Ocean Engineering 127 (2016) 349–356

Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng

An easy way to use artificial neural network model for calculating stability number of rock armors

Anzy Lee, Sung Eun Kim, Kyung-Duck Suh*

Department of Civil and Environmental Engineering, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea

ARTICLE INFO

Keywords:
Artificial neural network
Explicit method
Principal component analysis
Rock armors
Stability number

ABSTRACT

Since Van der Meer proposed a new empirical formula to calculate the stability number of rock armors based on his own experimental data in 1987, the data have been used in the development of artificial neural network (ANN) models. However, the ANN models are seldom used probably because they did not significantly improve the accuracy compared with the empirical formula and/or the engineers are not familiar with the ANN models. To resolve these problems, first we develop an ANN model combined with a principal component analysis (PCA) that markedly improves the accuracy of the model. The by-products of the model such as weights and biases are then used to develop an explicit method to calculate the stability number. The developed method is not as simple as the empirical formula but it can be easily used by an engineer who has basic knowledge of matrix operation without requiring knowledge of ANN and PCA. It is equally accurate as the developed ANN-PCA hybrid model and more accurate than previous empirical formula and ANN models.

1. Introduction

To determine the optimum weight of rock armors for sloping revetments and breakwaters, the Hudson (1959) formula has been widely used probably because of its simplicity. However, it 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. To solve the shortcomings in the Hudson formula, Van der Meer (1987) proposed a new design formula, based on a series of more than 250 model tests of Van der Meer (1988), which includes the influence of wave period, number of waves, armor grading, wave spectrum shape, groupiness of waves and the permeability of the core.

The model test data of Van der Meer (1988) have been used by many researchers in different types of research. Yoo et al. (2001) used the data to develop a new design formula, which is simpler than and comparable in accuracy to the Van der Meer formula (see Suh and Yoo, 2003). The data have also been used in the development of artificial neural network (ANN) model (Mase et al., 1995; Kim and Park, 2005; Balas et al., 2010), fuzzy model (Erdik, 2009), or M5′ model tree (Etemad-Shahidi and Bonakdar, 2009) to predict the damage level or stability number of rock armors. Especially Balas et al. (2010) developed a hybrid ANN model with principal component analysis (PCA) to improve the generalization performance of the ANN model. They used the five principal components (PCs) transformed from the five parameters used in the design formula of Van der Meer (i.e., permeability of core P; damage level S; number of waves N_w; structure slope $\cot\alpha$; and surf similarity parameter ξ_m) as the input variables of the ANN model for predicting the stability number. They also tested the model with four PCs transformed from four parameters (excluding the number of waves). They demonstrated that the predicting ability of neural network models is enhanced with the use of PCA when compared with the neural networks trained by the untreated data set and that it is also enhanced with increasing number of PCs.

The basic concept of choosing input variables in empirical formulas and ANN models is to select the variables that have great influence on the output variable and eliminate the others. In this 'select or eliminate', so-called variable selection method, especially in the case of ANN model, the variable set is often selected by comparing the results of all possible combinations of variable set. Although it can eliminate the less influential variables, it can lose the variables that have relatively small influence on the output 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. The PCA converts the high-dimension variable set to a low-dimension variable set that is mutually orthogonal. The orthogonality of PCs prevents the duplication of information by composing the variables into several independent components while maintaining the information in them. In this study, six parameters among the eleven parameters used in the tests of Van der Meer (1988) are transformed into five PCs by using the PCA. The

* Corresponding author.
E-mail address: kdsuh@snu.ac.kr (K.-D. Suh).

http://dx.doi.org/10.1016/j.oceaneng.2016.10.013
Received 9 November 2015; Received in revised form 30 September 2016; Accepted 4 October 2016

remaining five parameters are used as the input variables without any treatment. Using the five PCs and five untreated parameters as the input variable set for the ANN model is expected to yield better predictability than using the five PCs of Balas et al. (2010) as well as than previous ANN models without PCA.

Even though the ANN models are implemented in some areas of coastal engineering, e.g. estimation of wave overtopping discharge (http://nn-overtopping.deltares.nl/), they are seldom used in the design of armor layers probably because the engineers are not familiar with the ANN models. The ANN model is known as a black box model which does not explain the physical process on how the model output is produced from the input variables. However, the numerical process between the input and output variables in the ANN model can be found explicitly. The practicing engineers who are familiar with the use of empirical formulas may prefer an explicit calculation method. The numerical process in the ANN model could be used by the engineers as an explicit method, even though it is not as simple as an empirical formula. In this study, we present the 5×6 weight matrix that transforms the six parameters in the hydraulic experiment into five PCs. We also present the synaptic weights, biases, and activation functions in the ANN model that give the best agreement between the observed and calculated stability numbers. The engineers could then use the information to explicitly and easily calculate the stability number for given input parameters without sufficient knowledge of ANN. A worked example is provided.

