Coastal Engineering 59 (2011) 38–45

Contents lists available at ScienceDirect

Coastal Engineering

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

Stability of rubble-mound breakwater using H_{50} wave height parameter

Amir Etemad-Shahidi *, Meysam Bali

Enviro-hydroinformatics COE, School of Civil Engineering, Iran University of Science and Technology, Tehran, P.O. Box 16765-163, Iran

ARTICLE INFO

Article history:
Received 9 February 2011
Received in revised form 30 June 2011
Accepted 6 July 2011
Available online 20 August 2011

Keywords:
Breakwater
Stability number
Wave height
M5′ tree model
Armor

ABSTRACT

The prediction of rubble mound breakwaters' stability is one of the most important issues in coastal and maritime engineering. The stability of breakwaters strongly depends on the wave height. Therefore, selection of an appropriate wave height parameter is very vital in the prediction of stability number. In this study, H_{50}, the average of the 50 highest waves that reach the breakwater in its useful life, was used to predict the stability of the armor layer. First, H_{50} was used instead of the significant wave height in the most recent stability formulas. It was found that this modification yields more accurate results. Then, for further improvement of the results, two formulas were developed using model tree.

To develop the new formulas, two experimental data sets of irregular waves were used. Results indicated that the proposed formulas are more accurate than the previous ones for the prediction of the stability parameter. Finally, the proposed formulas were applied to regular waves and a wide range of damage levels and it was seen that the developed formulas are applicable in these cases as well.

1. Introduction

Conventional rubble mound breakwater is the most common type of breakwaters constructed around the world. It is composed of the core, filter and armor layer. One of the most important issues in breakwaters design is the determination of the armor block's weight using the stability number, N_s. Stability number is usually computed by the well-known empirical formulas of Hudson (1958) or Van der Meer (1988a, 1988b). According to the Hudson (1958) formula, the stability number depends on the significant wave height, armor type, damage level and slope of breakwater. Van der Meer (1988a) (hereafter VdM) considered type of breaker, permeability parameter, number of waves and surf similarity parameter in his formulas for irregular waves. His formulas were originally obtained for Rayleigh distribution in deep water, while most of the breakwaters are located in the intermediate or shallow waters which Rayleigh-distribution may be not valid (Vidal et al., 2006). Therefore, VdM suggested use of $H_{2\%}$ (average of the highest 2% of incident waves) instead of significant wave height H_s in these cases. Although this approach improved the predictions, $H_{2\%}$ does not fully consider the changes in wave height distribution in coastal zone (Vidal et al., 2006). Still there is a disagreement between the measured stability numbers and predicted ones (Kim and Park, 2005). As a result, a number of studies have been conducted to improve the prediction of stability number. Vidal et al. (1995) suggested use of H_n instead of H_s as an appropriate wave height parameter for the design of breakwaters in intermediate or shallow water. H_n is the average of the “n” highest waves that reach the breakwater during its useful life. First they suggested $n = 100$ and after some limited laboratory works with regular and irregular waves, Jensen et al. (1996) suggested $n = 250$. Finally, Vidal et al. (2006) (hereafter VML) replaced H_s with H_{50}. According to Thomson and Shuttler (1975) laboratory data, VML concluded that H_{50} can be used for the damage prediction and is independent of the wave distribution (Vidal et al., 2006). Soft computing models have also been used to improve the accuracy of stability number prediction. Mase et al. (1995) tried a new method for assessment of the stability number applying Artificial Neural Networks (ANNs) which was not successful completely. In another study, Kim and Park (2005) developed different ANN models that yield better results than VdM's when considering more governing parameters. Kim et al. (2008) examined the ability of probabilistic neural network methods to predict the stability number for specific numbers of waves. Fuzzy inference system was also used to improve the stability number prediction (Erdik, 2009) which led to the generation of 371 if–then rules. Although these soft computing methods were more accurate, they could not offer insight into the developed model. M5′ model tree (Wang and Witten, 1997) is a more transparent method that can provide formulas (e.g. Bhattacharya et al., 2007; Etemad-Shahidi and Mahjoobi, 2009). Etemad-Shahidi and Bonakdar (2009) (hereafter EB) proposed new formulas using M5′ algorithm which outperformed VdM's formulas. Nevertheless, it seems that the existing formulas may not be faultless, especially for intermediate and shallow waters, because of using H_s in the prediction of the stability number.

In the present study, first EB's formulas were modified by using H_{50} instead of H_s. Then, a set of new formulas were developed using M5′ algorithm. To develop the formulas, a combination of VdM's and

* Corresponding author. Tel.: +98 21 77240399; fax: +98 21 77240398.
E-mail address: etemad@iust.ac.ir (A. Etemad-Shahidi).

0378-3839/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.coastaleng.2011.07.002

VML's laboratory data for irregular waves with damage levels between 2 and 8 was used. Finally, the performances of the obtained formulas as well as other formulas were evaluated for experimental tests with regular waves and a wider range of the damage levels.

2. Existing formulas for stability number prediction

2.1. The formulas of Van der Meer (1988a, 1988b)

VdM proposed his method to determine the stability of breakwaters by analyzing a large number of irregular wave tests on rock stability (Van der Meer, 1988b). He proposed the following formulas for plunging breakers and surging breakers as:

$$N_s = 6.2S^{0.2}P^{0.18}N_w^{-0.1}\zeta_m^{-0.5} \quad \text{If } (\zeta_m<\zeta_{mc}) \text{ and } \cot\alpha \le 4 \tag{1a}$$

$$N_s = 1.0S^{0.2}P^{-0.13}N_w^{-0.1}\zeta_m^{p}\cot\alpha^{0.5} \quad \text{If } (\zeta_m\ge\zeta_{mc}) \text{ or } \cot\alpha \ge 4 \tag{1b}$$

where N_w is the number of wave attack, P is the nominal permeability of breakwater, ξ_m is the surf similarity parameter,cot α is slope angle, S is the damage level and N_s is the stability number defined as:

$$N_s = \frac{H_s}{\Delta D_{n50}}. \tag{2}$$

In the above equation D_{n50} is the nominal diameter of stone, $\Delta=\rho_s/\rho_w-1$ is the relative density of stone, ρ_s is the mass density of rock and ρ_w is the water density.

