APPLICIBILITY OF A NEW DATA MINING PREDICTION MODEL OF SCOURING AROUND PILES

Samaneh Ghazanfari Hashemi[1], Testsuya Hiraishi[2]

[1]Dept. of Civil and Earth Resources Engineering, Kyoto University, samanehghh@gmail.com
[2]Dept. Of Civil and Earth Resources Engineering, Kyoto University, hiraishi.tetsuya.2c@kyoto-u.ac.jp

Regarding the complexity of the scouring of piles due to waves, existing models do not always provide acceptable results in predictions of scour hole properties. This study addresses two alternative approaches to estimate the wave-induced scour depth around vertical piles. To assess the performance of two data mining approaches: Artificial Neural Networks (ANN) and Support Vector Machines, data sets collected in the filed or laboratory were used. Controlling parameters of scouring such as Keulegan-Carpenter, pile Reynolds number, etc. were used as inputs and the amounts of scouring around each set of pile data were predicted as outputs of the models. Results of data mining approaches were compared by those of empirical approaches to assess the applicability of data mining models. Results indicate that data mining models provide better prediction than other models.

Key Words: *scouring, piles, waves, soft computing models, Artificial Neural Networks, Support Vector Machines*

1. INTRODUCTION

Various arrangements of pile groups are widely being used as supports of marine structures. As piles are located on erodible beds of the sea, scouring is a threat to such structures and the scour depth around pile groups should be considered well in their designs. The arrangement of the piles in addition to their geometry, sediment and wave characteristics should be considered to estimate the scour depth around a group of vertical piles. Regarding the importance of prediction of scour depth, several empirical formulae have been presented to predict equilibrium scour depth around piles, but such approaches which are mostly based on dimensional analysis and data correlation of laboratory tests do not always produce reasonable results for field conditions or even for laboratory cases.

Prediction of wave-induced scour depth around a single pile has been studied by (Hebrich et al. 1984; Sumer et al. 1992, 1993; Kobayashi & Oda 1994, Ayoublu et al., 2010, etc).

Scour around groups of vertical piles due to waves has been investigated by (Sumer & Fredsoe, 1998; Bayram & Larson, 2000)

Experimental investigations made by Sumer & Fredsoe (1998) showed the importance of the arrangement in a pile group scour. Their study also showed that the equilibrium scour depth is governed by G/D and KC number, where G is the gap between the piles, D is the pile diameter and KC is Keulegan–Carpenter number

Considering the complexity of modeling and the importance of scouring phenomena this study aims to assess new alternative approaches to predict scouring around single piles as well as pile groups.

2. CONTROLLING PARAMETERS AND EMPIRICAL FORMULAS OF SCOURING DUE TO WAVES

Wave-induced scour around a single pile depends on several groups of variables such as the characteristics of the wave, the sediment properties and geometry of the pile. Thus, the following functional relationship can be used to describe the equilibrium scour depth for a single pile (Sumer et al., 1992b):

$$S = f\,(T, d_{50}, U_m, U_{fm}, D, s, \upsilon) \qquad (1)$$

where S is the equilibrium scour depth, T is the wave period, d_{50} is the medium sediment diameter,

U_m is the maximum undisturbed orbital velocity at the sea bottom just above the wave boundary, U_{fm} is the shear velocity at the undisturbed bed given by $U_{fm} = (0.5f)^{0.5} U_m$ in which f is the wave friction factor, D is the pile diameter, s is the specific gravity of sediments and υ is the kinematics viscosity. Using the dimensional analysis, the above relationship can be presented in non-dimensional form as follows (Sumer et al., 1992b):

$$\frac{S}{D} = f\,(Re, N_s, \theta, KC) \tag{2}$$

Where Re is the pile Reynolds number, Ns is the sediment number, θ is the Shield's parameter and KC is the Keulegan–Carpenter number defined as follows:

