ADVANCES IN MODELLING WAVE-STRUCTURE INTERACTION THROUGH ARTIFICIAL NEURAL NETWORKS

Barbara Zanuttigh[1], Sara Mizar Formentin[1] and Jentsje W. van der Meer[2,3]

This contribution describes a new Artificial Neural Network (ANN) able to predict at once the main parameters representative of the wave-structure interaction processes, i.e. the wave overtopping discharge, the wave transmission coefficient and the wave reflection coefficient. The development of this ANN started with the preparation of a new extended and homogeneous database (derived from CLASH database) which collects all the available tests including at least one of the three parameters, for a total amount of 16,165 data. Some of the existing ANNs were compared and improved, leading to the selection of a final ANN, whose architecture was optimized through an in-depth sensitivity analysis to the learning and training ANN features. The new ANN here proposed provides accurate predictions for all the three parameters, resulting in a tool that can be efficiently used for design purposes.

Keywords: artificial neural network, database, wave overtopping, wave reflection, wave transmission

INTRODUCTION

The design of coastal and harbour structures requires a systematic analysis of all the processes of wave-structure interaction, which takes into account the combined effects of wave overtopping, wave transmission and wave reflection. Indeed, all these phenomena should be considered as different outcomes of the same physical process, and therefore should be investigated contemporarily. However, based on the traditional approach most of the existing formulae are targeted to represent just one process and are fitted on specific (more or less wide) databases, addressing usually one or few structure types and having therefore a specific (more or less narrow) validity field.

The development and/or use of an Artificial Neural Network (ANN) is therefore particularly recommended in case of complicated structure geometries and variable wave conditions. This kind of predictive method requires however a homogeneous and "wide-enough" database: the number of data should be sufficient for training the ANN based on a number of ANN parameters and including a sufficiently wide number of data for all range of possible output values. There are specific cases that also an ANN cannot deal with, such as very complex walls, perforated caissons and double promenades, see for details EurOtop (2007) and specifically the methodology released within the PC-OVERTOPPING calculator (http://www.overtopping-manual.com/calculation_tool.html).

One of the most successful ANNs is the method available from CLASH (2004) and EurOtop (2007) for wave overtopping, which was specifically developed to predict the overtopping discharge for a wide range of coastal structures, including complex geometries (Van Gent et al., 2007).

After and during CLASH (2004) other ANNs have been developed. Specific ANNs were developed for the estimation of the wave overtopping discharge (Verhaeghe et al., 2008), of the wave transmission coefficient (Van Oosten and Peixó Marco, 2005; Panizzo and Briganti, 2007) and of the wave reflection coefficient (Zanuttigh et al., 2013).

Each of these ANNs actually proved to be able to overcome some of the limits imposed by the traditional empirical formulae, but they are still restricted to reproduce only one of the processes involved in the wave-structure interaction. Nevertheless, the assumption that all the processes are physically correlated implies that a unique set of physically-based parameters can be defined to represent all the phenomena.

Therefore, purpose of this work is to further develop the existing ANNs in order to deliver a tool which is able to estimate the wave overtopping discharge (q), the wave transmission and the wave reflection coefficients (K_r and K_t) at once, i.e. by means of just one set of input parameters and of the same ANN architecture.

Since the basis of a "good" ANN is essentially the database to be used for training, the first step is the arrangement of a "wide-enough" and homogeneous collection of tests, preferably organized following the same structure schematisation. Indeed, the data, the set-up and the structure parameters of CLASH (2004) have been used as a starting point to gather the other tests and organize them according to the same approach. The resulting final database is described in details in the first following section.

The second step consists of the choice of the best ANN layout, both considering the input parameters and the internal architecture. To achieve this goal, two of the existing ANNs have been analysed, tested and modified, specifically the original CLASH ANN (Van Gent et al., 2007) and the ANN for reflection proposed by Zanuttigh et al. (2013). Once selected the best input set, a sensitivity

[1] DICAM, Università di Bologna, Viale Risorgimento 2, Bologna, 40136, Italy
[2] Van der Meer Consulting bv, P.O. Box 11, Akkrum, 8490 AA, The Netherlands
[3] UNESCO IHE, Westvest 7, Delft, 2611 AX, The Netherlands

analysis to the parameters characterizing the ANN architecture, e.g., the number of the hidden neurons and the methods for improving generalization, is required. The process of selection and optimization of the input parameters is then reported.

A separate section illustrates then the ANN architecture, including a synthesis of the methodology followed to define the most relevant elements of the architecture itself. Since the existing ANNs propose different methods for the performance assessment, the employment of different techniques and its fallout on the valuation of the performance is discussed. Furthermore, the adoption of the bootstrap resampling technique and the effective optimal number of required resamples is described.

The performance of the ANN is then discussed, referring to all the output parameters, q, K_r and K_t: the outcomes of the ANN are analyzed for each separate process, by means of quantitative indexes of performance. The specific prediction of q is furthermore qualitatively investigated through plots and diagrams which report the comparison among experimental and predicted values and the distribution of the error computed by the ANN.

Conclusions and steps for further research are finally drawn.

THE DATABASE

A homogenous database of 16,165 tests has been gathered, starting from the original CLASH database (Van der Meer et al., 2008) and extending it in order to include other wave overtopping, reflection and transmission tests. The assemblage of the data has been carried out by keeping the set-up of the original database, i.e. by following the same schematization of the structures (see Fig. 1) and by maintaining the same geometric parameters, as well as the relevant climate parameters, already identified within CLASH project.

In addition with respect to the original CLASH database, the values of K_r and K_t have been included where available. Moreover, an additional geometric parameter has been introduced: the average unit size D representative of the structure elements, which has proved to be relevant in the prediction of K_r and K_t (Panizzo and Briganti, 2007; Formentin and Zanuttigh, 2013; Zanuttigh et al., 2013). Therefore, the complete database, in its final layout, consists of 13 hydraulic parameters, 18 structural parameters and 3 general parameters (the reliability and complexity factors and the identify name of the test). The most relevant parameters are sketched in Figure 1; for a more detailed description, see Van der Meer et al. (2009).

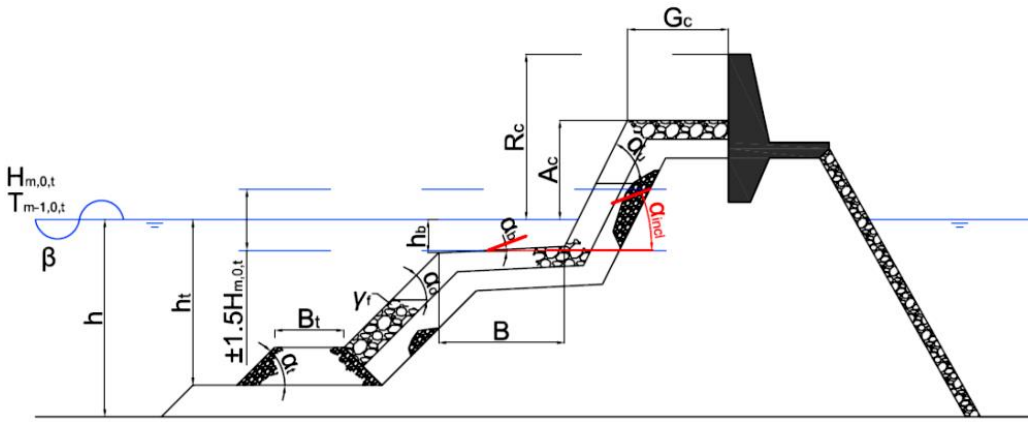

Figure 1. Structure schematization including the hydraulic and structural parameters, based on CLASH.

