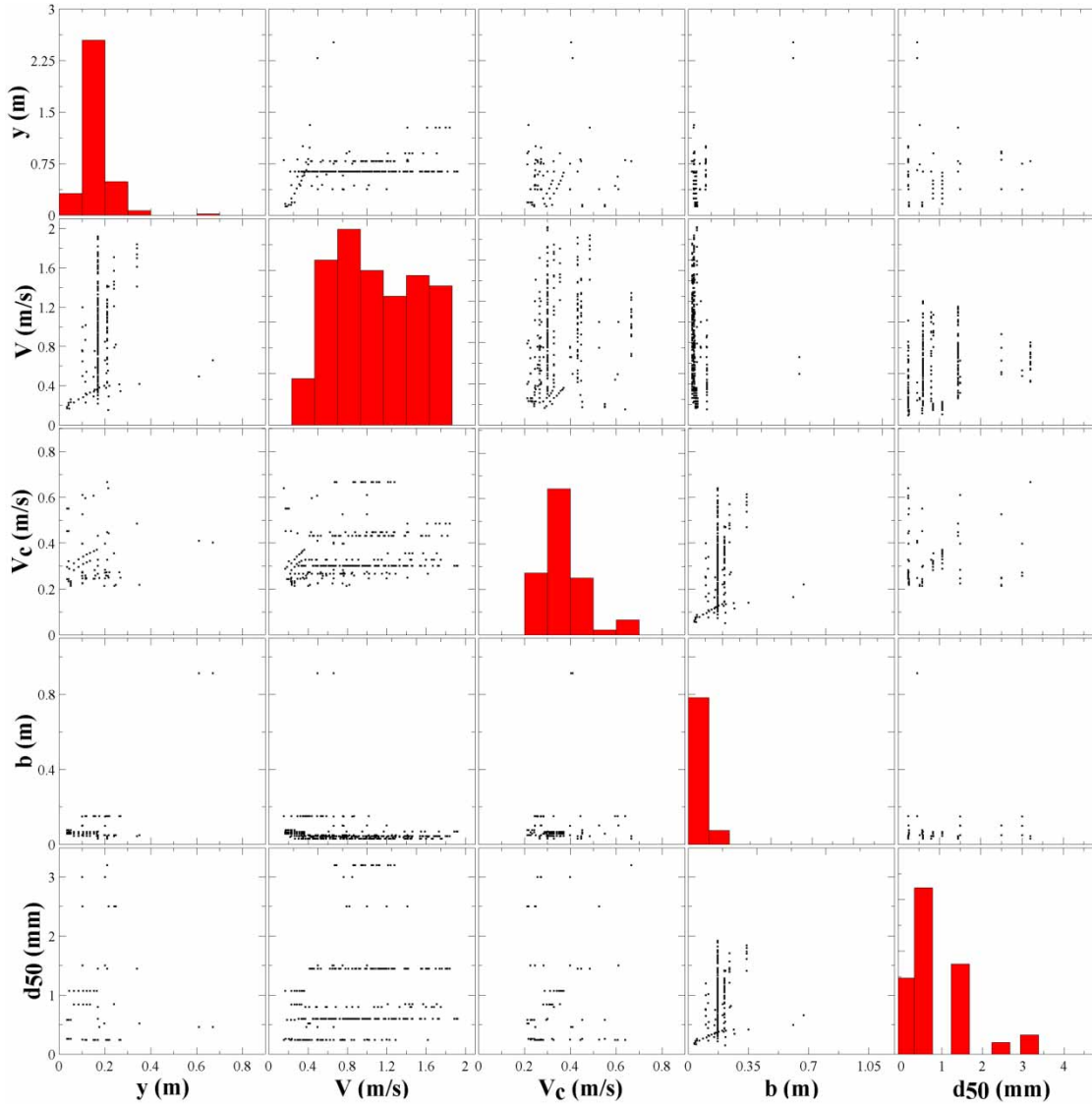


Figure 1 | Plots of dimensional laboratory local scour data.

Table 1 | Ranges of variables used in training

Variables	Range of data
$y$ (m)	0.035 to 0.671
$V$ (m/s)	0.152 to 1.2
$V_c$ (m/s)	0.206 to 0.626
$b$ (m)	0.05 to 0.914
$d_{50}$ (mm)	0.24 to 2.5
$d_{se}$ (m)	0.052 to 0.671

Table 2 | Ranges of variables used in validation (Yanmaz & Altinbilek 1991)

Variables	Range of data
$y$ (m)	0.045 to 0.165
$V$ (m/s)	0.167 to 0.362
$V_c$ (m/s)	0.339 to 0.43
$b$ (m)	0.047 to 0.067
$d_{50}$ (mm)	0.84 to 1.07
$d_{se}$ (m)	0.032 to 0.107

