New Relationship for Estimation of Wave Overtopping on Vertical Walls

ESKANDARI Hojjatollah*, KAMALIAN Ulrich Reza

Coastal Engineering Group, University of Qom, Qom 3716146611, Iran

Abstract

Soft computing tools in the form of combination of multiple nonlinear regression and M5′ model tree were used for estimation of overtopping rate at the vertical coastal structures. For reliable and precise estimation of overtopping rate, the experimental data available in the database CLASH were used. The dimensionless overtopping rate was estimated in terms of conventional dimensionless parameters including the relative crest freeboard R_c/H_s, seabed slope $\tan\theta$, deep water wave steepness S_{om}, surf similarity ξ_{om} and local relative water depth h_t/H_s. The accuracy of the new model was compared with other existing models and also evaluated with some field measurements. The results indicated that the model presented in this paper is more accurate than other existing models. With statistical parameters, it is shown that the accuracy of predictions in the new model is better than that of other models.

Key words: CLASH database, M5′ model tree, wave overtopping, vertical coastal structures, multiple nonlinear regression

---

1 Introduction

Seawalls are usually massive, vertical structures which are used to protect backshore areas from heavy wave action, and in lower wave energy environments, to separate land from water. For example, caisson type structure as a quay-wall was used before construction of breakwaters exposed to the open sea in Iran (Alikhani et al., 2003).

Reliable estimation of wave overtopping rate is required for the planning and safety considerations of coastal structures. Overestimation of overtopping rate will cause additional cost, and underestimation will result in irreparable damage. The wave overtopping rate (q, in m³/s per unit width) is a fundamental design factor for coastal structures that are designed to restrict q below an allowable amount. Therefore, the prediction of wave overtopping has been the study topic for so many researchers (van der Meer and Bruce, 2014).

Verhaeghe (2005) presents a summary of various proposed formulae for the prediction of overtopping rate. Goda et al. (1975) prepared design diagrams for evaluation of overtopping by using the equivalent deep water wave height H_0' as the main parameter. EurOtop (2007) presented several wave overtopping formulae for various kinds of coastal structures by employing the significant wave at the toe of the structure, H_s as the key parameter.

Recently, artificial neural network (ANN) as a soft computing method has been employed to predict the mean wave overtopping rate for coastal structures. These studies have been fulfilled within the European project CLASH (de Rouck et al., 2009). Van Gent et al. (2007) presented an ANN model for various types of coastal structures. Verhaeghe et al. (2008) developed a two-phase neural model to predict the wave overtopping rates.

Here, by using soft computing tools in the form of combination of multiple nonlinear regression and M5′ model tree, the overtopping rate at vertical coastal structures was studied with the data from the CLASH database and the results were compared with others as well as available real field data.

2 Existing formulas for estimation of overtopping

The most well-known experimental models have been proposed for vertical caisson breakwaters by Franco et al. (1994) and EurOtop (2007). Franco et al.'s formula based on 80 experimental models done in Delft Hydraulics Institute of the Netherlands has been suggested as:

$$\frac{q}{\sqrt{gH_s^3}} = a\exp\left(-b\frac{R_c}{H_s}\right), \tag{1}$$

where q is the overtopping rate (m³/s per meter width), g is the acceleration of gravity, H_s is the significant wave height and R_c is the crest height. a=0.082 and b=3.0 are constant coefficients proposed based on the experimental conditions. Another formula to calculate the overtopping rate over caisson breakwaters is proposed by EurOtop (2007). First, h_* can be defined as:

$$h_* = 1.35 \frac{h_\mathrm{s}}{H_\mathrm{s}} \frac{2\pi h_\mathrm{s}}{gT_\mathrm{m}^2}. \tag{2}$$

where h_s is local water depth and T_m is average wave period.