The complete database is organized into 7 "sections", labeled progressively from A to G, in order to distinguish the different type of structures and wave attack conditions: rock permeable straight slopes (group "A"), rock impermeable straight slopes (group "B"), armour units straight slopes ("C"), smooth and straight slopes ("D"), structures with combined slopes and berms ("E"), seawalls ("F") and oblique wave attacks ("G"). Following this partition, the database can be analyzed considering the available data for each of the specific processes (see Fig. 2):

- for q, a total amount of 11,825 tests are available; these data have been collected from: the CLASH database (Van der Meer et al., 2009), which consists of more than 10,000 irregular wave overtopping tests and includes dikes, rubble mound breakwaters, berm breakwaters, caisson

structures and combinations with complicated geometries; reports from LWI (Oumeraci et al., 2001, 2004 and 2007); additional data on smooth steep slopes (Victor, 2012).

- for K_r, 7,413 data are available, mainly derived from the extended wave reflection database (Zanuttigh et al., 2013), which collects more than 5,700 data, including the original wave reflection database based on CLASH database (Zanuttigh and Van der Meer, 2008) and additional data on seawalls (Oumeraci et al., 2001, 2004 and 2007), steep slopes (Victor 2012) and berm breakwaters (Lykke Andersen, 2006);
- for K_t, additional data structures with berms (Lissev, 1993) have been combined with the DELOS database on low crested breakwaters (Van der Meer et al., 2005; Panizzo and Briganti 2007), which consists of more than 3,000 data, including rubble mound structures, aquareefs, rubble mounds with different kind of armour units and smooth slopes, for a total amount of 3,366 tests.

The diagrams in Figure 2 represent the distribution of the data included within each of the three datasets, in order to illustrate the different kind of structures and wave conditions available for each output parameter. By comparing the pie-charts, it is evident that some typology of tests – e.g., the structures with berms and seawalls – are completely absent from the transmission dataset, while the reflection dataset, as well as the overtopping dataset, includes all the different kind of tests.

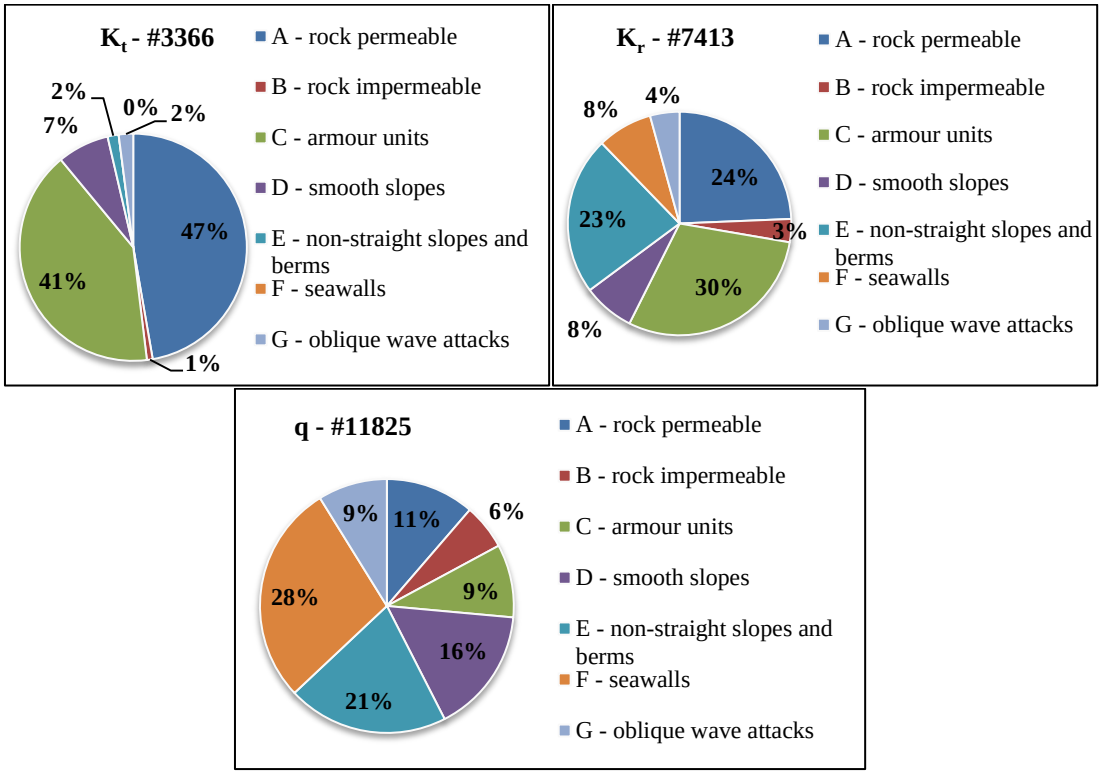

Figure 2. Pie charts representing the distribution of the data within, respectively - from left to right and from top to bottom - the wave transmission, the wave reflection and the wave overtopping datasets.

The three datasets are only partially overlapped, i.e. only sub-datasets comprehend more than one of the outputs parameters, and specifically:

- the value of q and K_r is contemporarily available for 1,996 tests, which represent approximately the 12% of the total data, the 27% of the reflection tests and the 17% of the overtopping tests;
- the value of K_r and K_t is contemporarily available for 2,303 tests, i.e. approximately the 14% of the total data, the 31% of the reflection tests and the 68% of the transmission tests;
- the percentage of tests for which both q and K_t are known is insignificant (less than 1%), and therefore also the percentage of tests for which all the output parameters are available is too limited.

The performance of prospective combined prediction of the output parameters (i.e. K_r and K_t; K_r and q) is expected to be influenced by the significantly narrower and less assorted "common databases" to train the ANN on. Depending on the possibility to gather more data on combined measurements, the development of contemporary predictions of more than one output parameter is therefore postponed to future research.

THE INPUT PARAMETERS

In order to define the most suitable set of input parameters and the best architecture for the ANN, two of the existing ANNs have been analysed, compared and modified:

- the ANN for wave overtopping developed within the CLASH project by Van Gent (2007), referred in the following as ANN (1);
- the ANN for reflection proposed by Zanuttigh et al. (2013), named ANN (2).

The ANN (1) was optimized for the evaluation of q and it was originally trained against the original CLASH database (Van der Meer et al., 2009). The experimental data were scaled according to the Froude Law by dividing each quantity with the significant wave height, $H_{m0,t}$,. Therefore the data were basically rescaled to a same prototype condition ($H_{m0,t}$, = 1 m). It included 14 "dimensional" input parameters, describing both geometric and climate characteristics. The output parameter q was also scaled and transformed in logarithmic scale, before applying the training process.