The damage level, S is defined by using the eroded area (A) of the breakwater cross-section as follows:

$$S = \frac{A}{D_{n50}^2}. \tag{3}$$

ζ_m is the surf similarity parameter based on the mean wave period (T_m):

$$\zeta_m = \frac{\tan\alpha}{\sqrt{2\pi Hs / gT_m^2}}. \tag{4}$$

The transition condition of surf similarity was expressed as (Van der Meer, 1988a)

$$\zeta_{mc} = \left(6.2P^{0.31}\tan\alpha^{0.5}\right)^{1/(p\,+\,0.5)}. \tag{5}$$

The permeability parameter in Eqs. (1a) and (1b) depends on the permeability of structure. The suggested values of P range from 0.1 for a relatively impermeable core to 0.6 for homogenous rock structures (Van der Meer, 1988b). Structures with filter layer between armor and core layers are represented by $P=0.40$–0.5. The factor 6.2 in Eq. (1a) and the factor 1.0 in Eq. (1b) are normally distributed with standard deviation 0.4 and 0.08, respectively

2.2. The formulas of Vidal et al. (2006)

VML's formulas are obtained by using H_{50} instead of H_s in the VdM's formulas.

$$N_{50} = 4.44S^{0.2}P^{0.18}\zeta_m^{-0.5} \quad \text{If } (\zeta_m<\zeta_{mc}) \text{ and } \cot\alpha \le 4 \tag{6a}$$

$$N_{50} = 0.716S^{0.2}P^{-0.13}\xi_m^{p}\cot\alpha^{0.5} \quad \text{If } (\zeta_m\ge\zeta_{mc}) \text{ or } \cot\alpha \ge 4 \tag{6b}$$

where N_{50} is defined as:

$$N_{50} = \frac{H_{50}}{\Delta D_{n50}}. \tag{7}$$

VML's formulas are independent of the storm duration and can be used for non Rayleigh-distribution cases, such as regular waves and shallow water waves (Vidal et al., 2006). VML showed that their formulas are more accurate than the VdM ones and can be used for any case (Vidal et al., 2006). The factors of 4.44 in Eq. (6a) and 0.716 in Eq. (6b) are normally distributed with standard deviation 0.29 and 0.06, respectively.

2.3. The formulas of Etemad-Shahidi and Bonakdar (2009)

EB, using M5′ algorithm, proposed a new set of formulas for prediction of stability number as follow:

$$Ns = 3.6N_w^{-0.09}P^{0.2}\cot\alpha^{0.17}S^{0.16}\xi_m^{-0.13} \quad \textit{If } S <= 2.95 \tag{8a}$$

$$Ns = 4N_w^{-0.08}p^{0.04}\cot\alpha^{0.02}S^{0.17}\xi_m^{-0.45} \quad \textit{If } S > 2.95 \textit{ and } \xi_m <= 1.52 \tag{8b}$$

$$Ns = 3.3N_w^{-0.07}P^{0.03}\cot\alpha^{0.02}S^{0.17}\xi_m^{-0.34} \quad \textit{If } S > 2.95 \textit{ and } \xi_m > 1.52 \textit{ and } P <= 0.22 \tag{8c}$$

$$Ns = 3.8N_w^{-0.06}P^{0.69}\cot\alpha^{0.3}S^{0.14}\xi_m^{-0.05} \quad \textit{If } S > 2.95 \textit{ and } \xi_m > 1.52 \textit{ and } p > 0.22. \tag{8d}$$

These equations were derived using VdM's laboratory data. The main advantage was their accuracy compared to that of VdM's.

3. Modification of EB's formula using H_{50}

VML showed that H_{50} is a more appropriate wave parameter in calculating the stability number. They proved that there is no need to consider number of waves provided that H_{50} is used instead of H_s.

For a rubble-mound breakwater with a certain damage level and surf similarity parameter, the relationship between the wave significant height and the number of waves in Eqs. (8a), (8b), (8c) and (8d) are given by:

$$H_sN_w^{0.09} = AK_1S^{0.16} \tag{9a}$$

$$H_sN_w^{0.08} = BK_2S^{0.17} \tag{9b}$$

$$H_sN_w^{0.07} = CK_3S^{0.17} \tag{9c}$$

$$H_sN_w^{0.06} = DK_4S^{0.14} \tag{9d}$$

where A, B, C and D are 3.6, 4, 3.3 and 3.8, respectively. K_1, K_2, K_3 and K_4 are constants. If H_{50} is used instead of H_s, N can be excluded from Eqs. (9a), (9b), (9c), and (9d):

$$H_{50} = A'K_1S^{0.16} \tag{10a}$$

$$H_{50} = B'K_2S^{0.17} \tag{10b}$$

$$H_{50} = C'K_3S^{0.17} \tag{10c}$$

$$H_{50} = D'K_4S^{0.14} \tag{10d}$$

A', B', C', D' are the new coefficients that can be determined by combining Eqs. (9a), (9b), (9c), and (9d) and (10a), (10b), (10c), and (10d) i.e.:

$$\frac{H_{50}}{H_SN_w^{0.09}} = \frac{A'}{A} \tag{11a}$$

$$\frac{H_{50}}{H_S N_w^{0.08}} = \frac{B'}{B} \quad (11b)$$

$$\frac{H_{50}}{H_S N_w^{0.07}} = \frac{C'}{C} \quad (11c)$$

$$\frac{H_{50}}{H_S N_w^{0.06}} = \frac{D'}{D}. \quad (11d)$$

The values of $\frac{A'}{A}, \frac{B'}{B}, \frac{C'}{C}, \frac{D'}{D}$ can be determined based on the waves distribution. For N_w varying between 500 and 5000 and assuming Rayleigh-distribution (Vidal et al., 2006), mean and standard deviation for values of $\frac{A'}{A}, \frac{B'}{B}, \frac{C'}{C}, \frac{D'}{D}$ can be calculated using Massel's (1996) approach. In this approach it is assumed that:

$$H_n = H_{\frac{n}{N_w}} = H_{\frac{1}{M}} = \left(\frac{M\sqrt{\pi}}{2} erfc\left(\sqrt{lnM}\right) + \sqrt{lnM}\right) H_{rms} \quad (12)$$

where $M = N_W/n$ and erfc are the complementary error function (Vidal et al., 2006). Fig. 1 shows the ratio of H_{50}/H_S against the variation of N_w. It is interesting to note that for N_W equals to 1000, this ratio becomes 1.41, i.e. the ratio between $H_{2\%}$ and H_S in Rayleigh distribution. This may justify the use of $H_{2\%}/1.41$ instead of H_S in shallow waters suggested by VdM.

In this way the mean and standard deviation were obtained as follows:

$$\left(\frac{A'}{A}\right) = 0.77 \pm 0.01 \quad (13a)$$

$$\left(\frac{B'}{B}\right) = 0.83 \pm 0.015 \quad (13b)$$

$$\left(\frac{C'}{C}\right) = 0.90 \pm 0.021 \quad (13c)$$

$$\left(\frac{D'}{D}\right) = 0.97 \pm 0.029. \quad (13d)$$

Using the mean values of $\frac{A'}{A}, \frac{B'}{B}, \frac{C'}{C}, \frac{D'}{D}$, new coefficients (A', B', C', D') were obtained and EB's formulas were modified as (here after called MT1):

$$N_{50} = 2.8P^{0.2} \cot\alpha^{0.17} S^{0.16} \xi_m^{-0.13} \; If\; S <= 2.95 \quad (14a)$$

$$N_{50} = 3.3p^{0.04} \cot\alpha^{0.02} S^{0.17} \xi_m^{-0.45} \; If\; S > 2.95 \; and\; \xi_m <= 1.52 \quad (14b)$$

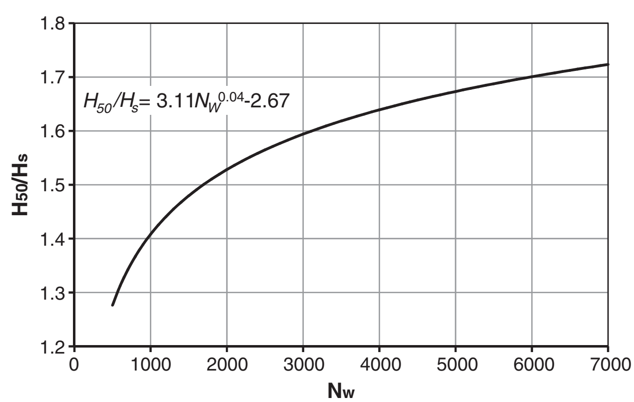

Fig. 1. The ratio of H_{50}/H_s against N_w.

$$N_{50} = 3.0P^{0.03} \cot\alpha^{0.02} S^{0.17} \xi_m^{-0.34} \; If S > 2.95 and\; \xi_m > 1.52\; and\; P <= 0.22 \quad (14c)$$

$$N_{50} = 3.7P^{0.69} \cot\alpha^{0.3} S^{0.14} \xi_m^{-0.05} \; If\; S > 2.95\; and\; \xi_m > 1.52\; and\; p > 0.22. \quad (14d)$$

4. The used data sets

A combination of VdM's and VML's laboratory data sets was used for the evaluation of the mentioned formulas. VdM's data contains 579 data for irregular wave's tests. VML's data consists of regular and irregular waves with three types of tests: (a) twelve regular wave tests with two surf similarity parameters (108 data), (b) twelve irregular wave tests with two surf similarity parameters (97 data) and (c) two long irregular wave tests with two surf similarity parameters (65 data). Overall, 270 data were extracted from the tests with these parameters: $P = 0.45$, $\cot\alpha = 1.5$, $\Delta = 1.7$, $D_{n50} = 2.95$ cm. More details of the tests are given in Vidal et al. (2006). Table 1 shows the ranges of different parameters of the data set. Tests with very low damage level ($S<2$) and very high damage level ($S>8$), which are not common in design practices, were not used at the first stage. Therefore, 265 data point out of VdM's data set and 41 data out of VML's were considered for the processing.

5. Evaluation of formulas

A total of 306 irregular wave data with $0 \le S \le 8$ were used for evaluation of different formulas. Fig. 2 shows the scatter between the measured and predicted stability numbers for the existing formulas. Fig. 2a and b show that both VdM's and EB's formulas underestimate the VML data and EB's formulas are more accurate in the prediction of VdM data compared to VdM's ones. Fig. 2b also shows EB's formulas underestimate VML's data but generally yield more accurate predictions for stability compared to those of VdM.

VML's formulas are accurate for their data, but they do not have adequate accuracy for VdM ones (Fig. 2c). In addition, VML formulas underestimate high stability numbers and are conservative in this range.

Fig. 2d indicates that generally the modified EB's formulas (MT1) yield more accurate predictions of the stability. Compared to other formulas, MT1 shows less scatter for both high and low values of the stability number. The higher performance of MT1 is mainly due to the use of wave parameter, H_{50}, suggested by VML instead of H_s in the formulas suggested by EB.