In fact, this number is the product of two dimensionless parameters: the ratio of the water depth to wave height and the ratio of the water depth to deep water wavelength, $2\pi/\left(gT_\mathrm{m}^2\right)$ being the wavelength in deep water. Then the formulae for estimation of overtopping rate are as follows:

$$\text{if} \quad h_* > 0.3, \quad \frac{q}{\sqrt{gH_\mathrm{s}^3}} = 0.04 \exp\left(-2.6\frac{R_\mathrm{c}}{H_\mathrm{s}}\right); \tag{3}$$

$$\text{if} \quad h_* < 0.2, \quad \frac{q}{h_*^2\sqrt{gH_\mathrm{s}^3}} = 1.5\times10^{-4}\left(h_*\frac{R_\mathrm{c}}{H_\mathrm{s}}\right)^{-3.1}. \tag{4}$$

Eq. (3) is valid for $0.1 < R_\mathrm{c}/H_\mathrm{s} < 3.5$ and Eq. (4) is valid for $0.03 < R_\mathrm{c}/H_\mathrm{s} < 1$.

Another formula is provided by Goda (2009) as:

$$\frac{q}{\sqrt{gH_\mathrm{s(toe)}^3}} = \exp\left[-\left(A + B\frac{R_\mathrm{c}}{H_\mathrm{s(toe)}}\right)\right], \tag{5}$$

with

$$A = A_0 \tanh\left[b_1\left(\frac{h_\mathrm{t}}{H_\mathrm{s(toe)}} + c_1\right)\right]; \tag{6}$$

$$B = B_0 \tanh\left[b_2\left(\frac{h_\mathrm{t}}{H_\mathrm{s(toe)}} + c_2\right)\right]. \tag{7}$$

The constant coefficients in these equations are determined according to Table 1.

Table 1   Constant coefficients of Goda's formula

Slope	Factor A			Factor B		
$\tan\theta$	A_0	b_1	c_1	B_0	b_2	c_2
1:10	3.6	1.4	0.1	2.3	0.6	0.8
1:20–1:1000	3.6	1.0	0.6	2.3	0.8	0.6

Note: θ is the seabed slope and h_t is the water depth at the toe of the structure.

3 M5′ model tree and multiple regression

Decision tree is one of the most powerful and popular tools for classification and prediction of the data. Unlike neural networks in which input data are considered only as numbers, in model tree there is no limit to the type of data. The most important advantage of the decision tree to the neural network is about producing of rules. Decision tree reveals the predictions in the form of a few rules. For validation of rules, a series of test data which are not used to build the tree, are utilized.

The basic form of the model tree was presented by Quinlan (1986) as a new soft computing tool. These models fit linear function to the subset of data. The most important advantage of model trees compared with other soft computing tools is that they provide clear mathematical relationships. Recently these models have been used in the field of sediment transport (Bhattacharya et al., 2007), prediction of wave spectrum (Sakhare and Deo, 2009), and prediction of the significant wave height (Etemad-Shahidi and Mahjoobi, 2009). In this study, with M5′ model tree, separate models

are presented to predict the average rate of wave overtopping on the vertical structures.

4 Data selected from the database CLASH

The data used in this study are selected from the database created in the project CLASH (de Rouck et al., 2009). In this project an extensive database on the wave overtopping including field measurements and laboratory data has been developed. There are many experimental data in the database that have been obtained from 163 independent experiments performed in institutes and laboratories around the world. Further details about the collection of the database are presented by Verhaeghe (2005) and van der Meer et al. (2009).

In this study selected are the data involving wave overtopping on the vertical structures, without any upper structure with reverse slope or protecting rubble mound structure in front of the caisson. By using correction factor for the permeability and roughness (γ_f) and the slope of the front face of the structure ($\cot\alpha$), the data for this type of structure are classified. As a part of the project CLASH, a series of tests have been conducted to evaluate the combined effect of roughness and permeability of various types of structures. Structures with $\gamma_\mathrm{f}=1$ have been considered as smooth and impermeable (data used in this study) and structures with $\gamma_\mathrm{f} \leqslant 0.6$ have been considered as rubble mound structures.

Overtopping wave studies within the project CLASH show that, because of the effects of the model, scale, and errors caused by different measurement techniques, the real condition cannot be simulated in a laboratory accurately. For example, the effects of sea currents or winds are generally not taken into account in the lab. If the large-scale as well as small-scale data are used for model training, scale effects may affect the accuracy the model. To avoid misleading model, data for large-scale tests (including tests with the significant wave height larger than 5.0 m and field measurements) is not used for training model.