$$Re = \frac{U_m D}{\nu} \tag{3}$$

$$N_s = \frac{U_m}{\sqrt{g(s-1)d_{50}}} \tag{4}$$

$$\theta = \frac{U_{fm}^2 D}{(s-1)g d_{50}} \tag{5}$$

$$KC = \frac{U_m T}{D} \tag{6}$$

The non-dimensional parameters should include the effect of various physical processes occurring during the scour, i.e. flow–seabed interaction, flow–structure interaction and sediment transport. In Equation (2) the Reynolds number and Keulegan–Carpenter number describe the flow pattern around piles, whereas the Shield's parameter and sediment number represent the mutual effects of flow on the seabed. For a group of vertical piles, in addition to the above parameters, spacing between the piles (G), number of piles normal to the flow (n) and number of piles parallel to the flow (m) are also important in the estimation of scour depth around pile groups. Considering these new parameters, the maximum scour depth S normalized by pile diameter D can be best expressed as follows (Ghazanfari-Hashemi et al. (2011)):

$$\frac{S}{D} = f\,(R_e, N_s, \theta, KC, \frac{G}{D}, \frac{m}{n}) \tag{7}$$

The existing empirical approaches for estimating scour depth at pile groups are given in **Table 1** for single piles and group piles.

Table 1 Empirical formulae applied in this study

Formula
$\frac{S}{D_e} = 1.3\{1-\exp[-0.043(KC-4.2)]\}$ *for* $KC \geq 4.2$ (a)* $\frac{S}{D_e} = 1.3\{1-\exp[-0.054(KC-3.3)]\}$ *for* $KC \geq 3.3$ (b)   * Myrhaug & Rue (2005) for single piles
$\frac{S}{D_e} = 0.023KC$ **   ** Bayram & Larson (2000) for pile groups

(1) DATA SET

In case of single piles, Laboratoray data of Subhasish et al. (2006) and Sumer et al. (1992), and in case of pile groups laboratory data of Sumer and Fredsoe (1998) and field data of Bayram and Larson (2000) were used respectively.

3. RESEARCH APPROACH

Two data mining approaches of Artificial Neural Networks (ANN) and Support Vector Machines as well as empirical formulae were applied to predict scour hole properties around piles due to waves. A brief introduction of each data mining method is given below:

(1) ARTIFICAIL NEURAL NETWORK (ANN)

ANN is the most famous data-mining approach that imitates some functions of the human brain. Neural networks are the general-purpose computing tools that can solve complex nonlinear problems The network comprises of a large number of simple processing elements linked to each other by weighted connections according to a specified architecture. These networks learn from the training data by adjusting the connection weights .he connection of neurons to each other can be carried out in various configurations. Hence, the simplest way of modeling a neural network consists of three layers: input layer, hidden layer and output layer. The optimum topology of ANN is usually determined by a trial-and-error procedure.

(2) SUPPORT VECTOR MACHINES (SVM)

Recently, a new soft computing approach named Support Vector Machines (SVM) has been successfully applied in problems such as the runoff modeling (Bray & Han, 2004), prediction of storm surge (Rjasekaran et al., 2008), and prediction of significant wave height (Mahjoobi & Mosabbeb, 2009).

In support vector regression the objective is to find a function f(x) which has at most ε deviation from the actually obtained targets y_i for all the training data $\{(x_1,y_1),...,(x_i,y_i)\}$ and at the same time is as flat as possible. In other words, errors are negligible as long as they are less than ε and any deviation larger than this is not accepted. *f(x)* can be expressed as (Smola & Scholkopf, 2003):

$$f(x) = (w,x) + b \quad , \quad w \in X \quad , b \in R \tag{8}$$

where w is a weight vector $(w \in R^n)$; b is additive noise $(b \in R)$ and (w,x) denote dot points in X. Flatness of the regression function $f(x)$ can be achieved by smaller values of w. To achieve that, the minimization problem can be written as a convex optimization problem:

$$\begin{cases} \min imize & \frac{1}{2}\|w\|^2 \\ subject\ to & \begin{cases} y_i - (w,x_i) - b \le \varepsilon \\ (w,x_i) + b - y_i \le \varepsilon \end{cases} \end{cases} \tag{9}$$

The optimization problem can be feasible if the error on any training data is less than ε

$$\min imize \ \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}(\xi_i + \xi_i^*)$$

$$subject\ \text{to} \begin{cases} y_i - (w_i,x_i) - b \le \varepsilon + \xi_i \\ (w_i,x_i) + b - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{cases} \tag{10}$$

where ζ_i and ζ^*_i are slack variables and the constant $C>0$ determines the trade-off between the flatness of $f(x)$ and the tolerable amount larger than ε which is defined by the user.