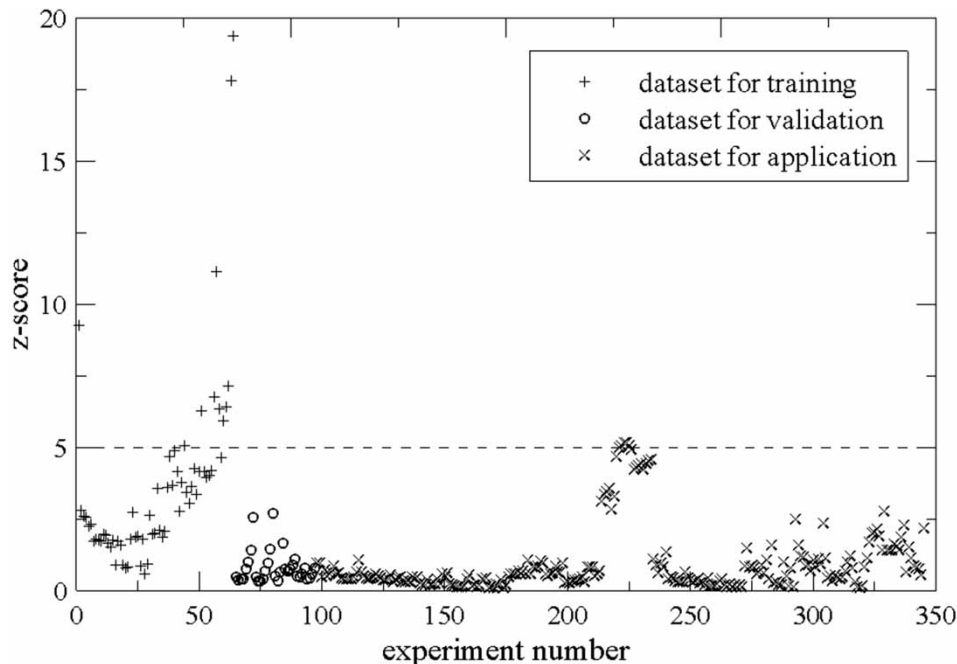
**Table 3** | Ranges of variables used in application (Chiew 1984)

Variables	Range of data
$y$ (m)	0.17 to 0.34
$V$ (m/s)	0.22 to 1.92
$V_c$ (m/s)	0.267 to 0.666
$b$ (m)	0.031 to 0.07
$d_{50}$ (mm)	0.24 to 3.2
$d_{sc}$ (m)	0.0111 to 0.118

ranges of variables of the training dataset include those of the validation dataset. It should be noted that some data for training were eliminated by the data quality assessment but no data for validation were eliminated.

Figure 5 shows the results of model validation. In the figure, the predicted scour depths are plotted against measured data, with a 45 degree line indicating a perfect prediction. For the measurement of the model's performance, the following mean absolute percent error (MAPE) was used:

$$\text{MAPE} = \frac{1}{n} \sum \frac{|d_{sc}^m - d_{sc}^p|}{d_{sc}^m} \times 100(\%) \quad (3)$$

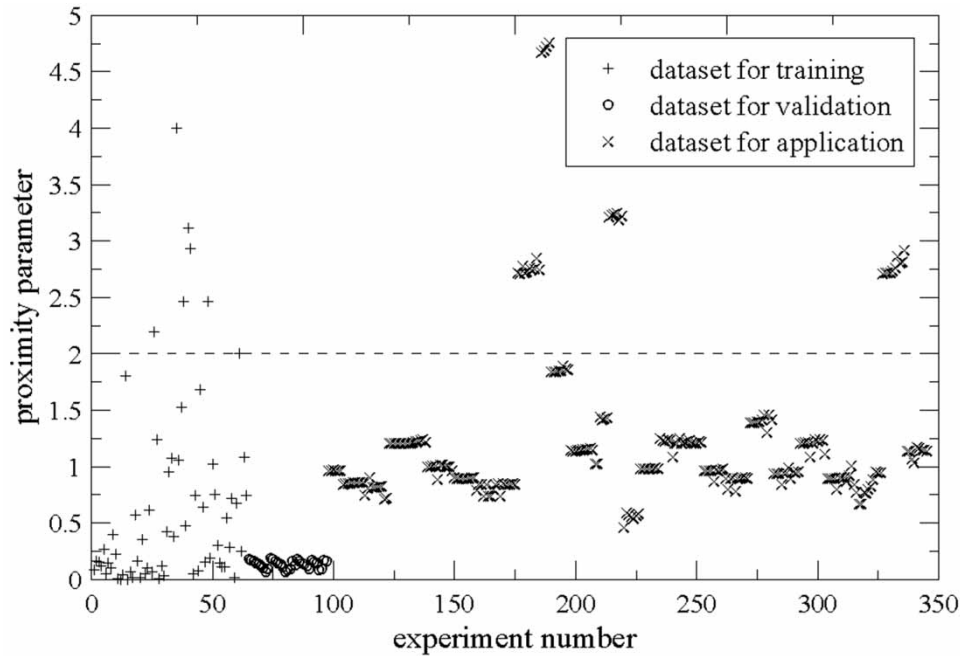
**Figure 2** | Plot of z-scores of the univariate method for laboratory data.

where  $d_{sc}^p$  and  $d_{sc}^m$  are the predicted and measured scour depths, respectively.

