Closure to Discussion of "Improving Prediction of Dam Failure Peak Outflow Using Neuroevolution Combined with K-Means Clustering" by Amir Hossein Eghbali, Kourosh Behzadian, Farhad Hooshyaripor, Raziyeh Farmani, and Andrew P. Duncan,

July 2017, Vol. 22, No. 6.

DOI: 10.1061/(ASCE)HE.1943-5584.0001505

By Farhad Hooshyaripor[1] and Kourosh Behzadian[2]

The authors would like to first thank the discusser for making three constructive comments which can be divided into two groups: (1) comments 1 and 2 related to clarity of a function used in the model developed and accuracy of the data collected; (2) comment 3 related to the ability of the developed model to predict new peak discharges of dam failure. To close the discussion and further clarification, the following are noted for each of the comments:

Comment #1:

The authors of the original paper (Eghbali et al. 2017) applied MATLAB tool as a platform to generate and combine artificial neural network (ANN) with genetic algorithm (GA) and *k*-means clustering. However, due to the combination of ANN and GA, the standard ANN as noted by the discusser was not used and instead all steps were coded in MATLAB. Thus, the tangent sigmoid (*tansig*) transfer function used in the paper for generating ANN was not the default MATLAB

[1] Assistant Professor, Dept. of Civil Engineering, Science and Research Branch, Islamic Azad Univ., 14778-93855 Tehran, Iran
[2] Senior Lecturer, School of Computing and Engineering, Univ. of West London, London W5 5RF, U.K. (corresponding author). ORCID: http://orcid.org/0000-0002-1459-8408, E-mail: kourosh.behzadian@uwl.ac.uk

function denoted as *tansig*(n) in MATLAB. Instead, The following *tansig* transfer function was used in the paper (Araghinejad 2014):

$$tansig(x) = 2/(1+e^{-ax})-1=(1-e^{-ax})/(1+e^{-ax}) \qquad a>0 \qquad (1)$$

where a = constant parameter which was considered 1.

The authors did not carry out sensitivity analysis for parameter a but even if the default MATLAB function is used instead (i.e. a=2), it is unlikely to lead to less accurate predictions although different weights and biases may be obtained for the same database.

Comment #2:

The discusser highlighted the importance of the discrepancy of peak flow rates from the failure of three dams (i.e. Oros Dam in Brazil, Banqio Dam in China and Hell Hole Dam in the United States) reported by the original paper and other publications. As the accurate date is of paramount importance to the results, the authors endeavoured to collect the data from different sources and impartially pick up those that were widely used. More specifically, failure peak flow rate of Oros Dam used in the paper (i.e. 9,630 m^3/s) has been reported by several sources (e.g. Wahl 1998; Xu and Zang 2009; Pierce et al. 2010; and Thornton et al. 2011) while the value expressed by the discusser (i.e. 58,000 m^3/s which is around 6 times larger) can be found in only few works (e.g. Wahl 2014). In addition, the peak flow rate of 56,300 m^3/s for the Banqiao Dam failure was only used by the discusser's publication while peak flow rate of 78,000 m^3/s was reported by many independent researchers (e.g. Fujia and Yumei 1994; Xu and Zhang 2009; Pierce et al. 2010;Thornton et al. 2011). Due to large amount of peak discharge in these two data samples (i.e. Banqiao and Oros Dams), we also agree that the major difference in the collected data can directly affect any developed model. For example, Banqiao Dam failure is an important data sample due to having the highest peak discharge in the database. Also, the original paper used the widely-reported value of 7,360 m^3/s for Hell Hole dam failure (MacDonald and Langridge-Monopolis 1984; Wahl 1998; Xu and Zhang 2009) while peak discharge of 17,000 m^3/s has only been used by the discusser. In addition, the frequency of the 92 data samples analysed in the paper shows that most of the observed peak flow discharges are less than 10,000 m^3/s (Hoosyaripor et al. 2014). In this dataset, there are only one peak discharge over 70,000 m^3/s, 2 cases over 60,000 m^3/s, 3 cases over 30,000 m^3/s and 7 cases over 20,000 m^3/s among 92 data samples.

It should also be noted that there is only one predictive model which is trained for all clusters using 90% of all data samples, not based on the data samples in one cluster (e.g. Oros and Banqiao Dams) as noted by the discusser. In addition, due to insufficient data available (92 samples), conventional model verification (i.e. dividing database into two subsets of training and test) was inefficient. Hence, the cross-validation technique was used in the paper, implying that all 92 data samples were participated in the evaluation of the test set (see "Assessment of Performance Indicators" section in the original paper).

Comment #3:

The comment challenges overfitting of the clustered ANN-GA model. As noted by the discusser, overfitting often occurs when the number of hidden neurons is large (i.e. model is excessively complex) while the proposed model has only four neurons which is far less than the number of data pairs (i.e. 82-83 equal to 90% of data samples). In other words, if the number of parameters in the ANN is much smaller than the total number of points in the training set which is the case in the paper, there is little chance of overfitting (MATLAB). Also, to avoid overfitting, conventional ANNs divide the database into two subsets of training and validation, in which the training dataset will only participate in the model training. Then, the ANN training will carry on to improve the fitness on training dataset until the mode performance on validation dataset (i.e. independent and unseen date) is deteriorating. Similarly, the cross-validation technique was used in the proposed model as unseen data to avoid overfitting during the model training (Eghbali et al. 2017).