The ANN (2) was developed for the prediction of K_r. It was trained against an extended version of the CLASH database, which included additional datasets on impermeable slopes, seawalls and 3D wave attack. The input vector comprehends 13 non-dimensional elements (Zanuttigh et al., 2013), reproducing a specific physical parameter or process, for example: $H_{m0,t}/L_{m-1,0,t}$ represents the wave steepness, $h_t/L_{m-1,0,t}$ the shoaling index, etc.. To this purpose, this ANN (2) uses two scaling parameters, i.e. a depth and a width, respectively $H_{m0,t}$ and $L_{m-1,0,t}$, instead of the Froude scaling as ANN (1). Due to the different scaling, the values in the extended database are processed in model scale.

Besides, since D is a key parameter for the wave reflection and wave transmission, a third ANN, concealed as a variation of ANN (1), has been introduced: this ANN, named ANN (1b), is absolutely identical to ANN (1), with the exception of the presence of D within the input vector, whose size consequently rises up to 15 parameters.

The input parameters of the three ANNs are displayed in Table 1, where the CLASH symbols and terminology (see Fig. 1) have been adopted. Table 1 also reports a last line which summarizes the minimum required number of tests to train each ANN; this value strictly depends on the number of input parameters, and, in particular, it exactly corresponds to the total number of connections among input parameters, hidden and output neurons and biases (see Fig. 2 and the following Section). Therefore the minimum required of tests, N_{tests}, can be easily determined as

$$N_{tests}=HN\cdot(I+2)+ON\cdot HN+b\cdot HN+(b+1)\cdot ON, \quad (1)$$

where: HN = number of hidden neurons; I = number of input parameters; ON = number of output neurons; b = bias.

Table 1. Input parameters of the three tested ANNs. ANN (1) is the original CLASH ANN; ANN (1b) is the variation of CLASH ANN including *D*; ANN (2) is the ANN developed by Zanuttigh et al. (2013). The total number of 15 parameters for ANN (2) includes the wave period $T_{m-1,0,t}$ and the water depth h in front of the structure implied in the calculation of the wave length $L_{m-1,0,t}$.

#	ANN (1)	ANN (1b)	ANN (2)
1	$H_{m,0,t}$	$H_{m,0,t}$	$H_{m,0,t}/L_{m-1,0,t}$
2	$T_{m-1,t}$	$T_{m-1,t}$	$h_t/L_{m-1,0,t}$
3	γ_f	γ_f	$B_t/L_{m-1,0,t}$
4	$cot\alpha_d$	$cot\alpha_d$	γ_f
5	$cot\alpha_u$	$cot\alpha_u$	$cot\alpha_d$
6	B	B	$cot\alpha_{incl}$
7	B_t	B_t	$D/H_{m,0,t}$
8	h	h	$R_c/H_{m,0,t}$
9	h_t	h_t	$B/L_{m-1,0,t}$
10	h_b	h_b	$h_b/H_{m,0,t}$
11	R_c	R_c	$G_c/L_{m-1,0,t}$
12	A_c	A_c	m
13	G_c	G_c	β [rad]
14	$tan\alpha_B$	$tan\alpha_B$	
15	β	β [rad]	
16		D	
TOT	15	16	15
Data required to train the ANN	**711**	**753**	**711**

Consequently, the first issue to be taken into account when comparing the performance of different ANNs is that the larger the input vector size, the larger the required number of training tests. Since the number of input parameters of ANN (1b) is one-element larger (16 instead of 15), this network requires more training tests. Besides, the different ANNs have to be tested separately on the prediction of each output parameter, comparing the numerical values to the experimental ones.

In this work, three error indexes have been employed: the root mean square error, *rmse*, the Willmott index, *WI* (Willmott, 1981*)* and the coefficient of determination, R^2, respectively defined in Eqs. 2, 3 and 4. The quantity "*X*" (specified in Eq. 5), represents the output parameter, while the subscript indexes "s" and "ANN" respectively refer to the experimental and the predicted quantity.

$$rmse = \frac{1}{500}\sum_{i=1}^{500}\left(\sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(X_{s_j}-X_{ANN,_j}\right)^2}\right) \tag{2}$$

$$\mathrm{WI} = \frac{1}{500}\sum_{i=1}^{500}\left(1-\frac{\sum_{j=1}^{N}\left(X_{s,j}-X_{ANN,j}\right)^2}{\sum_{j=1}^{N}\left[\left|X_{s,j}-\bar{X}_s\right|+\left|X_{ANN,j}-\bar{X}_s\right|\right]^2}\right) \tag{3}$$

$$R^2 = \frac{1}{500}\sum_{i=1}^{500}\left(1-\frac{\sum_{j=1}^{N}\left(X_{s,j}-X_{ANN,j}\right)^2}{\sum_{j=1}^{N}\left(X_{s,j}-\bar{X}_s\right)^2}\right) \tag{4}$$

$$\begin{cases} X = \begin{matrix} log\, q' = log\, q / \sqrt{g \cdot H_{m,0,t}^3} \\ K_r \\ K_t \end{matrix} & and \quad \bar{X}_s = \frac{1}{N}\sum_{j=1}^{N} X_{s,j} \end{cases} \tag{5}$$

The numerical values of the error indexes are reported in Table 2; these values correspond to the average results obtained from 50 bootstrap resamples of the database, and the uncertainty associated to each index is the standard deviation. Table 2 is organized into three parts according to the different applications.

Table 2. Comparison among the quantitative performance of the three ANNs considered and tested. ANN (1) is the original CLASH ANN; ANN (1b) is a variation to CLASH ANN, including the extra input parameter *D*; ANN (2) is the ANN developed by Zanuttigh et al. (2013). Average results from 50 bootstrap resamples of the database.

Overtopping discharge, *q*			
Error indexes	*rmse*	*WI*	R^2
ANN (1)	0.32 ± 0.01	0.971 ± 0.003	0.894 ± 0.008
ANN (1b)	0.30 ± 0.02	0.976 ± 0.004	0.91 ± 0.01
ANN (2)	0.28 ± 0.02	0.976 ± 0.005	0.91 ± 0.01
Wave reflection coefficient K_r			
Error indexes	*RMSE*	*WI*	R^2
ANN (1)	0.046 ± 0.003	0.985 ± 0.002	0.943 ± 0.008
ANN (1b)	0.044 ± 0.002	0.987 ± 0.002	0.949 ± 0.007
ANN (2)	0.035 ± 0.004	0.991 ± 0.003	0.966 ± 0.009
Wave transmission coefficient – K_t			
Error indexes	*RMSE*	*WI*	R^2
ANN (1)	0.035 ± 0.002	0.9943 ± 0.0006	0.978 ± 0.002
ANN (1b)	0.033 ± 0.001	0.9951 ± 0.0004	0.981 ± 0.002
ANN (2)	0.027 ± 0.002	0.9968 ± 0.0006	0.987± 0.002

The adoption of different indexes is introduced to take into account different aspects of the ANN performance and to compare the results of the different applications. The *rmse* shows the dispersion of the predicted values with respect to the measured ones, therefore the lower the *rmse*, the more accurate the prediction. Its numerical value depends on the order of magnitude of the associated output parameter. In fact, the *rmse* values associated to *q* result at least 2-orders of magnitude larger than the *rmse* for K_r and K_t (see Tab. 2), because the values of *q* are processed in logarithmic scale (see Eq. 5).