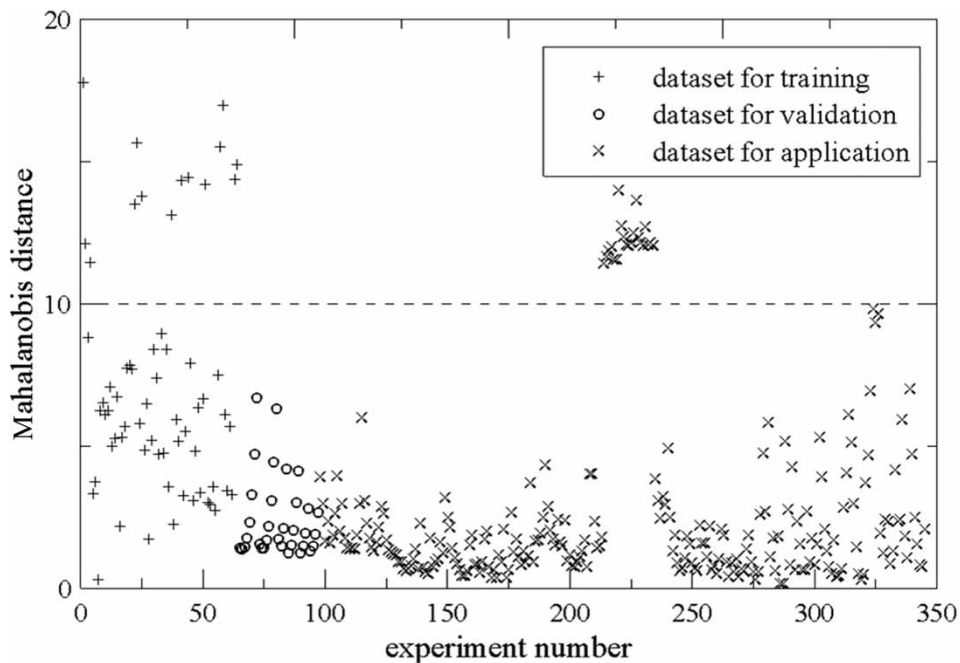
In Figure 5(a), the results for prediction using the model without data quality assessment are given. It can be seen that all three methods for data quality assessment improve the prediction of the ANN model. Specifically, the  $z$ -score method slightly improves the prediction by the ANN model. Multivariate methods were found to work better than the univariate method. This suggests that variables associated with local scour are not independent but are correlated with each other. MAPE was reduced by 16% in the prediction by the ANN model if outliers were identified by MDM, but were 10% in the case of EDM. This indicates that data quality assessment using MDM is slightly more efficient than by the EDM.

## Predictions and results

After training and validation, the ANN model was applied to data from Chiew's (1984) experiments. The  $z$ -score method, EDM, and MDM identified four, 30, and 21 outliers, respectively, in Chiew's data. Tables 1 and 3 indicate that the ranges of variables of the application



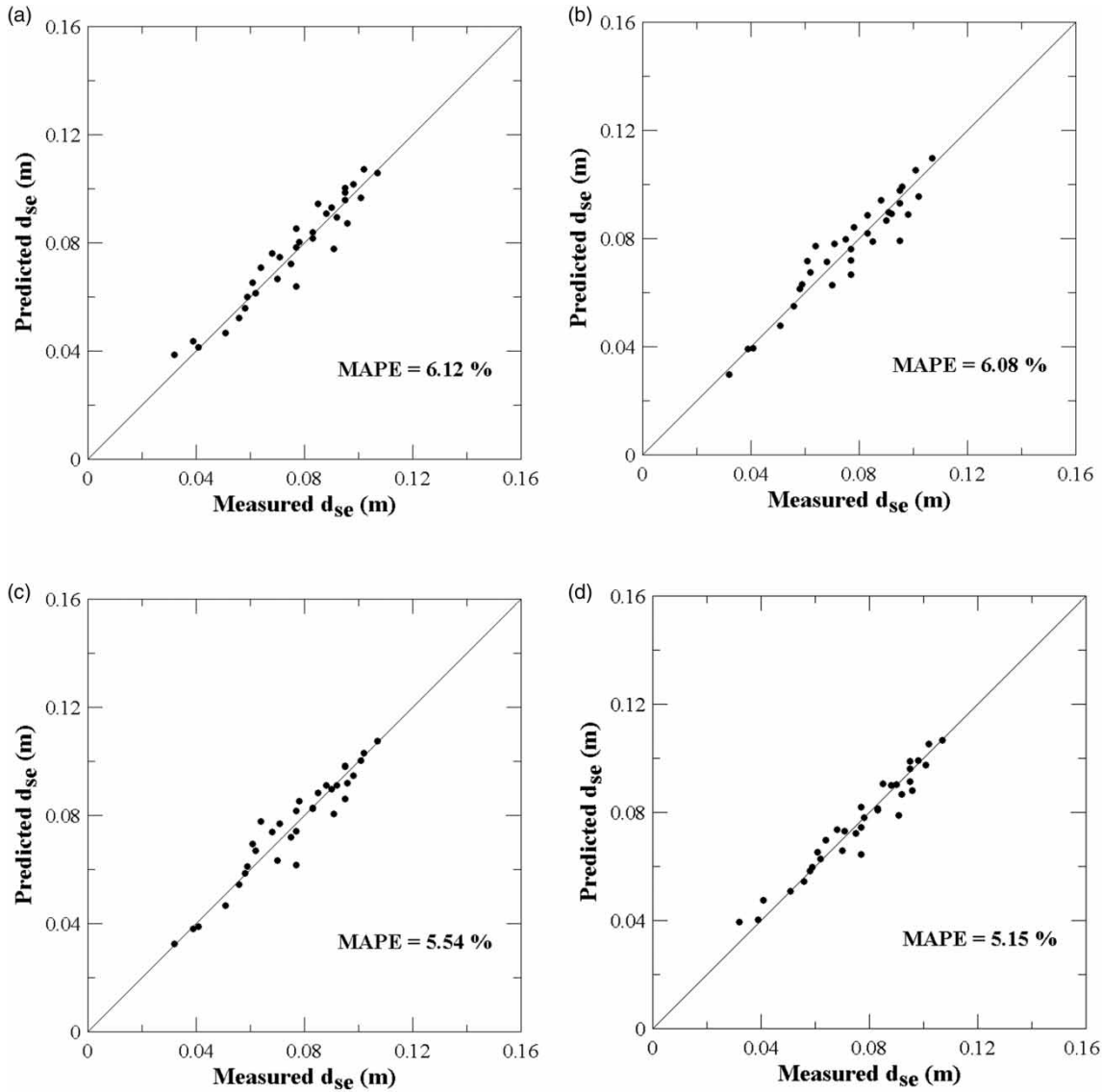
**Figure 3** | Plot of proximity parameter of the Euclidean distance method for laboratory data.