The Complexity-Factor (*CF*) and Reliability-Factor (*RF*) are allocated for each experiment in CLASH database. The complexity factor changes from one for structures with a very simple cross-section to four for structures with a very complex cross-section. Similarly, reliability factor varies from one to four for the very reliable to non-reliable tests, respectively. For this study the data with *CF*=4 and *RF*=4 have been excluded. Data concerning simple structures (without platform), single slope, without the reverse slope of the upper structure and the structure of the stone without protective structures have been selected.

5 Modeling

The experiments available in CLASH were performed in different scales (such as 1:20, 1:30, and 1:40). Therefore the measured overtopping rates are not comparable to each other, without converting to dimensionless parameters. Further-

more, the experiments were conducted in laboratories by small-scale models. Thus if a relationship is developed by these parameters, it is not applicable for large-scale models. Consequently we should use dimensionless parameters such as the relative crest freeboard R_c/H_s, seabed slope $\tan\theta$, deep water wave steepness S_{om}, surf similarity ξ_{om}, local relative water depth h_t/H_s and dimensionless width of structure crest G_c/H_s. The output dimensionless parameter is $q/\sqrt{gH_s^3}$ according to most of the existing formulae.

The models are trained and tested using small-scale

$$\text{if } \frac{R_c}{H_s} \leqslant 1.31, \quad \frac{q}{\sqrt{gH_s^3}} = \exp\left(-2.48\,\frac{R_c}{H_s} - 0.13\,\frac{h_t}{H_s} - 2.58\right); \tag{8}$$

$$\text{if } \frac{R_c}{H_s} > 1.31, \quad \frac{q}{\sqrt{gH_s^3}} = \exp\left(-1.02\,\frac{R_c}{H_s} - 0.075\,\frac{h_t}{H_s} - 5.08\right). \tag{9}$$

In order to overcome the limitation of the M5' model which can only give the linear relationships, we used SPSS

laboratory data selected from the database CLASH. EurOtop (2007) and some of the previous empirical formulae for the estimation of wave overtopping rate have an exponential form. Hence, the exponential form was used for the present model. The M5′ model tree can only produce linear relationships between the input and output parameters. To remove this restriction, the model was developed using $\ln\left(q/\sqrt{gH_s^3}\right)$ as the output parameter.

Using just M5′ model tree, the best results will be Eqs. (8) and (9):

software to make nonlinear regression modeling. Using just this method, the best formula is:

$$\frac{R_c}{H_s} \leqslant 1.31, \quad \frac{q}{\sqrt{gH_s^3}} = \exp\left[-1.84\,\frac{R_c}{H_s} - 4.38\sqrt{S_{om}}\left(\frac{h_t}{H_s}\right)^{2.36} \bigg/ \tan\theta - 3.39\right]. \tag{10}$$

Because in the nonlinear fitting method there is only one limited model which may not entirely cover the range, from

the combination of two above mentioned methods, we can obtain:

$$\frac{R_c}{H_s} \leqslant 1.31, \quad \frac{q}{\sqrt{gH_s^3}} = \exp\left[-2.26\,\frac{R_c}{H_s} - 2.88\sqrt{S_{om}}\left(\frac{h_t}{H_s}\right)^{2.96} \bigg/ \tan\theta - 2.91\right]; \tag{11}$$

$$\frac{R_c}{H_s} > 1.31, \quad \frac{q}{\sqrt{gH_s^3}} = \exp\left[-1.22\,\frac{R_c}{H_s} - 3.38\sqrt{S_{om}}\left(\frac{h_t}{H_s}\right)^{2.43} \bigg/ \tan\theta - 4.72\right]. \tag{12}$$

In fact M5' model uses a special algorithm to divide the domain and presents only linear models. This disadvantage is eliminated by using a combination of M5' model and nonlinear regression of SPSS. The method is illustrated in Fig. 1.
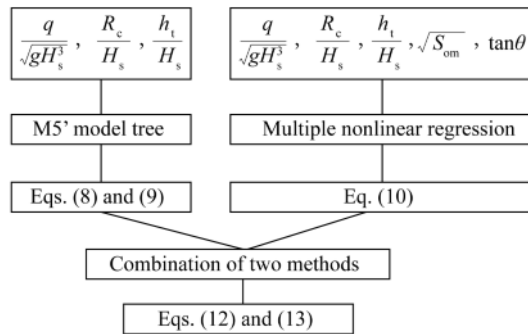

Fig. 1. Parameters and methods.