WI and R^2 are instead normalized indexes and therefore range between 0 and 1, being 1 the perfect correspondence. R^2 accounts for the distribution of the experimental values around the mean, while *WI* accounts also for the distribution of the prediction with respect to the same experimental mean. *WI* is

thus a symmetry indicator. The logarithmic transformation does not affect the normalized values of *WI* and R^2.

By comparing the results of the ANNs (Tab. 2) it is pretty evident that ANN (1) and ANN (2) perform best when predicting, respectively q and K_r, consistently with the fact that they have been originally optimized to predict these processes. ANN (1b) performs moderately but methodically better than ANN (1), confirming the relevance of the parameter D, not only for the wave reflection and transmission, but also for wave overtopping. ANN (2) yields a significant reduction of the error in terms of error indexes with respect to ANN (1) and ANN (1b). Therefore, number of input elements being equal or lower, ANN (2) returns the best outcomes for all the application tests and this can be explained by the physically based scaling (with wave height and wave length). The input elements of ANN (2) have been selected in the following for the prediction of all the output parameters.

THE ANN ARCHITECTURE

The realization of an ANN model does not only comprehend the definition of the input parameters by means of analysis of sensitivity. In fact, the "*black-box*" nature of this kind of mathematical models requires anyway a calibration process oriented to the choice and the optimization of several characteristic elements.

An ANN model essentially consists of layers (see Fig. 2). The input elements are "connected" to one or more hidden layers, each of them being composed by an undefined number of hidden neurons. The numerical information contained in the input is elaborated by the ANN and passed to the first (or unique) hidden layer through the so-called "hidden layer transfer function", a proper mathematic function which transforms the numerical values of the input neurons. Similarly, the following hidden layers - if existing - receive the information from the hidden layer through other transfer functions, up to the output layer (in this case, we have the "output neuron transfer function").

Each input element is connected to each hidden neuron, and the hidden neurons are likewise connected to the output. The process of training the ANN essentially consists in the definition of the numerical weights to be assigned to each connection: this process, governed by the training algorithm, is based on the iterative research of the minimum error among the outputs predicted by the ANN itself and the experimental target values. The velocity and the efficiency of the training depend on the transfer functions, the error type to minimize, the tolerance imposed and on the training algorithm.

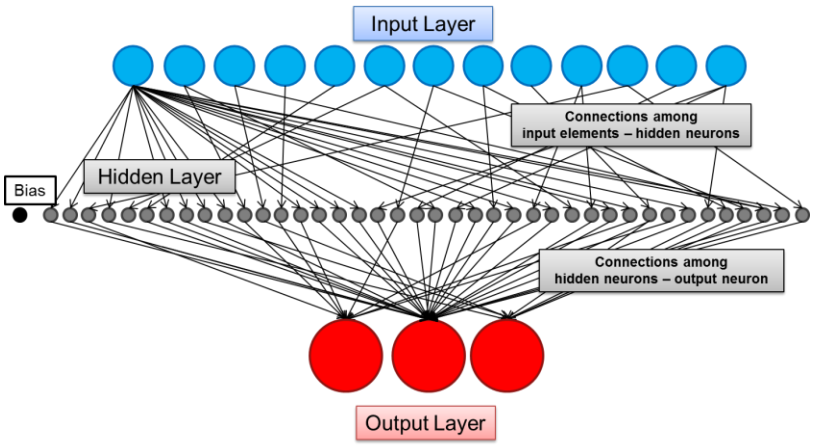

Figure 2. Schematization of the conceptual layout of an ANN, organized in layers. The number of the input elements (13), the hidden neurons (40, and 1 bias) and the output neurons (3) reflects the architecture of the network definitely chosen in this work, i.e. ANN (2).

The initial ANN architecture has been directly derived from ANN (2), based on the analysis presented already by Zanuttigh et al. (2013) and Formentin and Zanuttigh (2013), and then tested against the prediction of all the three processes. The resulting optimal characteristics of the ANN architecture are resumed in the following:

- multilayer network, based on a "feed-forward back-propagation" learning algorithm; 1 hidden layer, and 1 output layer, corresponding either to K_r, K_t or q.
- the hidden layer comprehends 40 hidden neurons; this number has been re-defined after a specific sensitivity analysis (see the reserved subsection later in this paper);

- training algorithm: *Levenberg – Marquardt* (Levenberg, 1994; Marquardt, 1963);
- hidden neurons transfer function: hyperbolic tangent sigmoid function;
- output neuron transfer function: linear transfer function.
- error type: *mse* (mean squared error);
- tolerance on the error (*goal*): $\sim 10^{-4}$ (K_r, K_t), $\sim 10^{-7}$ (log q);
- maximum number of iterations (*epochs*) allowed: 100;
- method to improve generalization: none, after testing (and deciding to discard) the "*early stopping*" method. The assessment of the ANN performance is attributed to the *bootstrap resampling* technique.

Although each of these parameters have been defined after calibration, only the most significant analysis are reported within this paper. In particular, the first subsection in the following, describes in detail the optimization of the number of hidden neurons, while the second one focuses on the different techniques to improve the capability of generalization of the ANN.

The bootstrap resampling technique

The bootstrap technique consists in several (N) resampling with replacement of the data to be selected for the training of the ANN. For each run of the ANN, one of the N bootstrapped databases is used for the training. The size of each bootstrapped database equals the original one, but the included data are differently assorted, since each selected test is randomly selected with replacement. In this work, the random selection is driven by the "Weight Factor" WF associated to each test and based on the Reliability and Complexity Factors, RF and CF, as already done within the CLASH project (Van Gent, 2007): WF=(4-RF)·(4-CF). The higher the WF, the higher the probability for a test to be selected. Each bootstrapped database may therefore include the same tests more than once, while some tests may never appear.

The bootstrap resampling of the database is principally adopted to assess the performance of the ANN. Each differently-trained ANN yields to differently evaluated output parameters, and the ensemble of the predicted outputs can be considered as a stochastic variable and therefore used to derive average indexes of performance and standard deviations.

Furthermore, if the number of resamples is large-enough to be statistically significant, it is possible to calculate the quantiles of the distribution and derive the confidence intervals. A paramount aspect is that a mean prediction is not only more significant from a statistical view-point, but is also more accurate since it adopt the commitment of several randomly-trained ANNs.

Tests have been carried out by means of different numbers of bootstrap resamples, ranging from 50 (to speed up the sensitivity analysis to other parameters) up to 500, based on the suggestions from previous works (Van Gent et al., 2007; Verhaeghe et al., 2008). The results of these tests did not show any significant dependence on the increase of the number of resamples above 50, proving that essentially the difference is made by the use or not of this technique. This can be justified because of the very large size of the original database.