Nonlinear support vector regressions can be used in complex and nonlinear problems. Solving nonlinear problems can be achieved by mapping the data into a higher-dimensional feature space with the help of kernel functions.

The problem of support vector regression in the feature space can be written as (11) in the feature space as;

$$f(x) = \sum_{i=1}^{m}(\alpha_i - \alpha_i^*)K(x_i,x) + b \tag{11}$$

where α and αi* are Lagrangian Parameters.

In addition to the choice of a kernel, SVM requires the setting up of kernel-specific parameters. Furthermore, optimum values of the regularization parameter C and the size of error in the sensitive zone need to be determined. The choice of these parameters controls the complexity of the prediction. SVM needs less computational time and fewer parameters than ANN which leads it to be trained more rapidly (Mahjoobi & Mosabbeb, 2009; Ghazanfari-Hashemi et al., 2011)

4. SCOUR DEPTH PREDICTIONS

The data set was separated into two input categories as inputs for prediction of the scour depth: single piles (*Re, Ns, ɵ, KC*) and group of piles (*Re, Ns, ɵ, KC, G/D, m/n*). Scour depth normalized by pile diameter (*S/D*) was predicted as the output.

Both sets of single and group pile data were divided into two subsets for training and testing phases. The total number of single pile data was 88 and those of group piles were 102 data points. In both cases 80% of the data were used for training phase and 20% were used in validation.

(1) ANN MODEL

A Multi-Layer Perceptron (MLP) with Back Propagation (BP) learning algorithm was used to train the network structure of the single pile scouring prediction. The number of neurons for input, hidden and output layers was 4, 2 and 1 respectively.

Likewise, a Multi-Layer Perceptron (MLP) with Back Propagation (BP) learning algorithm was used to train the network structure of the group pile scouring prediction. The number of neurons for input, hidden and output layers was 6, 1 and 1 respectively.

The performances of models were assessed quantitatively using the following statistical parameters: coefficient of correlation (*CC*), root mean square error (*RMSE*) and scatter index (*SI*) which are defined as the following relationships

$$CC = \frac{\Sigma(T_i - \overline{T_i})(O_i - \overline{O_i})}{\sqrt{\Sigma(T_i - \overline{T_i})^2 \Sigma(O_i - \overline{O_i})^2}} \tag{12}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(O_i - T_i)^2}{N}} \tag{13}$$

$$SI = \frac{RMSE}{\overline{T_i}} \tag{14}$$

$$Bias = \overline{O_i} - \overline{T_i} \tag{15}$$

In the above formulae, O_i and T_i represent target and network outputs for the i th output, respectively; $\overline{O_i}$ and $\overline{T_i}$ are the average of target and network outputs and Ni is the total number of data points. A higher value of CC and smaller values of $RMSE$ and SI mean a better model performance.

Fig. 1 and Fig. 2 show the comparison between the measured and predicted values of scour depth in case of single and group piles.

(2) SVM MODEL

The SVM model requires setting of a few user-defined parameters, such as the regularization parameter (C) and the type of kernel (polynominal or RBF).

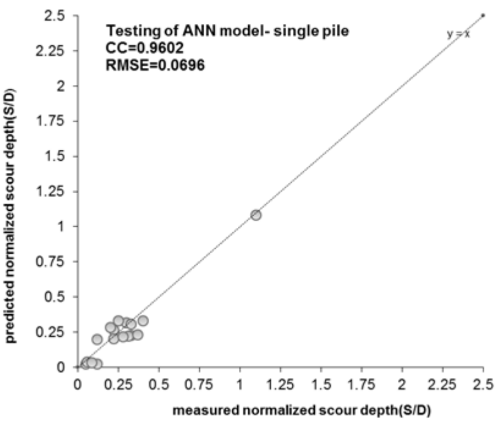

Fig. 1 Comparison between the observed and predicted normalized scour depth of single piles by the ANN model