**Figure 4** | Plot of the Mahalanobis distance for laboratory data.

dataset are larger than those of the training dataset. Figure 6 shows application results when outliers are not identified (Figure 6(a)), are identified by EDM (Figure 6(b)),

and are identified by MDM (Figure 6(c)). Values of MAPEs are provided to show the performance of the ANN model. It can be seen that the assessment of data



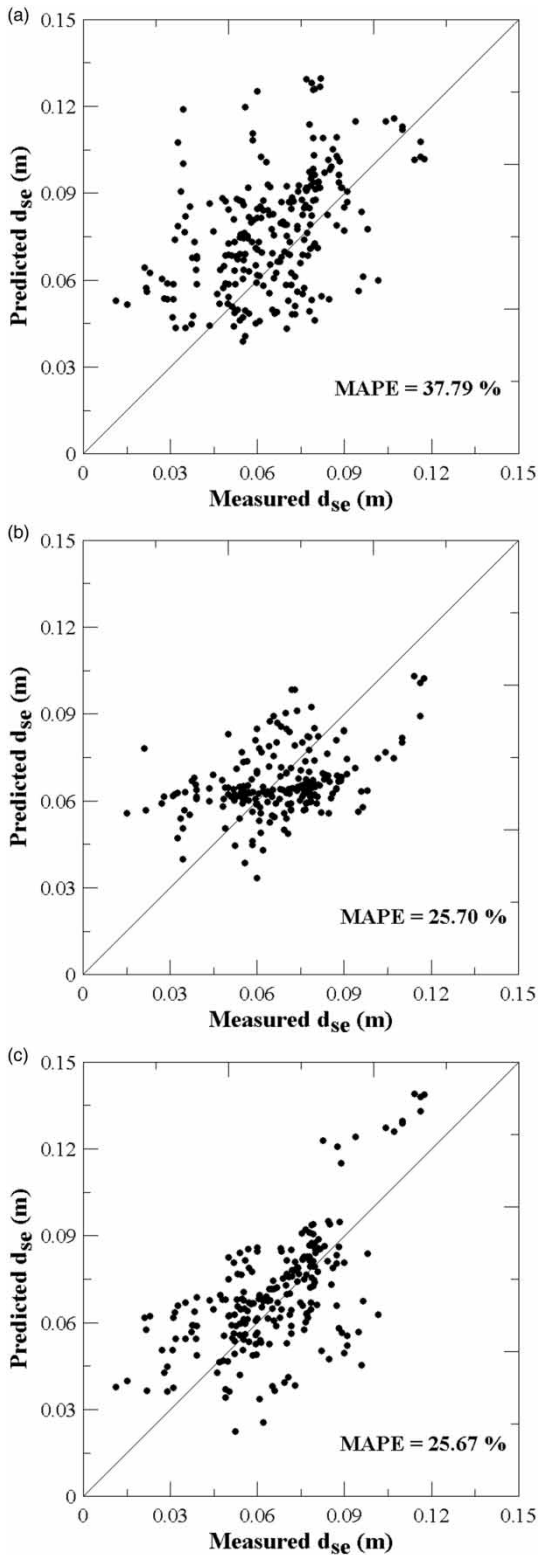
**Figure 5** | Validation of the ANN model: (a) without data quality assessment, (b) screened by z-score method, (c) screened by EDM, and (d) screened by MDM.

quality by the multivariate methods results in a significant improvement in predictions of the ANN model, largely by reducing the number of over-predictions. The value of MAPE by the two methods decreases by about 32%. Both EDM and MDM were found to improve the prediction at a similar level. It is interesting to note that scour depths in the range of 0.09–0.12 are under- and over-predicted by the ANN model if EDM and MDM are used to identify outliers, respectively.