Data range used to build the model is given in Table 2.

6 Results

Although the slope angle is embedded in the surf similarity parameter (Eq. (7)), the wave runup is separately proportional to the surf similarity parameter and slope angle due to the interaction between individual runup bores. There

often exists significant interaction between subsequent waves, such that when a wave reaches the shoreline and travels up a beach face it cannot always complete a full swash cycle before the next wave comes along. The second wave either overtakes the first wave during its uprush stage (catch up) or collides with the first wave during the backwash stage. This interaction between waves continues with each incoming wave, while following the swash zone hydrodynamics; as a result the maximum runup will not correspond to the uprush of the highest wave in the train. This is particularly true for the mild foreshore slopes, where the time for a swash lens to travel up and down is longer than that for steeper slopes. Thus, the slope angle has a significant effect in predicting the maximum runup height if the swash interaction is accounted for (Bakhtyar et al., 2008).

Table 2 Data range used to build models

Parameter	Definition	Range
R_c (m)	Crest freeboard	0.025–0.400
H_s (m)	Significant wave height at the toe of the structure	0.030–0.262
h_t (m)	Water depth at the toe of the structure	0.050–0.775
T_m (s)	Average wave period	0.632–2.750
$\tan\theta$ (-)	Seabed slope	0.001–0.100
q (m³/s/m)	Mean overtopping discharge per the structure width	1.11×10^{-6}–1.44×10^{-2}

According to Eqs. (11) and (12), reduction in the angle of slope at the toe of the structure, result in the increase of wave overtopping rate. Because when the angle decreases, the reduction of water depth rate decreases and therefore waves are less affected by the roughness of the bed.

On the other hand, when the angle is reduced, if the water depth is smaller than the depth of wave breaking, wave breaks farther from the crest. Thus, a few volumes of water reach behind the wall and small overtopping rate is recorded.

Similarly according to Eqs. (11) and (12), with the increasing depth of structures the overtopping rate is reduced. With the increasing depth of structure, the wave is less affected by the bed and placed under the shoaling phenomenon and its height is not increased. Moreover, the probability of breaking at the vicinity of structure highly decreases and thus water droplets caused by the break are not dispersed in the air and does not appear as the part of the overtopping rate. Instead, standing waves are formed in the toe of the structure, which leads to decreasing of runup and overtopping rate.

Also S_{om} with regard to the relationship $S_0 = H/L_0 = 2\pi H/(gT^2)$, is a function of the significant wave height and incident wave period. By increasing the height of the waves, the waves break at larger depths, thus, the distance between the breaking place and the structure increases and the probability of overtopping becomes less likely. Thus increasing the wave height results in increasing wave steepness and this causes the reduction of overtopping.

In EurOtop (2007) formula, the parameter of ξ_{om} is the surf similarity parameter (Iribarren number) defined as:

$$\xi_{om} = \frac{\tan\alpha}{\sqrt{S_{om}}}. \tag{13}$$

Thus, the wave steepness (S_{om}), is inversely proportional to the surf similarity parameter. With regard to the relationships presented in CEM (2006), the wave runup is directly related to the surf similarity parameter. It is means that the reduction of ξ_{om} reduced runup. Thus, by reducing the runup levels, a few waves can achieve the crest level and overflow. By increasing the wave steepness, ξ_{om} decreases. Reduction of ξ_{om} results in the reduction of overtopping.

Eqs. (11) and (12) have been obtained based on the average of input and output data. In other words, the exceedance level of the measured values from the predicted values is about 50%. Sometimes it is necessary to predict the overtopping rates based on less exceedance level. Therefore, the above equations' coefficients for cases of less exceedance level can be modified. To correct the coefficient, confidence interval is used. The confidence interval of 90% means that the exceedance level of the measured values in this range is 10%. If the line of the best fit is defined as $y=ax$, the line equations of the confidence intervals can be defined as follows:

$$y = (a \pm n\sigma)x, \tag{14}$$

in which a is the fit factor and σ is the standard deviation of the data. n depends on the distribution type of data and the confidence interval. For example, assuming a normal distribution of the data and the confidence interval of 90%, n is equal to 1.65. If we assume that the distribution of data around the line of the best fit is approximately symmetrical, for a confidence interval of 90%, the probability that the measured values exceed the predicted value using $y = (a + n\sigma)x$ is 5%. According to the above mentioned items and considering the standard deviation of the data, Eqs. (11) and (12) are rewritten as follows for the design conditions:

$$\frac{R_c}{H_s} \leqslant 1.31, \quad \frac{q}{\sqrt{gH_s^3}} = \exp\left\{(1-0.28n)\left[-2.26\frac{R_c}{H_s} - 2.88\sqrt{S_{om}}\left(\frac{h_t}{H_s}\right)^{2.96}\middle/ \tan\theta\right] - 2.91\right\}; \tag{15}$$

$$\frac{R_c}{H_s} > 1.31, \quad \frac{q}{\sqrt{gH_s^3}} = \exp\left\{(1-0.39n)\left[-1.22\frac{R_c}{H_s} - 3.38\sqrt{S_{om}}\left(\frac{h_t}{H_s}\right)^{2.43}\middle/ \tan\theta\right] - 4.72\right\}; \tag{16}$$

Values of n for the normal distribution and for various values of confidence intervals and exceedance level are given in Table 3. Putting the appropriate amount of n in Eqs. (15) and (16), we obtain the desired prediction equation based on the exceedance level. The above equations can be used as design relationships for the vertical structures.

For better understanding consider Fig. 2. In Fig. 2, the selected data from CLASH database (for data range $R_c/H_s \leqslant 1.31$) are indicated by blue dots. The predicted values using Eq. (11) are shown in black continuous line. that is the approximately average of data, as seen. However, if we are to obtain the overtopping rate with the exceedance level of 33% for the design consideration, putting $n=0.44$ (from Table 3 for exceedance level of 33%) in Eq. (15), the values shown by black dashed-line is obtained. As seen in

Fig. 2, about 33% of the measured data is larger than the data predicted by the black dashed-line.

7 Discussion

The comparison between the measured dimensionless overtopping discharge and the ones predicted by using the

Table 3 Values of n for a normal distribution for different values of confidence interval exceedance level

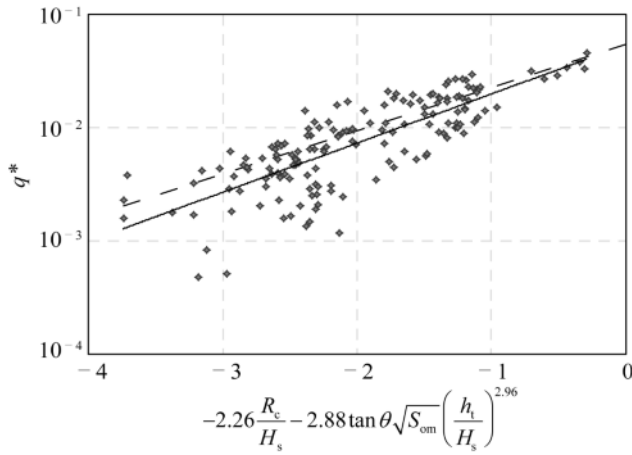| Confidence interval (%)	Exceedance level (%)	n
99.8	0.1	3.09
96.0	2.0	2.05
90.0	5.0	1.65
80.0	10.0	1.28
34.0	33.0	0.44
0.0	50.0	0.00

Fig. 2. Dimensionless measured overtopping discharge (blue dotes) and the predicted ones by using Eq. (11), average value (solid line), and exceedance level of 33% (dashed line).

new formulas (15) and (16) is shown by Fig. 3. The figure indicates that the new formulas yield accurate predictions of the overtopping. Especially for the high overpassing discharges which are more dangerous and may cause damage, the new model is more accurate.
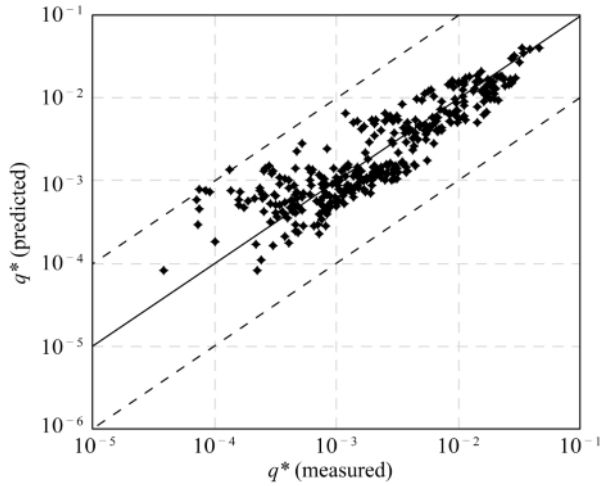

Fig. 3. Scatter diagram of the measured dimensionless overtopping discharge and the ones predicted by using the new formula.

