Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 116 (2015) 149 – 154

Procedia Engineering

www.elsevier.com/locate/procedia

8th International Conference on Asian and Pacific Coasts (APAC 2015)

Estimation of Stability Number of Rock Armor Using Artificial Neural Network Combined with Principal Component Analysis

Anzy Lee[a], Sung Eun Kim[a], Kyung-Duck Suh[a]*

[a]*Department of Civil and Environmental Engineering, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 151-744, Korea*

Abstract

In this paper, a hybrid artificial neural network (ANN) model is constructed to estimate the stability number of rock armor using the experimental data of Van der Meer (1988). Among the eleven input parameters in the experiment, the six parameters each of which is well distributed in a certain range are transformed into six principal components (PCs) by using a principal component analysis (PCA), which are then used as the input variables of the ANN. The remaining five parameters that vary among several different values (e.g. number of waves of 1000 or 3000) are directly used as the input variables of the ANN. Since the orthogonality of the PCs prevents the duplication of information by separating the variables into several independent components while maintaining the critical information in them, the hybrid ANN model combined with the PCA gives better results compared with the conventional ANN models.

Peer- Review under responsibility of organizing committee , IIT Madras , and International Steering Committee of APAC 2015

Keywords: *Armor stone; Artificial Neural Network; Pricipal Component Analysis; Stability Number*

1. Introduction

The stability number is an important variable in the design of sloping revetments and breakwaters, especially in determining the optimum weight of armor stones. The formula suggested by Hudson (1959) has been widely used probably because of its simplicity. However, it does not include the influence of the factors that affect the stability of coastal structures, e.g. wave period and random waves. In order to overcome this problem, a new design formula was proposed by Van der Meer (1987, 1988) based on the experimental data of Van der Meer (1988). It additionally

* Corresponding author. Tel.: +82-2-880-8760; fax: +82-2-873-2684.
E-mail address: kdsuh@snu.ac.kr

Peer- Review under responsibility of organizing committee , IIT Madras , and International Steering Committee of APAC 2015
doi:10.1016/j.proeng.2015.08.276

includes the influence of wave period, wave spectrum shape, number of waves, groupiness of waves, and the permeability of the core.

The Van der Meer's (1988) experimental data have been used for the development of artificial neural network (ANN) models (Mase et al. 1995; Kim and Park 2005; Balas et al. 2010). Particularly, in the hybrid ANN model developed by Balas et al. (2010), the five input parameters used in the Van der Meer (1987) formula were transformed into five or four principal components (PCs) using a principal component analysis (PCA), which were then used as the input variables of the ANN model. To be more specific, the five variables, permeability of core (P), damage level (S), number of waves (N_w), structure slope ($\cot\alpha$), and surf similarity parameter (ξ_m), were transformed into five or four PCs. They showed that the estimating ability of ANN models was enhanced with the use of PCA when compared with ANNs trained by the untreated data set. They also showed that using more number of PCs gives better result.

In this research, a hybrid ANN model with PCA, similar to the Balas et al.'s (2010) model, is developed using the experimental data of Van der Meer (1988). Among the eleven input parameters in the experiment, the six parameters that are well distributed in certain ranges are transformed into six PCs by using a PCA, which are then used as the input variables of the ANN. The remaining five parameters that vary among several values (e.g. number of waves of 1000 or 3000) are directly used as the input variables of the ANN. The estimating capability of the present hybrid ANN model is compared with those of the previous ANN models with or without PCA.