The number of hidden neurons

The definition of the number of hidden neurons – which represents one of the key-features of an ANN – is generally related to the number of input parameters and to the range of variability of the input data, but it cannot be defined *a priori*.

The common methodology (Van Gent, 2007; Panizzo and Briganti, 2007; Verhaeghe, 2005) to establish the optimal number of hidden neurons by testing the performance of the ANN as a function of the progressive increase of the number of the hidden neurons has been here adopted. Since the ANN is supposed to predict K_r, K_t and q, all the applications have been taken into account, and the optimal number has been finally established as the most suitable for all the outputs.

The results of the sensitivity analysis are graphically reported in Figure 3 in terms of average *rmse* values and standard deviation. These values were derived from the resulting prediction of the ANN after 50 bootstrap resampling of the training database. The computation of average errors allows to detect (Fig. 3) how the increase of the number of hidden neurons may induce not only a reduction of the *rmse* but also a decrease of the error band, i.e. of the standard deviation. The analysis of this latter aspect, which was not considered by previous works, shows that the ANN tends to systematically perform better and therefore to be less affected by the random selection of data for training by increasing the size of the hidden layer. An important consequence is the reduction of the uncertainty associated to the prediction, which corresponds to narrower intervals of confidence for each output estimation. The trend of the number of hidden neurons in predicting q (Fig. 3) would suggest to select 50 hidden neurons, both considering the *rmse* value and its standard deviation. However, the

predictions of K_t and K_r show that the number of 50 hidden neurons would not represent a good compromise among the complexity of the architecture and the performance obtained, since the reduction of *rmse* is negligible (K_t) or even opposite (K_r) when the number of hidden neurons is greater than 40.

The complexity of the architecture is indeed a paramount issue, since from one hand it characterizes the minimum number of tests to training the ANN (see Tab. 1), and from the other hand it may induce overtraining problems, see the following subsection regarding the improvement of the generalization. Taking into account all these correlated aspects, a final number of 40 hidden neurons was selected.

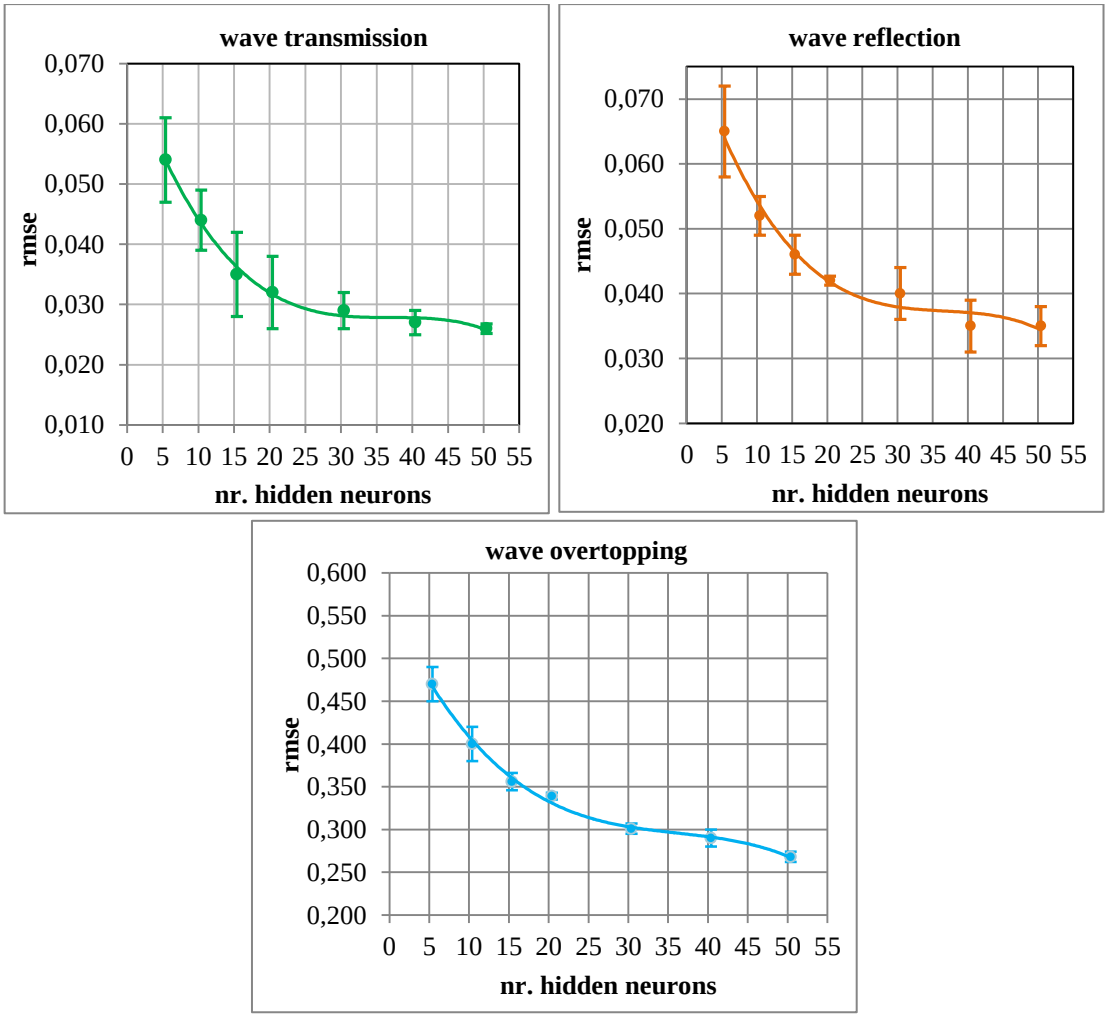

Figure 3. Mean *rmse* values (ordinate) and corresponding standard deviations as functions of the number of hidden neurons (abscissa); from left to right and from top to bottom: wave transmission, reflection and overtopping. Results averaged after 50 bootstrap resamples.

Improving generalization: the early stopping technique

One of the most important issues correlated to the actual performance of an ANN is represented by its capability of generalization, i.e. of overcoming the limits of the range of training tests. An ANN is said to be "over-trained" when it is able to reproduce very well its training data, but is not able to predict with sufficient accuracy beyond the training ranges. This generally occurs when a "too-large" number of hidden neurons has been applied and the architecture of the ANN is "too much focused" on replicate the trend of the training tests and does not learn the "general rule".

Since a pretty-large number of hidden neurons is anyway requested to obtain satisfactory results (see Fig. 3), especially if a wide database is employed, several techniques to improve generalization have been developed. The most common technique is the so-called "early-stopping": this methodology essentially consists in splitting the overall database into three datasets, training the ANN only over one

of them (the proper "training set") and using the remaining sets to stop if needed the training process before the achievement of the expected performance. The stopping is imposed when for several iterations consecutively the *rmse* on the validating set is not reduced, even if the training *rmse* continues to decay. In other terms, the "early-stopping" technique interrupts the training process before the "optimum", leading to a slightly worse performance, but ensuring a greater capability of generalization.