To assess the accuracy of the models, statistical indicators such as the geometric mean (x_G), geometric standard deviation (σ_{x_G})(Goda, 2009), the index of dispersion (*SI*), BIAS, correlation coefficient (R^2), root mean square error (*RMSE*) and discrepancy ratio (*DR*) are used. These indicators are defined as follows:

$$x_G = \exp\left[\frac{1}{n}\sum_{i=1}^{n}\ln\left(\frac{q^*_{est}}{q^*_{meas}}\right)\right] ; \qquad (17)$$

$$\sigma_{x_G} = \exp\left\{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[\left(\ln\frac{q^*_{est}}{q^*_{meas}}\right)^2 - (\ln x_G)^2\right]}\right\} ; \quad (18)$$

$$SI = \frac{\sqrt{\frac{1}{n}\sum(\log q^*_{est} - \log q^*_{meas})^2}}{\overline{\log q^*_{meas}}} ; \qquad (19)$$

$$BIAS = \frac{1}{n}\sum_{i=1}^{n}(\log q^*_{est} - \log q^*_{meas}) ; \qquad (20)$$

$$R^2 = \frac{\left[\sum\left(\log q^*_{est} - \overline{\log q^*_{est}}\right)\left(\log q^*_{meas} - \overline{\log q^*_{meas}}\right)\right]^2}{\sum\left(\log q^*_{est} - \overline{\log q^*_{est}}\right)^2 \sum\left(\log q^*_{meas} - \overline{\log q^*_{meas}}\right)^2} ; \quad (21)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\log q^*_{meas} - \log q^*_{est})^2} ; \qquad (22)$$

$$DR = \frac{q^*_{est}}{q^*_{meas}} , \qquad (23)$$

in which, q^*_{est} and q^*_{meas} are the estimated and measured dimensionless overtopping rates. n is the number of measurements. Table 4 shows the calculated errors for various models.

As it is seen in Table 4, all the error parameters indicate the improvement of the overtopping rate forecasted by the new model.

Table 4   Errors calculated from the values predicted by the models

Statistical indicators	Proposed formula			
	Franco et al. (1994)	EurOtop (2007)	Goda (2009)	Eqs. (11) and (12)
x_G	0.61	1.03	0.78	0.98
σ_{x_G}	2.91	2.29	2.13	1.99
SI	19:17	13:46	12:95	11:16
BIAS	−0.22	0.01	−0.11	−0.01
R^2	0.77	0.25	0.85	0.84
RMSE	0.51	0.36	0.35	0.30
Range of *DR*	0.007–18.2	0.16–93.3	0.03–25.1	0.22–28.8
1/5<*DR*<5(%)	84.02	93.40	92.84	95.32

Fig. 4 shows the changes of DR versus relative crest freeboard (R_c/H_s) for the different models. It's clear that discrepancy ratio of the new model is not as sensitive as other models for changes of R_c/H_s and therefore is more reliable.

To assess the applicability of new model to real cases, the model is evaluated with field data of a vertical wall breakwater in Samphire Hoe, UK (Pullen et al., 2009). Distributions of the measured values and the predicted overtopping rate by different models for this case are indicated in Fig. 5. Compared to other formulas, the new model shows more accurate predictions. Franco's predictions are underestimated and the formulas of EurOtop and Goda give overestimated outputs, especially for low overtopping rates that will result in uneconomical designs.