2. Rock Armor Stability Number

The stability number is a dimensionless number which measures the stability of the armor layer of a rubble mound structure. It is defined as

$$N_s \equiv \frac{H_s}{\Delta D_{n50}} \tag{1}$$

where H_s is the significant wave height in front of the mound structure, $\Delta = \rho_a / \rho - 1$ is the relative mass density, ρ_a is the mass density of stone, ρ is the mass density of water, and D_{n50} is the nominal size of the armor unit. The stability number indicates how stable the armor stone is and is used to determine the required weight of stones for a given wave height at a particular place. In order to estimate the stability number, it is essential to figure out the relationship between the stability number and other parameters which describe the characteristics of waves and structure. However, the physical mechanism between waves and the stability of armor units is so complicated that it is not practical to find the analytic solution between them. For this reason, a lot of experiments which consider various physical characteristics of structures and waves were conducted to propose empirical formulas among them.

Hudson(1959) proposed an empirical formula as follows.

$$N_s = (K_D \cot\alpha)^{1/3} \tag{2}$$

where K_D is the stability coefficient that depends on the shape of the armor unit, method of placement, the location at the structure (i.e. trunk or head), and whether the breaking of incident wave occurs before reaching or on the structure face. Even though it is very simple, the Hudson formula has been found to have a lot of shortcomings. It does not include, for example, the influence of wave period and does not take into account random waves. Thus, an extensive series of tests including the parameters which are considered to have significant effects on armor stability were conducted by Van der Meer(1988), and the empirical formula based on the experimental data was proposed by Van der Meer(1987, 1988) as follows.

$$N_s = \begin{cases} \dfrac{1}{\sqrt{\xi_m}}\left[6.2P^{0.18}\left(\dfrac{S}{\sqrt{N_w}}\right)^{0.2}\right] & \text{for } \xi_m < \xi_c \\ 1.0P^{-0.13}\left(\dfrac{S}{\sqrt{N_w}}\right)^{0.2}\sqrt{\cot\alpha}\,\xi_m{}^{P} & \text{for } \xi_m \geq \xi_c \end{cases} \quad \text{where } \begin{cases} \xi_m = \tan\alpha / \sqrt{2\pi H_s / gT_m{}^2} \\ \xi_c = \left(6.2P^{0.31}\sqrt{\tan\alpha}\right)^{1/(P+0.5)} \end{cases} \tag{3}$$

where T_m is the average wave period, and ξ_c is the critical surf similarity parameter indicating the transition from plunging waves to surging waves.

Later, several ANN models to estimate the stability number were developed by Mase et al. (1995), Kim and Park (2005), and Balas et al. (2010), yielding more or less the same agreement with the experimental data of Van der Meer(1988) compared with the Van der Meer's (1987) empirical formula.

3. Hybrid Artificial Neural Network Model

An ANN model is a representative data-driven model aiming to mimic the systematic relationship between input variable and output variable using a training algorithm. 'Training' is to modify the matrix that relates the input and output variable sets so that the output values from the model are as close as possible to the target values. An ANN model consists of a set of processing elements called 'neurons' carrying the information, and they are mutually connected with different weights indicating the correlation between input and output data. The general configuration of an ANN model is illustrated in Fig. 1.

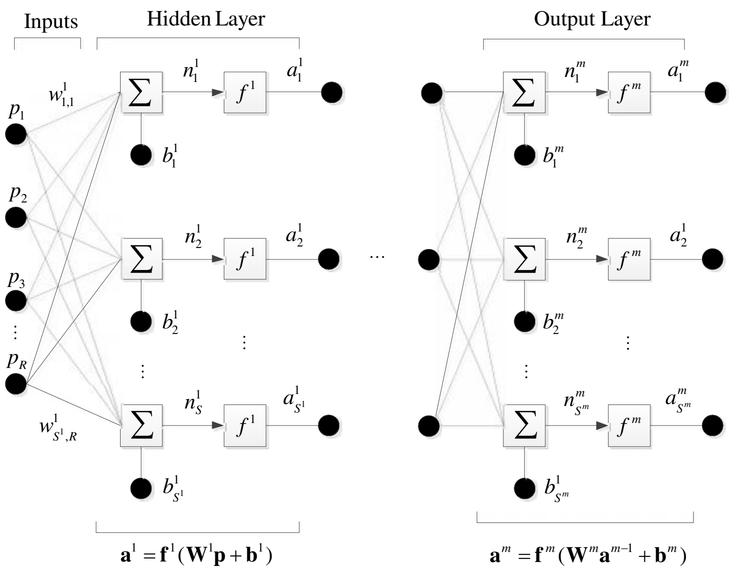

Fig. 1. General configuration of ANN model.