Therefore the validating set is employed to test the ANN capability of generalization, i.e. the capability to predict values not belonging to the training set, while the testing set is employed to verify the performance of the ANN at each iteration without affecting the training process.

The adoption of the bootstrap resampling of the database allows to assess the capability of generalization of an ANN avoiding the employment of specific methodologies, such as the early stopping (Verhaeghe, 2005). In fact, an ANN trained many times, each time on a randomly different (bootstrapped) database, produces "average" predictions and relative indexes of performance, such as the standard deviation or intervals of confidence (see the related section below).

The effects of the implementation of the "early-stopping" have been considered, through a specific sensitivity analysis. Table 3 synthetizes the performance of the ANN with respect to all the applications (q, K_r and K_t), comparing the results obtained with the "early-stopping" and without it. Similarly to the sensitivity analysis to the input parameters (see Tab. 2), the error indexes *rmse*, *WI* and R^2 have been calculated. The numerical values have been derived as average results from 50 different random splitting of the database into training-testing-validating sets (in case of early-stopping) and from 50 bootstrap resampling of the database.

The introduction of the "early-stopping" causes a sharp decrease of the ANN performance, as it can be appreciated by the increase of the *rmse*, while *WI* and R^2 decrease (Table 3). Therefore, the choice to exclude the early stopping and to assess the uncertainty of the predictions with the bootstrapping technique is straight-forward.

Table 3. Synthesis of the performance of the ANN obtained with and without the "early-stopping" technique. Average results from 50 bootstrap resamples of the database (case without early stopping) and 50 different random splitting of the database into training-testing-validating sets (case with early stopping).

Overtopping discharge, q			
Early stopping	*rmse*	*WI*	R^2
No	0.28 ± 0.01	0.976 ± 0.005	0.91 ± 0.01
√	0.33 ± 0.04	0.969 ± 0.008	0.89 ± 0.03
Wave reflection coefficient K_r			
Early stopping	*rmse*	*WI*	R^2
No	0.035 ± 0.004	0.991 ± 0.003	0.966 ± 0.009
√	0.046 ± 0.009	0.990 ± 0.007	0.95 ± 0.03
Wave transmission coefficient K_t			
Early stopping	*rmse*	*WI*	R^2
No	0.027 ± 0.002	0.9996 ± 0.0006	0.987 ± 0.002
√	0.033 ± 0.009	0.995 ± 0.005	0.98 ± 0.02

THE PERFORMANCE OF THE ANN

The final layout of the ANN has defined based on the result of the sensitivity analysis to the input parameters, the number of hidden neurons and the other characteristic features of the ANN (see Fig. 2 and the previous section). This section presents the main results obtained by the application of the "optimized" ANN to the prediction of q, K_r and K_t. The performance of the ANN has been investigated by means of:

- numerical error indexes, i.e. the *rmse*, the *WI* and the R^2;
- for the specific case of q, comparison among predicted and experimental values and study of the error distribution.

The comparison with existing ANNs and the assessment of the reliability of the predictions has been already presented previously, see Tables 2 and 3 respectively.

Average predictions

The quantitative results of the predictions of q, K_r and K_t are summarized in Table 4 in terms of the average values derived after 500 bootstrap resamples of the ANN. The choice to develop 500 resamples of the database is related to the assessment of the performance; this aspect has been described in the subsection dedicated to the bootstrapping technique.

From Table 4 it can be observed that the *rmse* value associated to the prediction of q presents the same order of magnitude of the *rmse* for K_r and K_t. In fact, in this case, the experimental values of q have been scaled in a different way with respect to Eq. 5.

$$q^* = \frac{log_{10}(q_{AD}) - min\{log_{10}(q_{AD})\}}{\left| min\{log_{10}(q_{AD})\} - max\{log_{10}(q_{AD})\}\right|}, \quad (6)$$

The target output for the ANN is the quantity q^* computed as in Eq. 6. The experimental values of q are previously non-dimensionalized to the quantity q_{AD}, by means of the factor $(g \cdot H_{m0,t}^{\ 3})^{1/2}$

$$q_{AD} = \frac{q}{\sqrt{g \cdot H_{m0,t}^3}} \quad (7)$$

q_{AD} is then transformed into logarithmic scale and finally q^* is obtained by scaling q_{AD} with respect to its maximum and minimum and translating it to the range [0; 1] (see eq. 6).

Actually, the ultimate aim of the transformation from q to q^* is to obtain target values of q between 0 and 1, similarly to the natural values of K_r and K_t. This objective is pursued to uniform the results and ease the comparison of the performance among the three application cases. Besides, the employment of q^* has revealed to allow a slight improvement of the performance, which can be detected by comparing the values of *WI* and R^2 between Table 2 (referring to the case of ANN (2), prediction of q) and Table 4. The *rmse* is not comparable being affected by the different scale of the outputs.

The increase of the error number of bootstrap resamples from 50 (Tab. 2) to 500 (Tab. 4) does not significantly affect the error indexes associated to K_r and K_t.

Table 4. Synthesis of the performance of the ANN in its final layout (13 non-dimensional input parameters, 40 hidden neurons, no early stopping). Average results from 500 bootstrap resamples of the database.

Output	*rmse*	*WI*	R^2
q	0.046 ± 0.008	0.98 ± 0.01	0.92 ± 0.05
K_r	0.036 ± 0.008	0.99 ± 0.01	0.96 ± 0.04
K_t	0.028 ± 0.005	0.996 ± 0.009	0.98 ± 0.04

It is particularly important to remark that the final outcome of the ANN is an "average" prediction: this means that the ultimate tool to be delivered to the users is represented by 500 ANNs, each of them trained on a different database. Each ANN is supposed to separately process the input parameters defined by the user and predict the output; therefore, the 500 output estimations are averaged and a mean result is delivered, together with the confidence interval associated to the prediction. As regards the specific prediction of q, it has to be clarified that the transformation applied to the target values before training the ANNs (Eqs. 6 and 7) is completely undone before supplying the results to the user.

Analysis of the error distributions

Since the optimized ANN definitely corresponds in large part to the ANN (2) proposed by Zanuttigh et al. (2013) for the wave reflection and since the same ANN was already applied to the prediction of the wave transmission (Formentin and Zanuttigh, 2013), hereafter the results relating to the prediction of q are mainly reported and discussed.

It has to be noticed in advance that, following the work of Van Gent et al. (2007), all the experimental values of q $<10^{-6}$ $m^3/(s \cdot m)$ were discarded and associated to "zero-overtopping". With reference to the pie-chart of Figure 4 – which displays the distribution of q values, divided into different classes according to the order of magnitude – the loss of tests with overtopping $< 10^{-6}$ $m^3/(s \cdot m)$ corresponds to 22% of the total amount of data. Therefore, in order to train the ANN on the complete database, the implementation of a classifier ANN able to distinguish among "zero-overtopping" and "non- zero-overtopping", similarly to the studies of Verhaeghe et al. (2008) for the CLASH ANN, can be subject for future research.