The performance of different approaches was also judged quantitatively using error measures such as *BIAS*, scatter index (*SI*), correlation coefficient (*CC*) and agreement index (I_a) defined as:

$$BIAS = \sum_{i=1}^{N} \frac{1}{N}(Y_i - X_i) \quad (15)$$

Table 1
Ranges of the used parameters.

Parameters	Range (train)	Range (test)	Mean	Median	Std
Nw	1000–3000	1000–3000	1771	1000	957.01
P	0.1–0.6	0.1–0.6	0.30	0.45	0.28
ξ_m	0.67–6.96	0.67–6.58	3.00	2.89	0.86
$\cot\alpha$	1.5–6	1.5–6	2.81	3.00	1.26
S	2–8	2–8	4.57	4.37	1.70
N_s	1.07–3.61	1.16–3.55	2.14	2.10	0.54
N_{50}	1.51–5.06	1.63–4.98	2.92	2.92	0.7

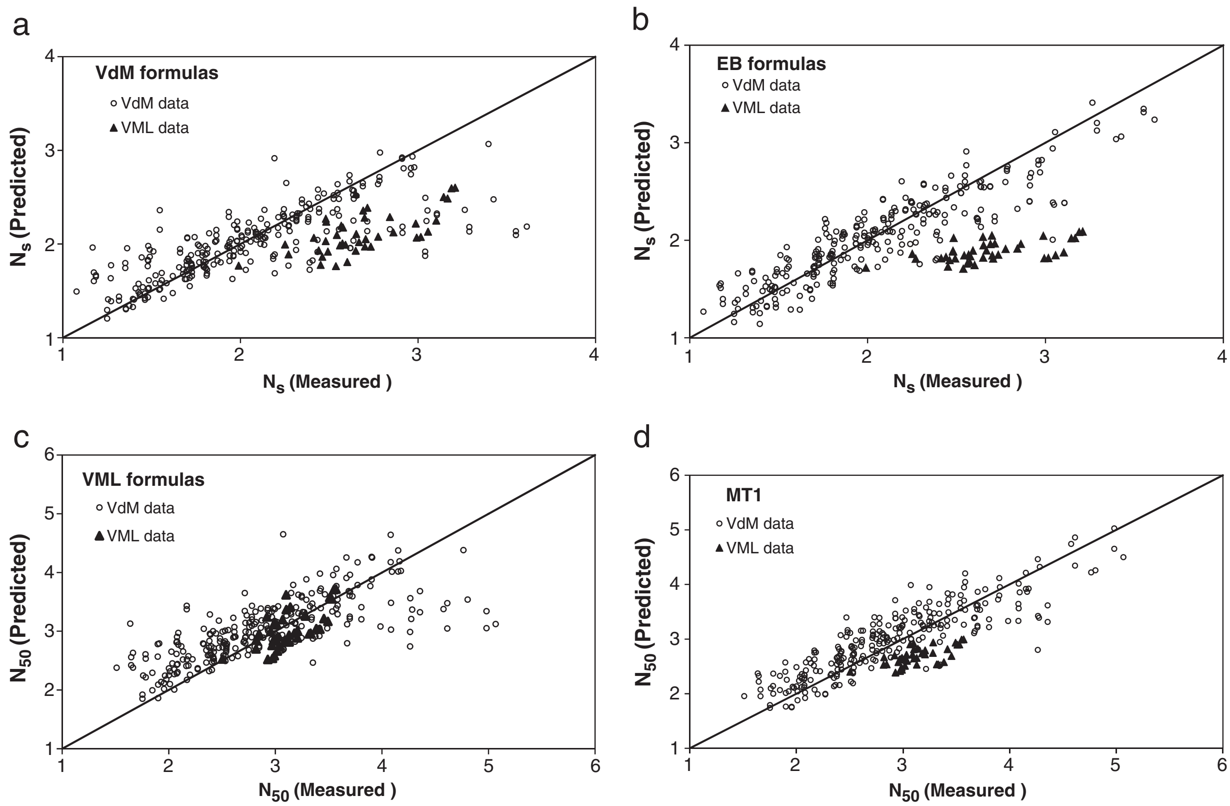

Fig. 2. Comparison between the measured and predicted stability numbers for $2 \leq S \leq 8$ (a) VdM's formulas (b) EB's formulas(c) VML's formulas and (d) MT1.

$$SI = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(Y_i - X_i)^2}}{\overline{X}_i} \quad (16)$$

$$CC = \frac{\sum_{i=1}^{N}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{N}(X_i - \overline{X})^2 \sum_{i=1}^{N}(Y_i - \overline{Y})^2}} \quad (17)$$

$$I_a = 1 - \frac{\sum_{i=1}^{N}(Y_i - X_i)^2}{\sum_{i=1}^{N}\left(|Y_i - \overline{X}| + |X_i - \overline{X}|\right)^2} \quad (18)$$

where X_i and Y_i denote the measured and predicted values respectively and N is the number of observations. $\overline{X}$ and $\overline{Y}$ are the mean values of the measured and predicted parameters, respectively. Table 2 shows the error measures of MT1 and previous equations for the used set data. It can be seen the *BIAS* and *SI* of MT1 are less than those of other formulas. The *BIAS* and *SI*are largely reduced using MT1. In addition, The *CC* and I_a of MT1 are larger than those of others. Generally, it can be concluded that MT1 is more accurate than the previous equations. Although acceptable results are obtained using H_{50} in EB's formulas, but a new set of formulas will be developed by model tree in the next section.