Owing to the property of data-driven models, the ANN model is highly sensitive to the correlation between input and output variables. Therefore, the selection of input dataset or data preprocessing is crucial to the design an ANN model. The basic concept of choosing variables in both empirical formulas and ANN models to estimate stability numbers is to select the variables which have a great influence on the stability number and eliminate the others. In this 'select or eliminate', so-called variable selection method, especially in the case of ANN models, the variable set is often selected by comparing the results of all possible combinations of variable set. Although it can eliminate the uninfluential variables, it can lose the variables that have relatively small effects on the target variable. Consequently, there is a possibility to lose the information included in the eliminated variables. The PCA is often used to evaluate the relative importance of variables while taking all the variables into account. For this reason, the unused variables in the previous studies are additionally considered and transformed to the set of orthogonal functions using a PCA in this research.

The parameters in the experiment of Van der Meer (1988) are given in Table 1. Noticing that a PCA is not adequate to analyze the data which vary among only a few values, the parameters are classified into 2 groups: Group 1 containing the parameters each of which is well distributed in a certain range (e.g. $0.0461 \le H_s \le 1.1800$ cm); and Group 2 containing the parameter each of which varies among only a few values (e.g. N_w = 1000 or 3000). A PCA was used to transform the parameters in Group 1 into six PCs. The eigenvalue and the percentage of the explained variance of each PC are given in Table 2 along with the cumulative explained variance. It can be seen that the last PC incorporated almost zero percent of the total variance. However, the ANN model with all the PCs gave somewhat better result than that calculated using the first five PCs without the last one. It seems that the last PC explains the outliers. Therefore, we included all the PCs in the ANN model. As a result, the six PCs obtained from the data in Group 1 and the five data values in Group 2 are used as the input data of the ANN model.

Table 1. Classification of the parameters in the experiment of Van der Meer (1988)

Group 1	Group 2
H_s : Significant wave height in front of the structure	N_w : Number of waves
T_m : Average period	h : Water depth
T_p : Peak period	P : Permeability of core
ξ_m : Surf similarity parameter	$\cot\alpha$: Slope angle
S : Damage level	SS : Spectral shape
h / H_s : Dimensionless depth	

Table 2. Eigenvalue, explained variance and cumulative explained variance of principal components

Principal Component	Eigen value	Explained variance (%)	Cumulative explained variance (%)
1	2.552	42.528	42.528
2	2.011	33.513	76.041
3	.848	14.140	90.181
4	.337	5.611	95.793
5	.231	3.843	99.636
6	.022	.364	100.000

4. Application and Result

In this study, the 579 experimental data of Van der Meer (1988) excluding the data of low-crested structures were used. The randomly selected 100 data were used to train the network, and the remaining 479 data were used to test the model, as done by Mase et al. (1995). For the training, the Levenberg-Marquardt algorithm was used and the training was stopped when the epoch reached to 5000. In addition, to take into account the sensitiveness of initial weights, 100 ANN models were generated in parallel and the model whose output variables gave the largest correlation coefficient against the target (or observed) variables was selected. Furthermore, in order to obtain the optimal number of neurons, ANN models which have different number of hidden neurons (i.e. 3, 6, 9, 12, 15, and 18) were generated and compared one another. The correlation coefficients between the ANN model output and the observed values for the test data are given in Table 3 for different numbers of hidden neurons, indicating that the ANN model gives the best result when the number of hidden neurons is three. Fig. 2 shows the regression plots for the training data and test data in the case of three hidden neurons. The training data show almost perfect agreement between the model and observation. The agreement for the test data is also very good.