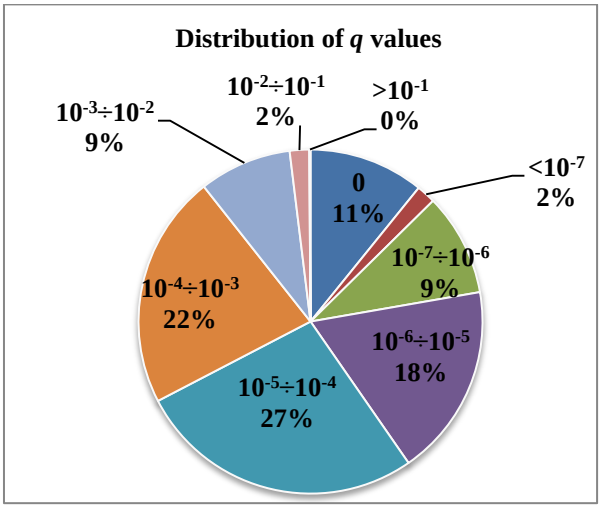

Figure 4. Pie charts representing the distribution of the experimental values of q, divided into different classes according to the order of magnitude.

An overall presentation of the ANN performance is provided in Figure 5 through the qualitative comparison among the experimental (q_s) and the predicted (q_{ANN}) values of q: in the diagram the bisector line represents the ideal condition $q_{ANN} = q_s$, while the external lines are the 95% confidence bands. The scatter is satisfactorily limited and most of the predictions fall within the 95% confidence bands. However, Figure 5 shows that the distribution of the points is not completely symmetric, as the ANN tends to systematically overestimate the low values of q ($q < 10^{-5}$ m^3/(s·m)). This aspect is enlightened by Figure 6, which represents the distribution of the relative error $(q_s\text{-}q_{ANN})/q_s$ as a function of the experimental values q_s. From this diagram, it is evident that the best ANN predictions are obtained for values of $q_s \approx [10^{-5}; 10^{-3}]$ m^3/(s·m), where the scatter is reduced and the symmetry higher.

The principal reason for the asymmetry of the error distribution may be attributed to the non-uniform distribution of the values of q_s within the complete interval $[10^{-6}; 1]$ m^3/(s·m). Most of the values of q_s (49%) actually falls in the range $[10^{-5} ; 10^{-3}]$, while only the 11% of q_s is greater than 10^{-3} and 18% is lower than 10^{-5} (see Fig. 4). The "bias" of the experimental database is expected to be reflected into a bias of the prediction, i.e. to be the cause of the tendency to overestimate the low values of q_s.

Since the relative crest freeboard $R_c/H_{m0,t}$ represents one of the key parameters in the determination of q (see, for example, EurOtop, 2007; Van der Meer et al., 2013), the dependency of the error on this parameter has been furthermore investigated. Similarly to Fig. 6, Figure 7 displays the distribution of the relative error $(q_s\text{-}q_{ANN})/q_s$ against the increasing values of $R_c/H_{m0,t}$. The greatest scatter is concentrated around the condition of zero-freeboard, and therefore the error decreases with increasing $R_c/H_{m0,t}$. The error is quite symmetrically distributed around the ideal condition (i.e. $(q_s\text{-}q_{ANN})/q_s = 0$), revealing that the ANN predictions are not biased with respect to the relative crest freeboard.

Generally, the diagram of Fig. 7, as well as Fig. 6, demonstrates that the performance of the ANN seems significantly affected by the number of the available tests representative of a given wave and structure condition.

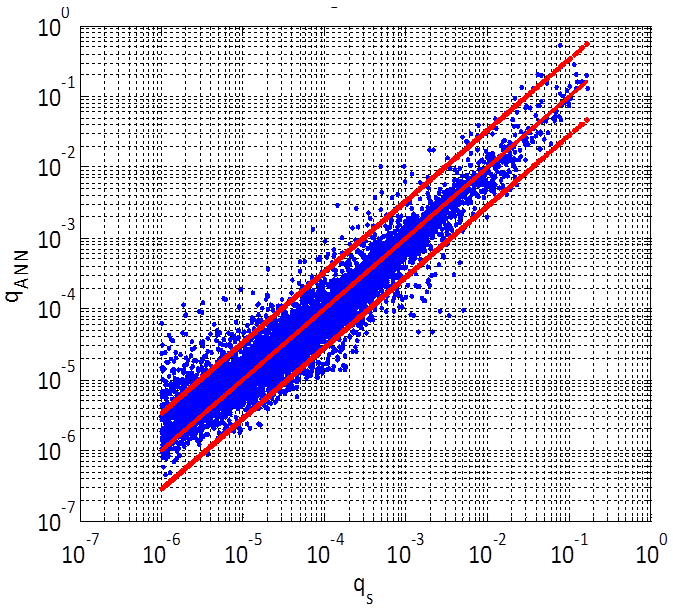

Figure 5. Comparison among the predicted values of q (q_{ANN}, ordinate) and the corresponding experimental values (q_s, abscissa); the bisector represents the ideal condition ($q_{ANN} = q_s$), while the external lines refer to the 95% confidence levels. Predictions resulting from the average of 500 bootstrap resamples of the ANN. Values in model scale.

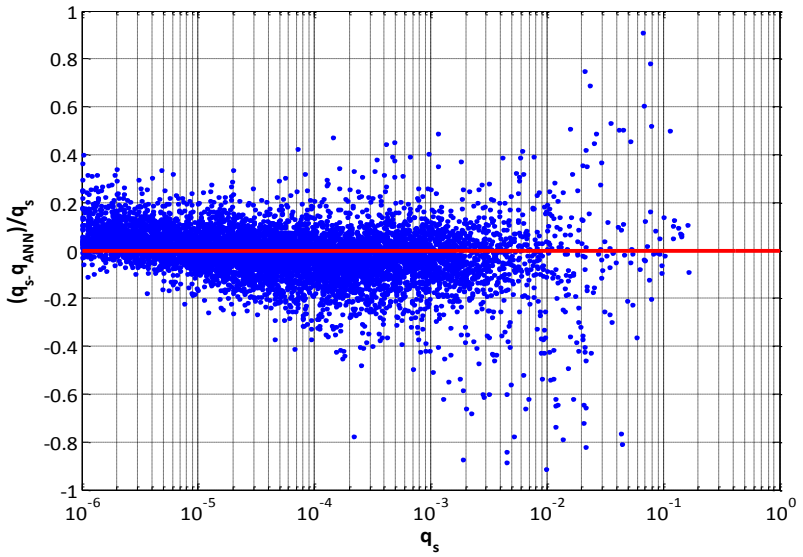

Figure 6. Distribution of the relative error $(q_s-q_{ANN})/q_s$ (ordinate) as a function of the experimental values q_s (abscissa). The continuous line represents the ideal condition of 0-error. Values in model scale.