8 Conclusions

New relationship for the prediction of overtopping rate for vertical walls introduced using combination of multiple
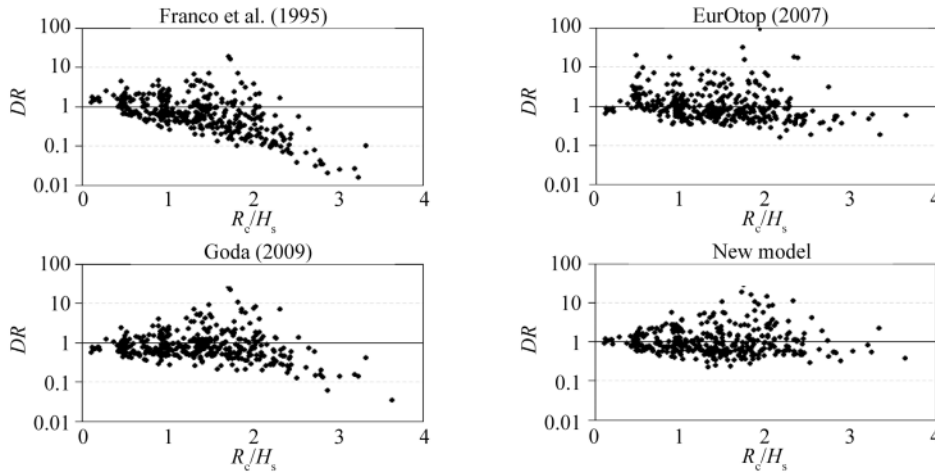
Fig. 4.   shows the changes of DR versus relative crest freeboard (R_c/H_s) for the different models. It is clear that discrepancy ratio of the new model is not as sensitive as other models for changes of R_c/H_s and therefore is more reliable.
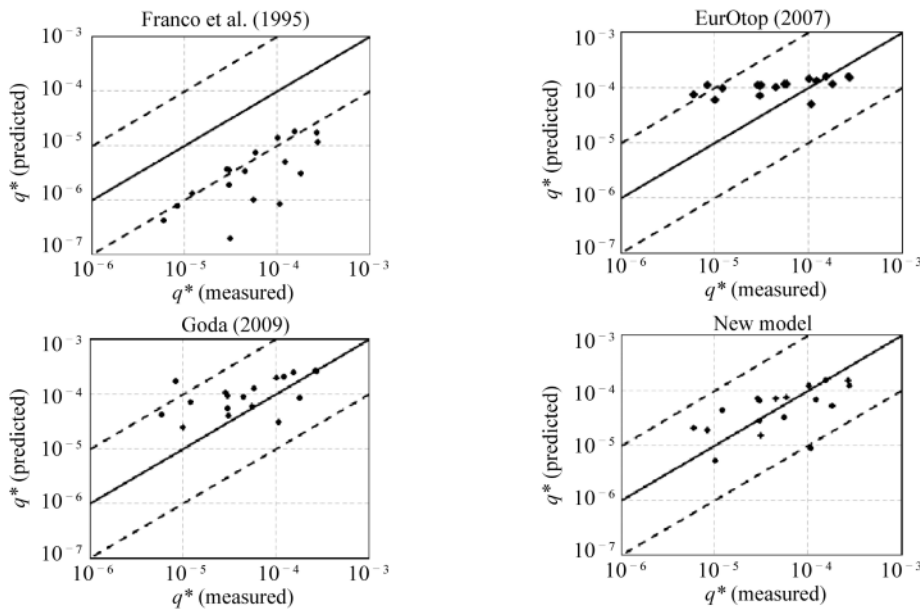

Fig. 5.   Distribution of the measured values and the predicted overtopping rate by different models for field data on the breakwater Samphire Hoe.

nonlinear regression and M5′ model tree. For this purpose, the related data from database CLASH for caisson type structures and other vertical walls were investigated. Data for breakwaters with a smooth surface were used and effective parameters were described. Among several output relationships created using different formations of effective parameters, finally the most accurate and the simplest one was presented with pronouncing nonlinearity of the relative water depth effect on wave overtopping of vertical structures for two classes of the relative freeboard namely 1.31 being the threshold value of the change. This was interpreted for the effect of wave breaking by interaction of reflected waves by incident one and the amount of energy losses caused by overtopped water. The accuracy of the model to experimental data has been compared with that of

other existing models and it shows that the proposed model is more accurate than other existing models. The performances of the models with field measurements are evaluated using statistical parameters and show that the accuracy of these predictions has been improved. The coefficients of the models for the design likely exceeding the lesser of 50% were also modified. As the model provided more accurate results than other existing models and also the simplicity and convenience of use compared with other soft computing tools such as network and nervous, can the case of these relationships be well used.