6. Model tree

The concept of model tree (MT) approach is based on dividing complex problems into smaller sub problems and solving each sub problem (Bhattacharya et al., 2007). The input parameter space is divided into smaller subspaces and providing a multiple linear regression model for them. M5 algorithm is one of the most commonly used algorithms of MT which was first proposed by Quinlan (1992). It was then improved and a more complete algorithm called M5′ was introduced by Wang and Witten (1997). This algorithm has similar structures to the M5 algorithm, but is able to deal effectively with missing values (Jang et al., 2009). After dividing the input space into sub-space, a set of piecewise linear models is considered as the final solution of the problem. M5′ model tree includes three steps: building, pruning and smoothing the tree. Standard deviation reduction, *SDR* factor, is used to grow the basic tree at first. *SDR* is defined as follows:

$$SDR = sd(T) - \sum_i \frac{|T_i|}{|T|} \times sd(T_i) \quad (19)$$

where T is the set of data point before splitting, T_i is data point which is the result of splitting the space and fall into one sub-space according to the chosen splitting parameter and *sd* is the standard deviation. Standard deviation is used as an error measure for the data points of a sub-space. M5′ model tree tests different splitting points by calculating *sd* for sub-spaces before dividing the space. As *SDR* is maximized in a point, the point is selected as the splitting point

Table 2
Error measures of different formulas.

Formula	*BIAS*	*SI*	*CC*	I_a
VdM	−0.11	0.18	0.73	0.80
VML	0.10	0.17	0.72	0.81
EB	−0.13	0.17	0.76	0.83
MT1	0.02	0.13	0.84	0.91
MT2	−0.02	0.10	0.89	0.94
MT2, test data	−0.02	0.11	0.88	0.93

(node). The splitting stops when *SDR*changes is less than a certain value or a few data points remain in sub-space. The accuracy of the model for training set increases uniformly as the tree grows. However, over-fitting may be inevitable while a model tree is being built. Hence, in the second step, pruning is used to avoid over-fitting. Prediction of the expected error at each node for the test data is used for the pruning procedure. For each training data, the predicted value is calculated. To prevent underestimation of the expected error, the predicted value is multiplied by the factor of $(n+\nu)/(n-\nu)$ where n is the number of training data that reach the node and ν is the number of parameter in the model that represents the value at the node (Wang and Witten, 1997). The sub-space can be pruned if predicted error is lower than the expected one (Witten and Frank, 2005). The final step, smoothing, is the regularization process to compensate any probable discontinuities among adjacent linear models. The smoothing procedure uses the models built in each sub-space (leaf) to compute the predicted value. The obtained value is then modified along the route back to the root of the tree by smoothing it at each node. The value at each node is combined with the output value by the linear model for that node as$P'=(np+kq)/(n+k)$, where $P'$is prediction passed up to the next higher node, p is prediction passed node from below, q is the value predicted by the model at this node, n is the number of training data points that reaches the node below and k is a constant (Wang and Witten, 1997). The predicted value by the leaf is combined with that of linear model for each node on the top of the leaf to the root. This smoothing process usually improves the predictions (Quinlan, 1992).

7. Development of formulas

The combined data set of irregular waves with $0\leq S\leq 8$ (306 data points) was used to develop and evaluate the stability number predictor model. From these data points, 204 data points were selected for model training. The ranges of different parameters used for test and train are presented in Table 1. As seen, the selected parameters are the same as those used by previous investigators and have wide ranges. The M5′ model tree produced only linear relationships between the input and output parameters. Previous studies have shown that the relationship between the governing parameters and the stability number is not linear, i.e. $N_{50}=aP^bS^c\xi_m^d\cot\alpha^e$.To overcomes this limitation, the parameters have been used in their logarithmic forms. Dimensionless parameters used to develop the formulas include N_{50}, P, S, ζ_m and $\cot\alpha$. The developed formulas (MT2) were:

$$\text{If } \xi_m \leq 2 \quad N_{50} = 4.18P^{0.16}S^{0.18}\xi_m^{-0.53} \tag{20a}$$

$$\text{If } \xi_m > 2 \quad N_{50} = 3.57P^{0.2}S^{0.18}\xi_m^{-0.24}. \tag{20b}$$

The splitting parameter and its value do not necessarily have any physical interpretation since they are obtained by minimizing the prediction error (Bhattacharya et al., 2007). However, it is seen that M5′ was successful at distinguishing between different regimes. The surf similarity parameter, an indicator of breaker type, is the only splitting parameter in the derived equations. This is due to the importance of wave breaking condition on the stability of rubble mound structures. As discussed by Van der Meer and Janssen (1994), when ζ_m is less than about 2, waves break on the slope of the structure and when ζ_m is greater than 2 the waves do not break on the slope (see also Van der Meer, 1988a, 1988b). Hence, the obtained regimes are justified and basically the model gives different formulas for plunging and surging waves. These simple and compact formulas are in good agreement with engineering sense and previous knowledge about role of different parameters in the stability of armor layer. They indicate that the stability number increases by an increase of the permeability of the breakwater and/or the accepted damage level.