To compare the performance of the present ANN model with the previous models, the correlation coefficients for the test data are summarized in Table 4. Even though there are some differences among the authors for the input data of the models and the experimental data used for the construction and test of the models, the present hybrid ANN model with PCA gives much larger correlation coefficient than other models.

Table 3. Correlation coefficient between the output data from ANN model and the observed data

Number of hidden neurons	Correlation coefficient (R)
3	**0.9742**
6	0.9501
9	0.9371
12	0.9254
15	0.9378
18	0.9241

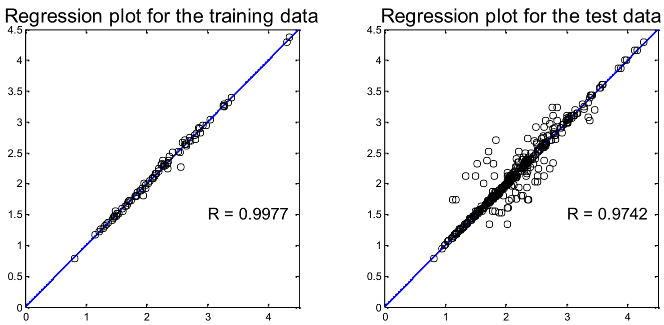

Fig. 2. The regression graph of output variable and target variable

Table 4. Correlation coefficients of different empirical formula or ANN models for test data

Author	Correlation coefficient	Remarks
Van der Meer (1987)	0.92	Empirical formula, Eq. (3) in this paper
Mase et al. (1995)	0.91	Also including the data of Smith et al. (1992)
Kim and Park (2005)	0.902~0.952	Including the data for low-crest structures

Balas et al. (2010)	0.906~0.936	Hybrid ANN model with PCA
Present study	0.974	Hybrid ANN model with PCA

5. Conclusion

A hybrid ANN model combined with PCA was constructed for prediction of stability number of rock armor based on the experimental data of Van der Meer (1988). The input parameters in the experiment were classified into two groups, one consisting of the parameters each of which is well distributed in a certain range and the other consisting of the parameters each of which varies among several different values. The six parameters in Group 1 were transformed into six PCs by a PCA, which were then used as the input variables of the ANN model. The five parameters in Group 2 were directly used as the input variables of the ANN model. The present hybrid ANN model compared against the observed stability numbers gave much larger correlation coefficient than the previous empirical formula or ANN models, indicating the better performance of the present model.

Acknowledgements

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (NRF-2014R1A2A2A01007921). The research was conducted at the Engineering Research Institute of Seoul National University.

References

Balas, C.E., Koc, M.L., Tür, R., 2010. Artificial neural networks based on principal component analysis, fuzzy systems and fuzzy neural networks for preliminary design of rubble mound breakwaters. Applied Ocean Research 32, 425-433.

Hudson, R.Y., 1959. Laboratory investigation of rubble-mound breakwaters. Journal of Waterways and Harbors Division 85(WW3), 93-121.

Kim, D.H., Park, W.S., 2005. Neural network for design and reliability analysis of rubble mound breakwaters. Ocean Engineering 32. 1332-1349.

Mase, H., Sakamoto, M., Sakai, T., 1995. Neural network for stability analysis of rubble-mound breakwaters. J. Waterway, Port, Coastal and Ocean Eng. 121(6), 294-299.

Smith, W.G., Kobayashi, N., Kaku, S., 1992. Profile changes of rock slopes by irregular waves. Proc. 23rd Int. Conf. on Coastal Engrg., ASCE, New York, N.Y., 1559-1572.

Van der Meer, J.W., 1987. Stability of breakwater armor layers − Design formulae. Coastal Engineering 11, 93-121.

Van der Meer, J.W., 1988. Rock slopes and gravel beaches under wave attack. Delft Hydraulics, Communication No. 396.