## APPLICATION TO FIELD DATA

In this section, the ANN model is applied to [Mueller & Wagner's \(2005\)](#) field dataset. [Mueller & Wagner \(2005\)](#) collected 493 data of scour depths in 65 mid- to large-size rivers in the USA. They measured scour depths around various types of 106 bridge piers, including single piers and groups of piers. In the present study, a total number of 390 data for scour depths for a single pier were used.





**Figure 6** | Applications to Chiew's (1984) dataset: (a) without data quality assessment, (b) screened by EDM, and (c) screened by MDM.

### Assessment of data quality

Figures 7 and 8 show proximity parameters and Mahalanobis distances with experiment number for EDM and MDM, respectively. It can be seen in Figure 7 that a group of proximity parameters clearly range within a cutoff value. However, in Figure 8, most data for Mahalanobis distance are concentrated near the origin. As previously noted, the total number of eliminated data versus proximity parameter or Mahalanobis distance is plotted, and cutoff values of 10 for EDM and MDM were selected. These cutoff values were found to identify the total number of outliers, namely, 100 and 20 for EDM and MDM, respectively.

### Predictions and results

Figure 9 shows the prediction results by the ANN model when the laboratory data in Table 1 were used for training. The total number of data used for the application was 195, half the total number of Mueller and Wagner's data. Figure 9(a) shows the predicted versus measured scour depths without data quality assessment, and Figures 9(b) and 9(c) show the same figures when EDM and MDM were used for identifying outliers, respectively. The use of EDM and MDM resulted in the elimination of 57 and 32 data, respectively. Comparisons of MAPE reveal that both EDM and MDM results in slightly improved predictions, i.e., by about 15%. The level of improvement is less than that of the previous application in Figure 6 because the ranges of variables of the training dataset are somewhat narrower than those of the application dataset.

Figure 10 shows the same results but the field data were used to train the ANN model. In fact, Mueller and Wagner's data were divided into two parts, namely one part consisting of 195 data for the training and the other of 195 data for applications. Figure 10(a) shows the predicted versus measured scour depths without data quality assessment. It can be seen that training with the field data improves the predictions, when compared with Figure 9(a). Figures 10(b) and 10(c) show the prediction results if EDM and MDM are used, respectively, to identify outliers. After the data quality assessment, the numbers of data used in training and prediction were 152 and 138, respectively, for EDM, and 167 and 163, respectively, for MDM. It can be seen

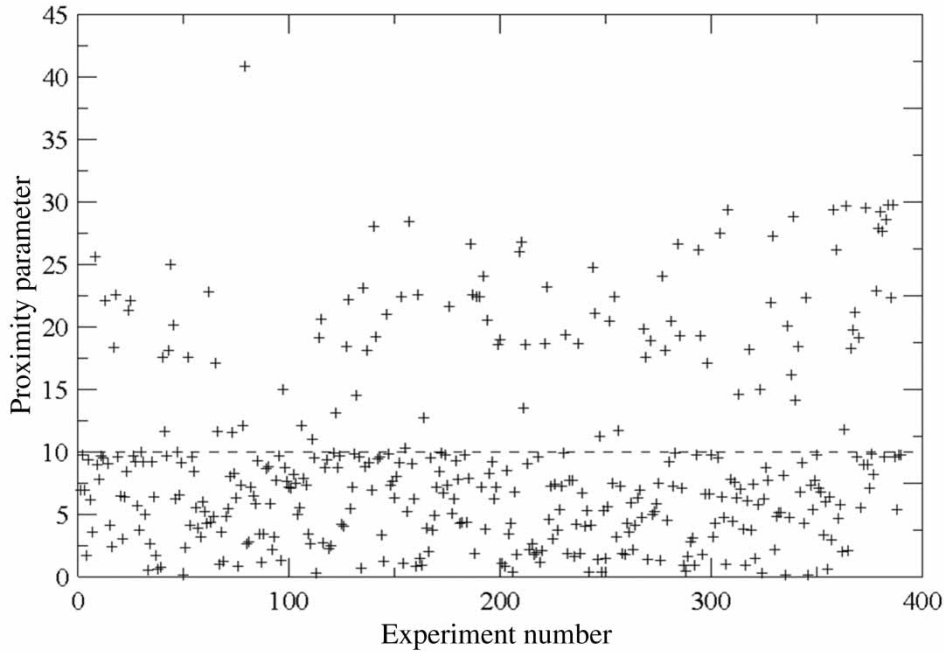


Figure 7 | Plot of proximity parameters of the Euclidean distance method for field data.