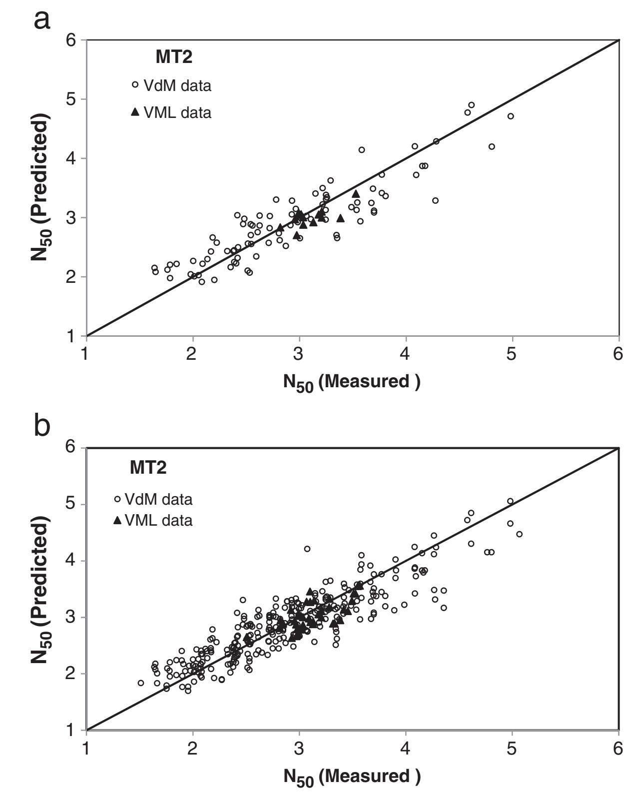

Fig. 3. Comparison between the measured and predicted stability numbers, for $2\leq S\leq 8$ MT2 (a) test data and (b) all data.

They also show that by increasing the surf similarity parameter (increase of the structure slope or wave period) the stability decreases. The role of damage level is the same for plunging and surging waves. However, for surging waves, the importance of ζ_m becomes less and that of permeability slightly increases since the surging waves move more up and down (on the slope). The obtained formulas are also supported by the previous findings regarding the exponents of the governing parameters. For example, both the exponents of P and S are about 0.18, a value very close to the suggested values of VdM and EB.

Fig. 3 shows the scatter between the measured and predicted stability numbers by Eqs. (20a) and (20b) (MT2). As seen, the prediction of the stability number is improved even though the formulas are simpler than the previous ones. The accuracy of this model was evaluated as well and the error measures showed an improvement in the prediction of N_S (Table 2). In comparison with MT1, the scatter index is reduced about 15% while the correlation coefficient and agreement index values are increased by 6% and 5%, respectively. In addition, the performance of the developed formulas is nearly the same for different experimental datasets and generally the results are promising. The uncertainties in the developed formulas are also minimal. Factors of 4.18 in Eq. (20a) and 3.57 in Eq. (20b) are normally distributed with standard deviations of 0.41 (coefficient of variation 0.1) and 0.40 (coefficient of variation 0.11), respectively.

Table 3
Ranges of parameters for all irregular wave data.

Parameters	Range	Mean	Median	Std
Nw	500–3000	1766	1000	54.77
P	0.1–0.6	0.30	0.45	0.32
ξ_m	0.2–7.58	2.91	2.77	2.04
$\cot\alpha$	1.5–6	2.67	2.00	1.41
S	0–32.9	6.11	4.08	2.55
N_s	0.72–4.37	2.12	2.07	1.15
N_{50}	0.78–6.15	2.87	2.77	1.35

Table 4
Error measures of different formulas for all irregular wave data.

Formulas	*BIAS′*	*SI*	*CC*	I_a
VdM	−0.14	0.21	0.72	0.82
VML	0.09	0.19	0.77	0.87
EB	−0.17	0.19	0.8	0.87
MT1	0.02	0.14	0.87	0.93
MT2	−0.05	0.12	0.91	0.95

These values can be used for probabilistic design of armors (Van der Meer, 1988b). For example if only 5% risk is acceptable, the obtained rock diameters should be multiplied by a factor of 1.17.

7.1. Evaluation for a wider range of damage levels and regular wave data

In this section, the performances of different formulas were evaluated for a wider range of damage levels, i.e. $0<S\leq32.9$. Zero damage level is not considered for the design purposes; therefore it was omitted from the analysis. Hence, 704 data were considered for the evaluation. Table 3 shows that the ranges of parameters are wider than those used for developing MT2 formulas. Fig. 4 displays the comparison between the measured and predicted dimensionless stability numbers by different formulas. As seen, VML's equations overestimate the stability parameter and other formulas mostly underestimate it and are more conservative. In addition, most of the data points are concentrated on the line of the perfect agreement when using Eqs. (20a) and (20b) and MT2 predicts the stability much better than others' formulas. Table 4 displays the error measures of all formulas for the prediction of stability number. As can be seen, even the MT2 is not developed for this range, and the accuracies are similar to those of previous case ($0\leq S\leq8$). *BIAS* and *SI* of MT2 are less than those of other formulas in this wider range of parameters; and the magnitude of correlation coefficient and agreement index is higher.

In the previous part, it was shown that H_{50} is an appropriate wave height parameter to calculate the stability number for irregular waves.