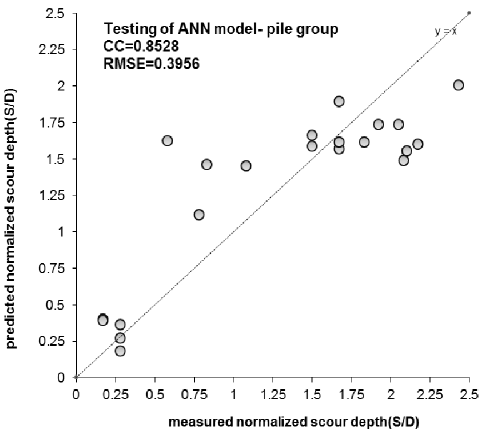

Fig. 2 Comparison between the observed and predicted normalized scour depth of group piles by the ANN model

In this study, the regularization parameter C and the size of error in sensitive zone parameters control the complexity of prediction. A value of C=350 and ε=0.010 were selected based on the process of error minimizing in case of single pile and the value of C=3.44 and ε=0.00001 in case of pile groups.

Fig. 3 and Fig. 4 show the comparison between the measured and predicted values of scour depth for single pile. In **Fig. 5** and **Fig. 6** the comparisons between the measured and predicted values of scour depth by different approaches are shown.

5. RESULTS AND DISCUSSIONS

Table 2 presents a summary of the results of all approaches (data mining and empirical approaches) and their error statistics. Comparison of results and their related statistical errors indicate that data mining approaches outperform the empirical ones not only in single piles but also in case of pile groups. The comparisons of all applied approaches were also shown in **Fig. 5** and **Fig. 6.**

Regarding **Table 2**, it can be concluded that

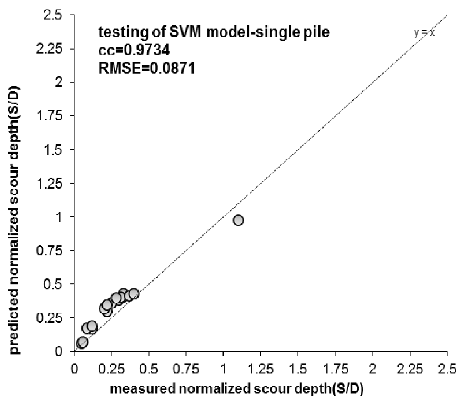

Fig. 3 Comparison between the observed and predicted normalized scour depth of single piles by the SVM model

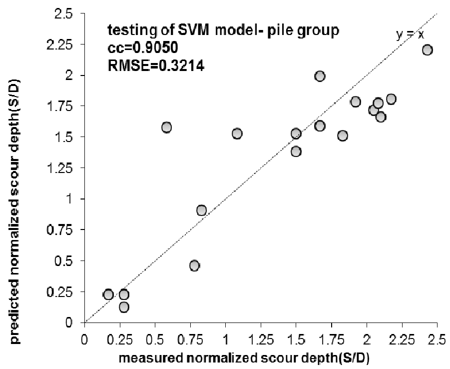

Fig. 4 Comparison between the observed and predicted normalized scour depth of group piles by the SVM model

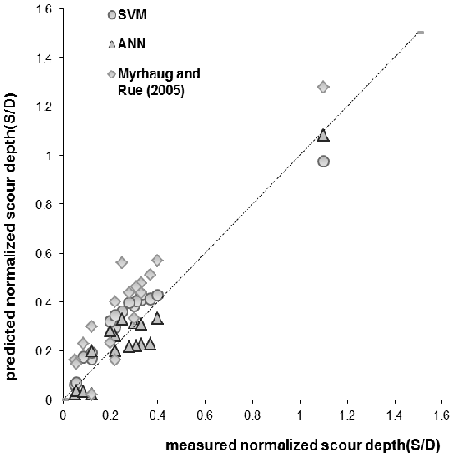

Fig. 5 Comparison of measured and predicted normalized scour depth for single pile using different approaches.