Furthermore, the authors totally agree with the trends of the profile traces shown in the discussion for the model developed. However, the reason for the unexpected functional responses in some profile traces cannot be attributed to a flaw or overfitting in the developed model but it is related to the discrepancy of the collected data. More specifically, cluster #1 has only two members (i.e. Oros and Banqiao Dams) in which the profile trace for constant value of H_w is the opposite of the expected function response (Fig. 1a in the discussion paper). When looking at the data of these two dams, it is apparent that given relatively similar H_w around 33m for both dams and a large water volume above the breached invert (V_w) in Oros Dam (660 mcm) compared to Banqiao Dam (607.5 mcm), the peak flow discharges are considerably opposite (78,100 m^3/s for Banqiao Dam compared to 9,630 m^3/s for Oros Dam). The other unexpected functional response is related to variation of Q_p with Vw in cluster #3 (Fig. 1c in the discussion paper) which has 3 members. Similar discrepancy can be observed within the dataset of these members. More specifically, coefficient of determination (R^2) between V_w and Q_p is significantly low (i.e. R^2=0.21) (Eghbali et al. 2017). Interestingly, the correlation between H_w and Q_p in the members of the same cluster is very strong (i.e. R^2=0.98) which also confirms the profile trace shown in Fig. 2c in the discussion paper. Similar correlation is in place for the last expected functional response (i.e. Fig. 1d of the discussion paper related to variation of Q_p with Vw in cluster #4. Although this cluster has relatively large number of data samples (i.e. 18 members), a weak correlation is observed between the members (i.e. R^2=0.02).

As can be seen in the above discussion, most of the inaccuracies and unexpected functional responses are mainly referred to the discrepancies between the data collected. Although the highly controversial data (e.g. the data in clusters #1 and #3 and some of the uncorrelated/outlier data in cluster #4) can be simply removed and the problem can be apparently solved, the authors do not recommend it due to the limited number of data available. Instead, the key message of the original paper is to identify the similar attributes of the data and conduct data clustering to recognize different specifications and predict their peak failure flows more accurately than the previously developed models including conventional regression models.

In addition, it seems to be inappropriate to stand by for future dam failures to enrich the quality and quantity of the database of dam failure as it will be unlikely to observe these catastrophic phenomena frequently in the future due to the advance in monitoring systems. Therefore, although some data collected from various breach cases may seem to be statistically chaotic, every piece of data may reveal information of high relevance (Gupta and Singh 2012) and hence they should not be changed/removed in favour of achieving a better correlation of the developed model for dam failure analyses.

References

Araghinejad, S. (2014). Data-Driven Modeling: Using MATLAB: registered: in Water Resources and Environmental Engineering. Water Science and Technology Library, vol(67), page 146.

Eghbali, A. H., Behzadian, K., Hooshyaripor, F., Farmani, R., & Duncan, A. P. (2017). Improving prediction of dam failure peak outflow using neuroevolution combined with k-means clustering. Journal of Hydrologic Engineering, 22(6), 04017007.

Fujia, T., and Yumei, L. (1994). “Reconstruction of Banqiao and Shimantan dams.” Int. J. Hydropow. Dams, 49–53.

Gupta, S. K. and Singh V. P. (2012). Discussion of “Enhanced Predictions for Peak Outflow from Breached Embankment Dams” by Christopher I. Thornton, Michael W. Pierce, and Steven R. Abt, J. Hydrol. Eng., DOI: 10.1061/(ASCE)HE.1943-5584.0000288.

Hooshyaripor, F., Tahershamsi, A., and Golian, S. (2014). “Application of copula method and neural networks for predicting peak outflow from breached embankments.” J. Hydro-Environ. Res., 8(3), 292–303.

MATLAB [Computer software]. MathWorks, Natick, MA

MacDonald, T. C., and Langridge-Monopolis, J. (1984). “Breaching characteristics of dam failures.” J. Hydraul. Eng., 1105, 567–586.

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

Thornton, C. I., Pierce, M. W., and Abt, S. R. (2011). “Enhanced predictions for peak outflow from breached embankment dams.” J. Hydrol. Eng., 10.1061/(ASCE)HE.1943-5584.0000288., 81–88.

Wahl, T. L. (1998). "Prediction of embankment dam breach parameters—A literature review and needs assessment." U.S. Bureau of Reclamation Dam Safety Rep. No. DSO-98-004, Denver.

Wahl, T. L. (2014). "Evaluation of erodibility-based embankment 144 dam breach equations." 145 *Hydraulic Laboratory Report HL-2014-02*, U.S. Department of the Interior, Bureau of 146 Reclamation, Denver, Colorado

Xu, Y., and Zhang, L. M. (2009). "Breaching parameters for earth and rockfill dams." J. Geotech. Geoenviron. Eng., 10.1061/(ASCE)GT .1943-5606.0000162, 1957–1970.