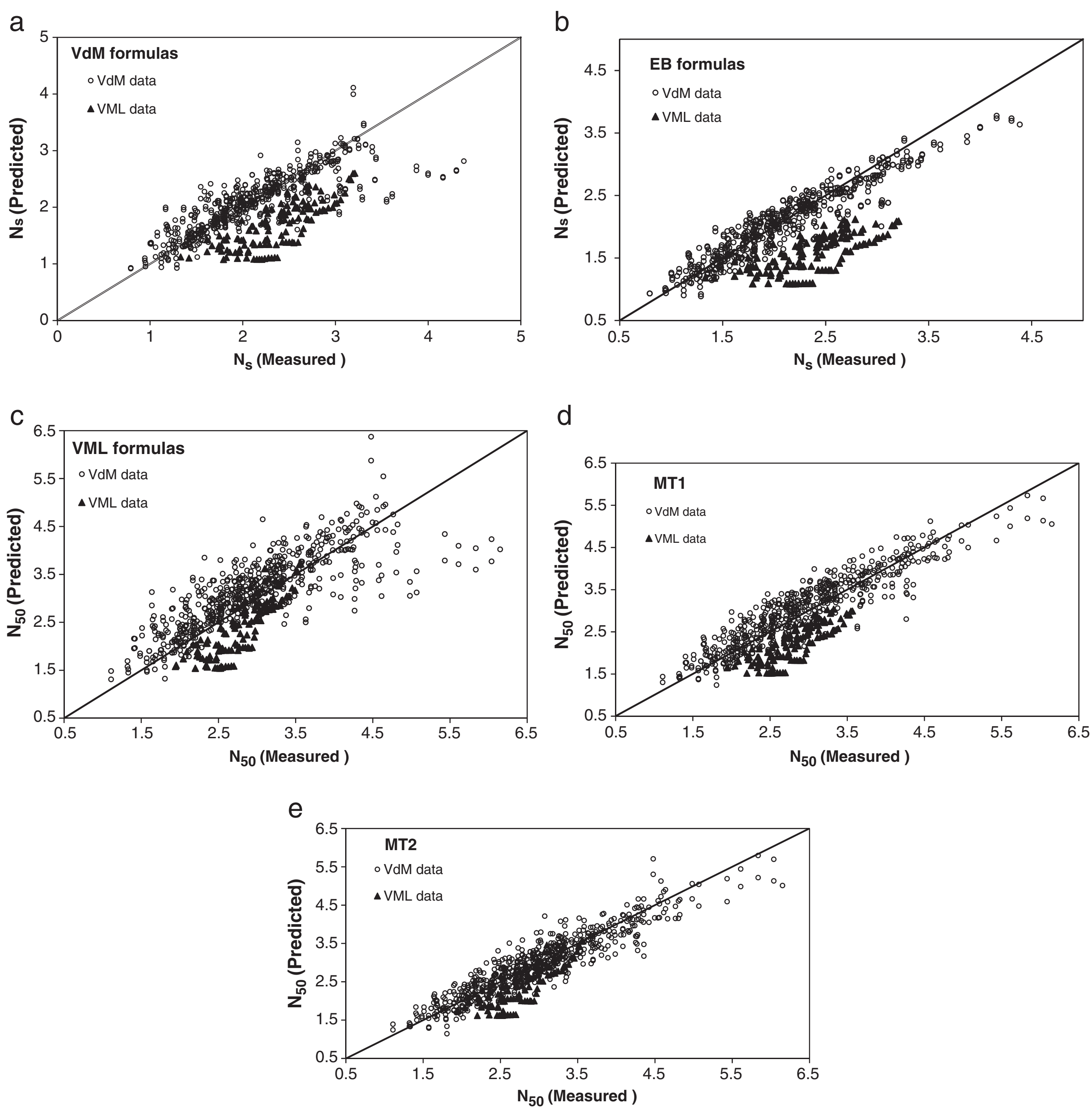

Fig. 4. Comparison between the measured and predicted stability numbers for $0<S\leq32.9$ (a) VdM's formulas (b) EB's formulas (c) VML's formulas (d) MT1 and (e) MT2.

Table 5
Ranges of parameters for regular wave data.

Parameters	Range	Mean	Median	Std
Nw	500	500	500	0.00
P	0.45	0.45	0.45	0.00
ξ_m	2.53–3.85	3.16	2.74	0.55
$cot\ \alpha$	1.5	1.5	1.5	0.00
S	0.1–15.25	2.73	1.68	3.40
N_s	1.42–3.3	2.41	2.48	0.39
N_{50}	1.47–3.3	2.50	2.56	0.40

The aim, here, is to examine the performance of the developed formulas for non Rayleigh-distribution such as regular waves. Overall, 66 regular wave data with non zero damage levels were extracted from the VML's tests for this purpose. Table 5 shows the ranges of parameters for regular wave data. Table 6 shows the error measures of different formulas for regular tests. As seen, all formulas under-predict the stability numbers and are conservative for the regular wave condition. However, the performance of MT2 is better than those of others. Although CC value is not the highest one, the highest value of I_a and lowest values of $BIAS$ and SI of the developed formulas show its skill.

Fig. 5 shows the difference between different formulas for the stability number in typical design conditions ($S=2$, $P=0.45$, $Nw=3000$, $cot\ \alpha=2$ and Rayleigh distribution). As seen, the results obtained using formulas of VdM and VML are the same since the wave height distribution is assumed to be Rayleigh's. The other two formulas, i.e. MT2 and EB yield nearly similar results specially for surging breakers. By increasing the surf similarity parameter, the runup/rundown and the flow velocity increase (CEM, 2006). Therefore, the stability number becomes less. This point is clearly seen in stability numbers predicted by MT2 at high values of ξ_m which is mainly due to the increase of (both upward and downward) water velocities.

8. Summary and conclusion

Wave height is the most important parameter in the prediction of structural response of breakwaters. In this study, the effect of using H_{50}, the average of the 50 highest waves that reach the breakwater, in the prediction of stability number of rubble-mound breakwater was investigated. First, EB's formulas were modified successfully by using H_{50} instead of H_s. The results showed that modified EB formulas (MT1) improve the accuracy of the predictions. For further improvement, a new model was presented using M5′ model tree. A combination of VdM's and VML's laboratory data sets with damage levels between 2 and 8 were used to develop the new formulas (MT2). These compact and simple formulas were obtained in forms of power law and were in conformity with the previous knowledge of the effects of different parameters on stability number. It was also shown that MT2 is more accurate and simpler than MT1 and the previous empirical formulas. MT2 showed high performance for all cases leading to the reduction of systematic errors. In addition, the uncertainties of the given formulas were given for the application in probabilistic design of breakwaters' armors when a certain level of risk is desired. Although MT2 was developed for irregular waves with$2\leq S\leq 8$, it was shown that it can be used for a wider range of damage levels as well as regular waves. This study indicates that coastal engineering design practices can be improved by using the suggested approach, i.e. combining modern soft computing tools and physical arguments, in place of traditional statistical methods.