Finally, it may be worthwhile to introduce some ANN studies related to environmental issues, even though they are not directly related to the stability of armor units. Sahoo et al. (2006) used ANNs to assess flash floods and their attendant water quality parameters using measured data at a Hawaiian stream. He et al. (2011) applied ANNs to the estimation of monthly river total nitrogen concentrations in 59 river basins in Japan. Recently, Zhou et al. (2015) combined the PCA and ANN to improve the accuracy and stability of the ANN model for predicting chlorophyll-*a* concentration in Dianshan Lake, Shanghai, China. Kim and Seo (2015) classified the monitoring data of turbidity and chlorophyll-*a* according to the patterns of daily change by applying several clustering methods, and then developed ANNs using each classified data to forecast the one-time ahead values of turbidity and chlorophyll-*a* concentrations.

2. Previous studies for estimation of stability number

To protect a rubble mound structure from severe erosion due to wave attack, an armor layer is placed on the seaward side of the structure. The stability of the armor units is measured by a dimensionless number, so-called stability number, which 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 structure, $\Delta = \rho_s/\rho_w - 1$ is the relative mass density, ρ_s and ρ_w are the mass densities of armor unit and water, respectively, and D_{n50} is the nominal size of the armor unit. As shown in Eq. (1), the stability number is defined as the ratio of the significant wave height to the size of armor units. A larger stability number, therefore, signifies that the armor unit with that size is stable against higher waves, that is, the larger the stability number, the more stable the armor units against waves.

To estimate the stability number, it is required to determine the relationship between the stability number and other variables which would describe the characteristics of waves and structure. Unfortunately, the physical mechanism of displacement of armor units due to waves is so complicated that the analytic solution is hardly found. For this reason, plenty of experiments which include various physical factors of waves and structure were conducted to propose empirical formulas explaining the relationship. Hudson (1959) suggested an empirical formula:

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

where K_D is the stability coefficient which depends on the shape of the armor unit, placement method, the location at the structure (i.e. trunk or head), and whether the structure is subject to breaking wave or non-breaking wave. Even though it is very simple, the Hudson formula has been found to have a lot of shortcomings.

To solve the main shortcomings of the Hudson formula, Van der Meer (1988) conducted an extensive series of tests including the parameters which are considered to have significant effects on armor stability, and the empirical formula based on the experimental data was proposed by Van der Meer (1987) as follows.

$$N_s = \frac{1}{\sqrt{\xi_m}}\left[6.2P^{0.18}\left(\frac{S}{\sqrt{N_w}}\right)^{0.2}\right] \quad \text{for} \quad \xi_m < \xi_c \tag{3a}$$

$$N_s = 1.0P^{-0.13}\left(\frac{S}{\sqrt{N_w}}\right)^{0.2}\sqrt{\cot\alpha}\,\xi_m{}^P \quad \text{for} \quad \xi_m \ge \xi_c \tag{3b}$$

where $\xi_m = \tan\alpha/\sqrt{2\pi H_s/gT_m{}^2}$ is the surf similarity parameter based on the average wave period T_m, and $\xi_c = (6.2P^{0.31}\sqrt{\tan\alpha})^{1/(P+0.5)}$ is the critical surf similarity parameter indicating the transition from plunging waves to surging waves.

On the other hand, with the recent developments in computational intelligence, particularly in the area of machine learning, various data-driven models have been developed, based on the extensive experimental data of Van der Meer (1988), as described in the Introduction. A brief summary is given here only for the ANN models. Mase et al. (1995) constructed an ANN by the randomly selected 100 experimental data set of Van der Meer (1988) and by 5000 learning iteration. They used 579 experimental data excluding the data of low-crested structures. They employed six input variables: P, N_w, S, ξ_m, h/H_s, and the spectral parameter, where h is the water depth in front of the structure. Kim and Park (2005) followed the Mase et al. (1995) approach, but they used 641 data including low-crested structures. Believing that the predictability of a neural network increases as the input dimension increases, they split the surf similarity parameter into wave steepness and structure slope, and further the wave steepness into wave height and period. They showed that the ANN gives better performance as the input dimension increases. It is known that in general the bias error and variance error decreases and increases, respectively, with the increase of input dimension. If the decreasing rate of bias error is greater than the increasing rate of variance error, the overall error decreases, and vice versa (Geman et al., 1992). It seems that the former is true for the data of Van der Meer (1988). On the other hand, Balas et al. (2010) developed hybrid ANN models with PCA based on 554 data of Van der Meer (1988). They developed four different models by systematically reducing the data from 554 to 166 by using PCA or by using the PCs as the input variables of the ANN. Table 1 shows the correlation coefficients of different studies, which will be compared with that of the present study later.

Table 1
Correlation coefficients of different empirical formula or ANN models.

Author	Correlation coefficient	Remarks
Van der Meer (1987)	0.92 (Mase et al., 1995) 0.947 (Balas et al., 2010)	Empirical formula, Eq. (3) in this paper
Mase et al. (1995)	0.91	Including data of Smith et al. (1992)
Kim and Park (2005)	0.902–0.952	Including data of low-crested structures
Balas et al. (2010)	0.906–0.968	ANN-PCA hybrid models