References

Alikhani, A., Burcharth, H.F., Moadelli, N. and Sotoodeh, B., 2003. Design and construction of a caisson quay wall in the south pars port project, *Asian Journal of Civil Engineering* (*Building and*

Housing), 4(2–4), 163–171.

Bakhtyar, R., A., Yeganeh B., and Ghaheri, A., 2008.Application of neuro-fuzzy approach in prediction of runup in swash zone, *Appl. Ocean Res.*, 30(1), 17–27.

Bhattacharya, B., Price, R.K. and Solomatine, D.P., 2007. Machine learning approach to modeling sediment transport, *Journal of Hydraulic Engineering-ASCE*, 133(4), 440–450.

CEM, 2006. *Coastal Engineering Manual*, US Army Corps of Engineers.

De Rouck, J., Verhaeghe, H. and Geeraerts, J., 2009. Crest level assessment of coastal structures-General overview, *Coast. Eng.*, 56(2), 99–107.

Etemad-Shahidi, A. and Mahjoobi, J., 2009. Comparison between M5' model tree and neural networks for prediction of significant wave height in Lake Superior, *Ocean Eng.*, 36(15–16), 1175–1181.

EurOtop, 2007. *European Manual for the Assessment of Wave Overtopping*, Pullen, T., Allsop, N.W.H., Bruce, T., Kortenhaus, A., Schüttrumpf, H. and van der Meer, J. W. (Eds.), HR Wallingford, Wallingford, U.K.

Franco, L., de Gerloni, M., and van der Meer, J.W., 1994. Wave overtopping on vertical and composite breakwaters, *Proceedings of the 24th International Conference on Coastal Engineering*, Kobe, Japan, Vol. 1, 1030–1045.

Goda, Y., 2009. Derivation of unified wave overtopping formulas for seawalls with smooth, impermeable surfaces based on selected CLASH datasets, *Coast. Eng.*, 56(4), 385–399.

Goda, Y., Kishira, Y. and Kamiyama, Y., 1975. Laboratory investigation on the overtopping rate of seawalls by irregular waves, *Report of the Port and Harbour Research Institute*, 14(4), 3–44. (in Japanese)

Nørgaard, J.H., Burcharth, A.T. and Hans, F., 2014. Distribution of individual overtopping wave volumes in deep and shallow-water wave conditions, *Coast. Eng.*, 83, 15–23.

Pullen, T., Allsop, W., Bruce, T. and Pearson, J., 2009. Field and laboratory measurements of mean overtopping discharges and spatial distributions at vertical seawalls, *Coast. Eng.*, 56(2), 121–140.

Quinlan, J.R., 1986. Induction of decision trees, *Machine Learning*, 1(1), 81–106.

Sakhare, S. and Deo, M.C., 2009. Derivation of wave spectrum using data driven methods, *Mar. Struct.*, 22(3), 594–609.

Van der Meer, J.W. and Bruce, T., 2014. New physical insights and design formulas on wave overtopping at sloping and vertical structures, *J. Waterw. Port Coast. Ocean Eng.*, ASCE, 140(6), 04014025. doi:10.1061/(ASCE)WW.1943-5460.0000221

Van der Meer, J.W., Verhaeghe, H. and Steendam, G.J., 2009. The new wave overtopping database for coastal structures, *Coast. Eng.*, 56(2), 108–120.

Van der Meer, J.W., Verhaeghe, H. and Steendam, G.J., 2005. *Database on Wave Overtopping at Coastal Structures*, CLASH WP2 database, Infram, Marknesse, The Netherlands.

Van Gent, M.R.A., van den Boogaard, H.F.P., Pozueta, B. and Medina, J.R., 2007. Neural network modelling of wave overtopping at coastal structures, *Coast. Eng.*, 54(8), 586–593.

Verhaeghe, H., 2005. *Neural Network Prediction of Wave Overtopping at Coastal Structure*, Ph.D. Thesis, Department of Civil Engineering, Ghent University.

Verhaeghe, H., de Rouck, J. and van der Meer, J.W., 2008. Combined classifier-quantifier model: A 2-phases neural model for prediction of wave overtopping at coastal structures, *Coast. Eng.*, 55(5), 357–374.