Table 6
Error measures of different formulas for regular wave data.

Formula	$BIAS$	SI	CC	I_a
VdM	−0.65	0.29	0.86	0.62
EB	−0.73	0.32	0.79	0.54
VML	−0.14	0.17	0.85	0.85
MT1	−0.25	0.16	0.80	0.82
MT2	−0.04	0.14	0.83	0.87

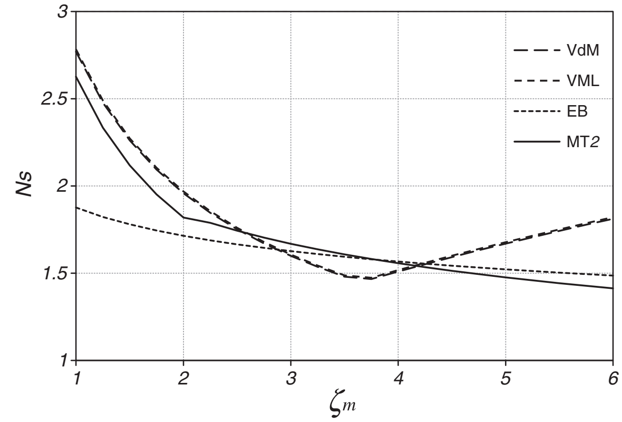

Fig. 5. Stability number by different formulas ($S=2$, $P=0.45$, $Nw=3000$, $cot\ \alpha=2$, Rayleigh distribution).

Acknowledgments

The authors express thanks to Professors Van der Meer and Vidal for their comprehensive and freely database of the stability of rubble-mound breakwater. The authors also thank Mohammad Kazeminezhad, Liasham Bonakdar, and Milad Rezaee for their help in improving the manuscript.

References

Bhattacharya, B., Price, R.K., Solomatine, D.P., 2007. Machine learning approach to modeling sediment transport. Journal of Hydraulic Engineering ASCE 133, 440–450.
Coastal Engineering Manual (CEM), 2006. Coastal Engineering Research Center. Department of Army Corps of Engineers, US. (Chapter VI, Part 5).
Erdik, T., 2009. Fuzzy logic approach to conventional rubble mound structures design. Expert Systems with Applications 36, 4162–4170.
Etemad-Shahidi, A., Bonakdar, L., 2009. Design of rubble mound breakwater using M5′ machine learning method. Applied Ocean Research 31, 197–201.
Etemad-Shahidi, A., Mahjoobi, J., 2009. Comparison between M5′ model tree and neural networks for prediction of significant wave height in lake superior. Ocean Engineering 36, 1175–1181.
Hudson, R.Y., 1958. Design of quarry stone cover layer for rubble mound breakwaters. Research Report No. 2–2, Waterways Experiment Station, Coastal Engineering Research Centre, Vicksburg, MS.
Jang, N., Popescu, I., Kelderman, P., Solomatine, D., Price, R., 2009. Application of model trees and other machine learning techniques for algal growth prediction in Yongdam reservoir, Republic of Korea. Journal of Hydroinformatics 1–13.
Jensen, T., Anderson, H., Gronbech, J., Mansard, Davis, M.H., 1996. Breakwater stability under regular and irregular wave attack. In: ASCE (Ed.), Proc. 25th Int. Conf. On Coastal Eng. (ICCE), pp. 1678–1962.
Kim, D.H., Park, W.S., 2005. Neural network for design and reliability analysis of rubble mound breakwaters. Ocean Engineering 32, 1332–1349.
Kim, D., Kim, D.H., Chang, S., 2008. Application of probabilistic neural network to design breakwater armor blocks. Ocean Engineering 35, 294–300.
Mase, H., Sakamoto, M., Sakai, T., 1995. Neural network for stability analysis of rubble-mound breakwaters. Journal of Waterway, Port, Coastal and Ocean Engineering 121, 294–299.
Massel, S.R., 1996. Ocean surface waves: their physics and prediction. World Scientific, 495 pp.
Quinlan, J.R., 1992. Learning with continuous class. In: Adams, Sterling (Eds.), Proceeding of AI'92. World Scientific.
Thomson, D.M., Shuttler, R.M., 1975. Riprap design for wind wave attack: a laboratory study on random waves. Report 61. HRS Wallingford, UK.
Van der Meer, J.W., 1988a. Stability of breakwater armour layers — design formula. Coastal Engineering 11, 219–239.
Van der Meer, J.W., 1988b. Rock slopes and gravel beaches under wave attack. PhD Thesis, Delft University of Technology, Delft, The Netherlands. Vidal, C., Losada, M.A., E.P.D., 1995. Suitable Wave Height Parameter for Characterizing Breakwater Stability. Journal of Waterway, Port, Coastal and Ocean Engineering, ASCE 121 (2), 88–97.
Van der Meer, J.W., Janssen, J., 1994. Wave run-up and wave overtopping at dikes and Revetments, Delft Hydraulics, publication no. 485.

Vidal, C., Losada, M.A., Mansard, E.P.D., 1995. Suitable wave height parameter for characterizing breakwater stability. Journal of Waterway, Port, Coastal and Ocean Engineering, ASCE 121 (2), 88–97 March/April.
Vidal, C., Medina, R., Lomonanco, P., 2006. Wave height parameter for damage description of rubble mound breakwater. Coastal Engineering 53, 712–722.
Wang, Y., Witten, I.H., 1997. Induction trees for predicting continuous classes. Proceeding of the Poster Papers of the European Conference on Machine Learning, 1997 Prague. University of Economics. Faculty of Informatics Statistics.
Witten, I.H., Frank, E., 2005. Data Mining — Practical Machine Learning Tools and Techniques. Morgan Kaufmann, San Francisco.