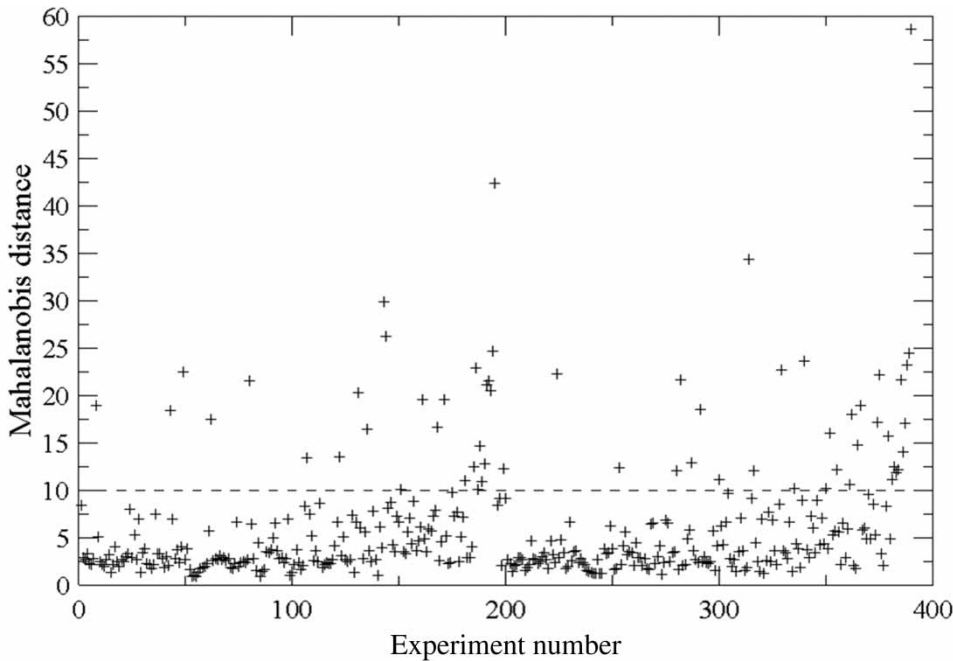
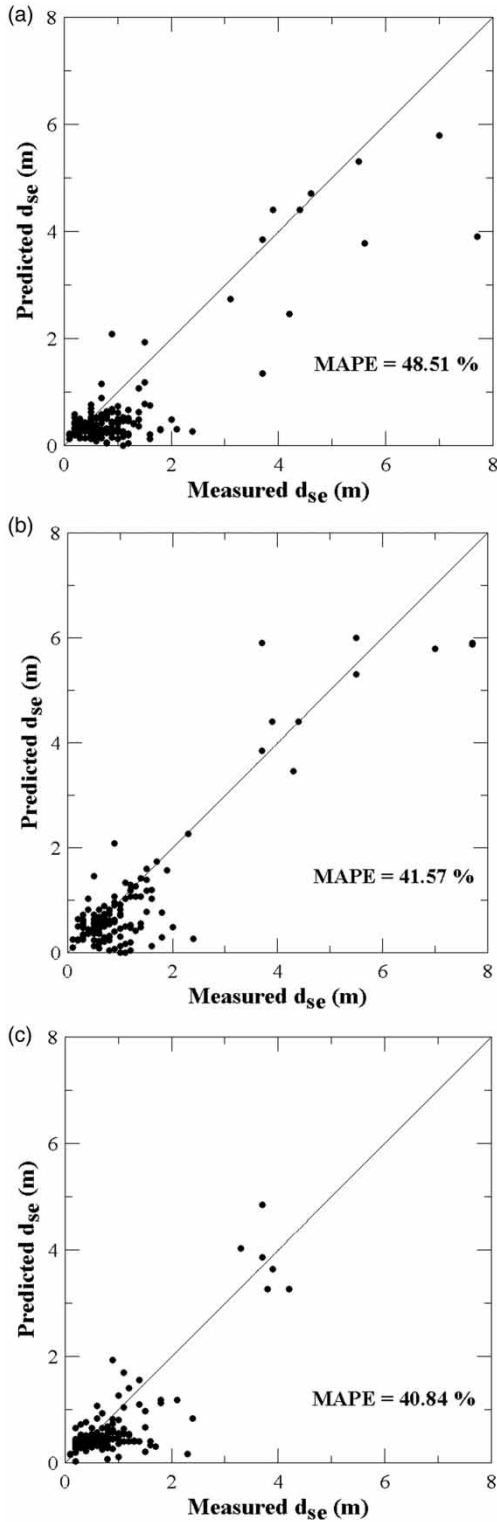


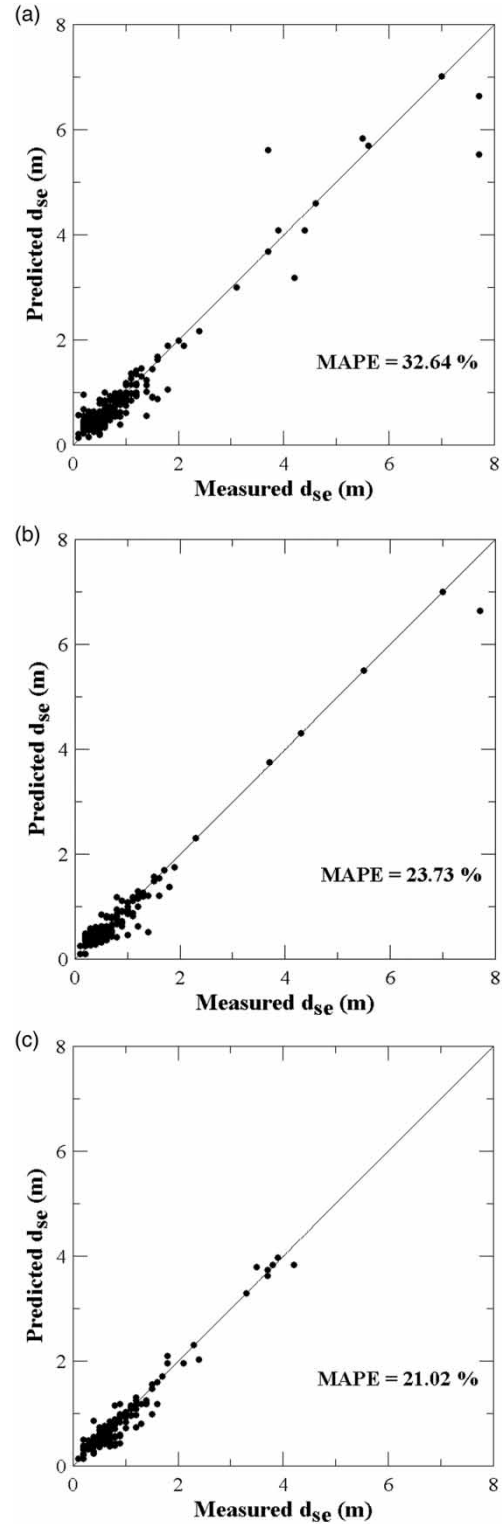
Figure 8 | Plot of the Mahalanobis distance for field data.