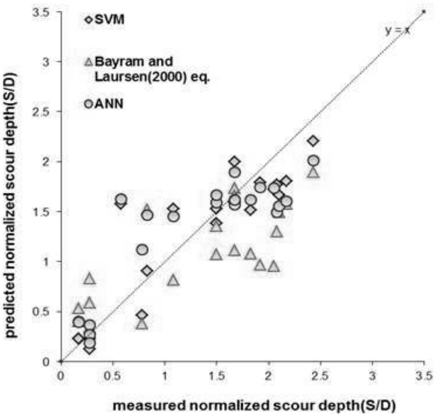

Fig. 6 Comparison of measured and predicted normalized scour depth for pile groups using different approaches

Table 2 Summary of results by all approaches

Approach		CC	RMSE	BIAS
Single pile	SVM	0.9734	0.0871	0.0643
	ANN	0.9602	0.0696	-0.0884
	Myrhaung &Rue (2005)	0.9425	0.1492	0.1405
Pile group	SVM	0.9050	0.3214	-0.0463
	ANN	0.8528	0.3956	0.0150
	Bayram &Larson (2000)	0.1176	1.2526	0.0154

SVM and ANN can be introduced as high reliable alternative approaches for empirical ones in prediction of pile scour depths due to waves.

REFERENCES

Ayoubloo, M.K, A. Etemad-Shahidi, J.Mahjoobi (2010): Evaluation of regular wave scour around a circular pile using data mining approaches, *Journal of Applied Ocean Research, VOl.*32, pp. 34-39.

Bayram, A. and M. Larson (2000): Analysis of scour around a group of vertical, *Journal of Wat. Port, Coastal Ocean Engng,* Vol. 126, pp.215-220.

Bray, M. and D. Han (2004): Identification of support vector machines for runoff modelling, J. *Hydroinf.* Vol6, No. 4, pp. 265–280.

Ghazanfari-Hashemi,S., A. Etemad-Shahidi, M.H. Kazeminezhad, A.R. Mansoori(2011): Prediction of pile group scour in waves using support vector machines and ANN, *Journal of Hydroinformatics,* Vol. 13 No. 4, pp. 609–620.

Hebrich, J.B., R.E.J. Schiller, R.K.Watanabe, W.A. Dunlap (1984): Sea Floor Design Guidelines for Ocean-founded Structures. Marcel Dekker, New York.

Kobayashi, T. and K. Oda (1994): Experimental study on developing process of local scour around a vertical cylinder. *Proc. 24th Int. Conf. on Coastal Engineering vol 2. ASCE, New York*, pp. 1284–1297.

Mahjoobi, J. and E. Mosabbeb (2009): Prediction of significant wave height using regressive support vector machines. *Journal of Ocean Engng*. Vol. 36, pp.339–347.

Myrhaug, D. and H. Rue (2005): Scour around group of slender vertical piles in random waves. *Journal of Applied Ocean Res*. Vol. 27, No.1, pp. 56–63.

Rjasekaran, S., S.Gayathri, T.L.Lee (2008): Support vector regression methodology for storm surge predictions. *Journal of Ocean Engng.* Vol. 35, pp.1578–1587.

Subhashish, D., M. Sumer, J. Fredsoe,(2006): Control of Scour at Vertical circular Piles under Waves and Current. *Journal of Hydraulic Engineering,* Vol.132, No. 3, pp. 270-279.

Smola, A.J. and B. Scholkopf (2004): A tutorial on support vector regression, *Journal of Statistics and computing* Vol 14, No.3, pp. 199-222.

Sumer, B.M. & J. Fredsoe (1998): Wave scour around a group of vertical piles, *J. Wat. Port Coastal Ocean Engng.* Vol. 124, pp. 248–256.

Sumer, B.M., J. Fredsoe, N. Christiansen (1992): Scour around vertical pile in waves. *J. Wat. Port Coastal Ocean Engng.* Vol. 117, pp. 15–31.

Sumer, B.M., J. Fredsoe, N. Christiansen (1993): Influence of cross section on wave scour around piles. *J. Wat. Port Coastal Ocean Engng.* Vol. 119, pp. 477–495.

Zounemat-Kermani, M., A.A.Beheshti, B. Ataie-Ashtiani, S.R. Sabbagh-Yazdi (2009): Estimation of current-induced scour depth around pile groups using neural network, adaptive neuro-fuzzy inference system. *Journal of Applied Soft Computing,* Vol 9, pp. 746–755.

(Received June 15, 2012)