that the use of both EDM and MDM results in moderately improved predictions. The level of improvement is larger than the previous case where the laboratory data were

used to train the ANN model. This is because the ranges of the variables of the training dataset in this case are similar to those of the application dataset. In addition, it is



**Figure 9** | Applications to field data when the laboratory data are used for training: (a) without data quality assessment, (b) screened by EDM, and (c) screened by MDM.



**Figure 10** | Applications to field data when the filed data are used for training: (a) without data quality assessment, (b) screened by EDM, and (c) screened by MDM.

noteworthy that EDM identified more outliers as seen in Figures 7 and 8 and MDM identified larger values of scour depths as outliers. However, the levels of improvement by the two methods are similar.

## CONCLUSIONS

This study presents an evaluation of the use of data quality assessment to an ANN model that is capable of predicting scour depths around bridge piers. Five-dimensional variables such as approach velocity, flow depth, particle size, critical velocity of the particles, and pier width were used for the ANN modeling. Both univariate and multivariate methods were applied and the predicted results were compared. For the multivariate methods, the Euclidean distance method and Mahalanobis distance method were used.

First, laboratory data were used to train and validate the ANN model. Comparisons of the prediction results revealed that the applications of the multivariate methods improved the predictions made using the ANN model better than those of the univariate method. The ANN model was then applied to laboratory experiments by Chiew (1984). The findings indicate that both the Euclidean distance method and the Mahalanobis distance method significantly improve the prediction of the ANN model. The resulting values of MAPE were reduced by about 32%.

Lastly, the model was applied to field data, and the results were also compared with the case in which the model was trained with a part of field data. The ANN model appeared to predict better, provided it was trained using the field data. In addition, both the Euclidean distance method and the Mahalanobis distance method were found to improve the predictions made using the ANN model at a level similar to the application to the laboratory data, by reducing the value of MAPE by about 31%.

In summary, it was revealed that data quality assessment such as the Euclidean distance method and the Mahalanobis distance method significantly improves the prediction of the ANN model. That is, the use of the data quality assessment reduced values of MAPE in both laboratory and field data at a similar level. However, the uncertainty due to the subjective choice of the cutoff value should be minimized in the data quality assessment.

## ACKNOWLEDGEMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea Government (NRF-2014R1A2A1A11054236).

## REFERENCES

- Alameddine, I., Kenney, M. A., Gosnell, R. J. & Reckhow, K. H. 2010 Robust multivariate outlier detection methods for environmental data. *J. Environ. Eng.* **136** (11), 1299–1304.
- Azamathulla, H. Md., Ghani, A. A. & Guven, A. 2010 Genetic programming to predict bridge pier scour. *J. Hydraul. Eng. ASCE* **136** (3), 165–169.
- Baranya, S., Olsen, N. R. B., Stoesser, T. & Sturm, T. 2012 Three-dimensional RANS modeling of flow around circular piers using nested grids. *Eng. Appl. Comput. Fluid Mech.* **6** (4), 648–662.
- Bateni, S. M. & Jeng, D. S. 2007 Estimation of pile group scour using adaptive neuro-fuzzy approach. *Ocean Eng.* **34** (8), 1344–1354.
- Bateni, S. M., Jeng, D. S. & Melville, B. W. 2007a Bayesian neural networks for prediction of equilibrium and time-dependent scour depth around bridge piers. *Adv. Eng. Softw.* **38** (2), 102–111.
- Bateni, S. M., Borghei, S. M. & Jeng, D. S. 2007b Neural network and neuro-fuzzy assessments for scour depth around bridge piers. *Eng. Appl. Artif. Intell.* **20** (3), 401–414.
- Brandimarte, L., Montanari, A., Briaud, J. L. & D'Odorico, P. 2006 Stochastic flow analysis for predicting river scour of cohesive soils. *J. Hydraul. Eng.* **132** (5), 493–500.
- Breusers, H. N. C., Nicollet, G. & Shen, H. W. 1977 Local scour around cylindrical piers. *J. Hydraul. Res.* **15** (3), 211–252.
- Briaud, J. L., Chen, H. C., Kwak, K. W., Han, S. W. & Ting, F. C. K. 2001 Multiflood and multilayer method for scour rate prediction at bridge piers. *J. Geotech. Geoenviron. Eng.* **127** (2), 114–125.
- Chabert, J. & Engeldinger, P. 1956 *Etude des affouillements autour des piles des ponts [Study on scour around bridge piers]*. Laboratoire National d'Hydraulique, Chatou, France.
- Chiew, Y. M. 1984 Local scour at bridge piers. PhD Thesis, School of Engineering, The University of Auckland, New Zealand.
- Choi, S.-U. & Cheong, S. 2006 Prediction of local scour around bridge piers using artificial neural networks. *J. Am. Water Resour. Ass.* **42** (2), 487–494.
- Dargahi, B. 1990 Controlling mechanism of local scouring. *J. Hydraul. Eng.* **116** (10), 1197–1214.
- Dey, S., Bose, S. K. & Sastry, G. L. N. 1995 Clear water scour at circular piers: a model. *J. Hydraul. Eng. ASCE* **121** (12), 869–876.
- Firat, M. & Gungor, M. 2009 Generalized regression neural networks and feed forward neural networks for prediction of

- scour depth around bridge piers. *Adv. Eng. Softw.* **40** (8), 731–737.
- Gaudio, R., Grimaldi, C., Tafarojnoruz, A. & Calomino, F. 2010 Comparison of formulae for the prediction of scour depth at piers. In: *Proc., 1st IAHR European Division Congress*. Heriot-Watt University, Edinburgh, UK.
- Ge, L. & Sotiropoulos, F. 2005 3D Unsteady RANS modeling of complex hydraulic engineering flows. I: numerical model. *J. Hydraul. Eng. ASCE* **131** (9), 800–808.
- Ge, L., Lee, S. O. & Sturm, T. 2005 3D Unsteady RANS modeling of complex hydraulic engineering flows. II: model validation and flow physics. *J. Hydraul. Eng. ASCE* **131** (9), 809–820.
- Jain, S. C. & Fischer, E. E. 1979 *Scour Around Circular Bridge Piers at High Froude Numbers*. FHWA-RD-79-104, Federal Highway Administration, US Department of Transportation, Washington, DC, USA.
- Johnson, P. A. 1995 Comparison of pier-scour equations using field data. *J. Hydraul. Eng.* **121** (8), 626–629.
- Johnson, R. A. & Wichern, D. W. 1998 *Applied Multivariate Statistical Analysis*. Prentice-Hall, Englewood Cliffs, NJ, USA.
- Keshavarzi, A., Gazni, R. & Homayoon, S. R. 2012 Prediction of scouring around an arch-shaped bed sill using neuro-fuzzy model. *Appl. Soft Comput.* **12** (1), 486–493.
- Khan, M., Azamathulla, H. M. & Tufail, M. 2012 Gene-expression programming to predict pier scour depth using laboratory data. *J. Hydroinform.* **14** (3), 628–645.
- Kirkil, G., Constantinescu, G. & Ettema, R. 2009 Detached eddy simulation investigation of turbulence at a circular pier with scour hole. *J. Hydraul. Eng. ASCE* **135** (11), 888–901.
- Melville, B. W. & Chiew, Y. M. 1999 Time scale for local scour at bridge piers. *J. Hydraul. Eng.* **125** (1), 59–65.
- Melville, B. W. & Coleman, S. E. 2000 *Bridge Scour*. Water Resources Publications, LLC, Colorado, USA.
- Mueller, D. S. & Wagner, C. R. 2005 *Field Observations and Evaluations of Streambed Scour at Bridges*. Office of Engineering Research and Development, Federal Highway Administration, McLean, Virginia, USA.
- Muzzammil, M. 2010 ANFIS Approach to the scour depth prediction at a bridge abutment. *J. Hydroinform.* **12** (4), 474–485.
- Najafzadeh, M. & Azamathulla, H. M. 2013 Group method of data handling to predict scour depth around bridge piers. *Neural Comput. Appl.* **23** (7–8), 2107–2112.
- Raudkivi, A. J. & Ettema, R. 1983 Clear-water scour at cylindrical piers. *J. Hydraul. Eng.* **109** (3), 338–350.
- Robinson, R. B., Cox, C. D. & Odom, K. 2005 Identifying outliers in correlated water quality data. *J. Environ. Eng.* **131** (4), 651–657.
- Shen, H. W., Schneider, V. R. & Karaki, S. 1969 Local scour around bridge piers. *J. Hydraul. Div. ASCE* **95** (HY5), 1919–1940.
- Sheppard, D. M., Demir, H. & Melville, B. 2011 *Scour at Wide Piers and Long Skewed Piers*. National Cooperative Highway Research Program Rep. 682, Transportation Research Board, Washington, DC.
- Sheppard, D. M., Melville, B. & Demir, H. 2014 Evaluation of existing equations for local scour at bridge piers. *J. Hydraul. Eng. ASCE* **140** (1), 14–23.
- Yanmaz, A. M. & Altinbilek, H. D. 1991 Study of time dependent local scour around bridge piers. *J. Hydraul. Eng. ASCE* **117**, 1247–1268.

First received 12 September 2014; accepted in revised form 13 May 2015. Available online 20 June 2015