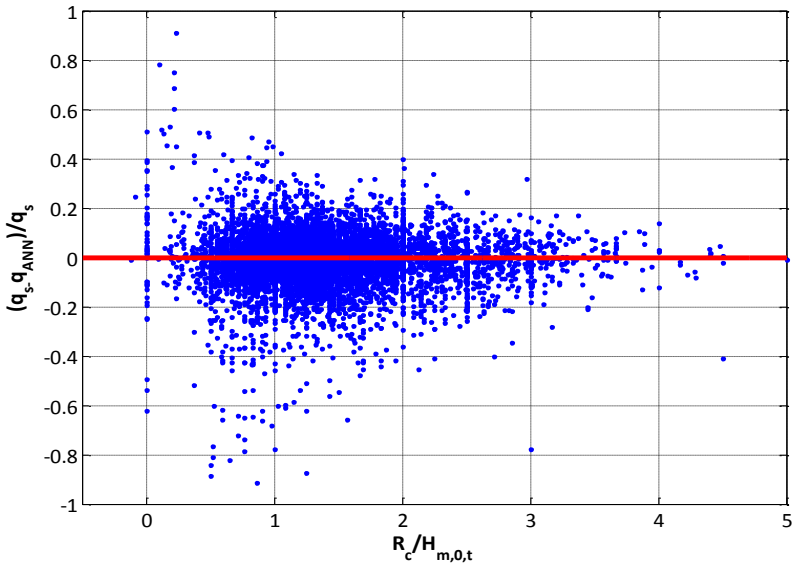

Figure 7. Distribution of the relative error (q_s-q_{ANN})/q_s (ordinate) as a function of the relative crest freeboard $R_c/H_{m,0,t}$ (abscissa). The continuous line represents the ideal condition of 0-error.

CONCLUSIONS

This paper presented an advanced ANN able to accurately predict the overtopping discharge q, the wave transmission coefficient K_t and the wave reflection coefficient K_r for a wide range of (complicated) structure geometries and wave conditions.

The ANN was trained on a new extended database based on CLASH (2004) and then extended, and consisting of more than 16,000 data. The database was prepared following the same structural and hydraulic schematization as in CLASH and including the following additional information: K_t and K_r, where available, and the average unit size D as representative of the structure elements.

Three set of input elements for the ANN were examined, i.e. the original CLASH ANN, the CLASH extended ANN with the inclusion of the parameter D and the ANN proposed by Zanuttigh et al. (2013) to verify the relevance of the D and to test the dimensionless treatment of the input elements.

The optimal architecture of the new ANN was investigated, by means of a careful sensitivity analysis to several characteristic elements which lead the training and the learning processes, i.e. the training algorithm, the number of hidden neurons and the early-stopping technique. The novel and exportable finding from this research is that the employment of bootstrap resampling technique allows to improve the capability of generalization of the ANN, i.e. the capability to predict outputs beyond the range of the training parameters, and to assess its performance.

The results of the ANN, with reference to the parameters q, K_r and K_t, were analyzed and the comparison among predicted and experimental values was carried out, revealing that the predictions are satisfactorily accurate, providing values of the mean square error in the range of [0.03; 0.05]. The analysis of the distribution of the errors underlined the importance of the homogeneity and extension of the database.

In particular, with regard to the prediction of small q, the enlargement of the overtopping dataset with new tests of modest or null values of q could improve the performance of the ANN. It is planned to verify the effects of introducing a classifier-quantifier (following Verhaeghe et al., 2008), in order to extend the capability of the ANN to deal with cases of low or zero overtopping.

ACKNOWLEDGMENTS

The support of the European Commission through Contract 244104 THESEUS ("Innovative technologies for safer European coasts in a changing climate"), FP7.2009-1 Large Integrated Project, is gratefully acknowledged.

REFERENCES

CLASH, 2004. Crest Level Assessment of coastal Structures by full scale monitoring, neural network prediction and Hazard analysis on permissible wave overtopping. EC-contract EVK3-CT-2001-00058. www.clash-eu.org.

EurOtop, 2007. European Manual for the Assessment of Wave Overtopping. T. Pullen, N.W.H. Allsop, T. Bruce, A. Kortenhaus, H. Schüttrumpf and J.W. van der Meer. At: www.overtopping-manual.com.

Formentin, S. M. and Zanuttigh B. 2013. Prediction of wave transmission through a new artificial neural network developed for wave reflection. *Proceedings of 7th International Conference on Coastal Dynamics*, http://www.coastaldynamics2013.fr/pdf_files/057_Formentin_SaraMizar.pdf.

Levenberg, K., 1944. A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics 2*, 164–168.

Lykke Andersen, T., 2006. *Hydraulic Response of Rubble Mound Breakwaters: Scale Effects - Berm Breakwaters*. Aalborg University, PhD thesis, 429 pp.

Marquardt, D., 1963. An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Mathematics* 11 (2), 431–441.

Oumeraci, H., Kortenhaus A. and Burg S. 2007. Investigations of wave loading and overtopping of an innovative mobile flood defence system: Analysis of model tests and design formulae, Report Nr. 949, *Technische Universität Braunschweig*, Leichtweiß-Institut für Wasserbau, Abteilung Hydromechanik und Küsteningenieurwesen.

Oumeraci, H., Kortenhaus A. and Haupt R. 2001. Untersuchung zur Abminderung des Wellenüberlaufs bei senkrechten Wänden durch Wellenabweiser, Report Nr. 865, *Technishe Universität Braunschweig*, Leichtweiß-Institut für Wasserbau, Abteilung Hydromechanik und Küsteningenieurwesen.

Oumeraci, H., Kortenhaus A. and Haupt R. 2004. Überarbeitung des Bemessungskonzepts für die Hochwasserschutzwände des privaten Hochwasserschutzes im Hamburger Hafen, Report Nr. 860, *Technishe Universität Braunschweig*, Leichtweiß-Institut für Wasserbau, Abteilung Hydromechanik und Küsteningenieurwesen.

Panizzo, A. and Briganti R. 2007. Analysis of wave transmission behind low crested breakwaters using neural networks. *Coastal Engineering*, 54, 643–656.

Van der Meer, J.W., Bruce T., Allsop W., Franco L., Kortenhaus A., Pullen T. and Schüttrumpf H. 2013. EurOtop revisited. Part 1: sloping structures. *Proceedings of ICE, Coasts, Marine Structures and Breakwaters*, Edinburgh, UK.

Van der Meer, J.W., Verhaeghe, H. and Steendam, G.J. 2009. The new wave overtopping database for coastal structures. *Coastal Engineering* 56, 108–120.

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. *Coastal Engineering* 54, 586–593.

Van Oosten, R.P. and Peixò Marco J. 2005. Wave transmission at various types of low-crested structures using neural networks. MsC Thesis, *TU Delft, Faculty of Civil Engineering and Geosciences, Hydraulic Engineering*.

Verhaeghe, H. 2005. Neural network prediction of wave overtopping at coastal structures. PhD dissertation, *Ghent University*, Belgium.

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. *Coastal Engineering* 55, 357–374.

Victor L., 2012. Optimization of the hydrodynamic performance of overtopping wave energy converters: experimental study of optimal geometry and probability distribution of overtopping volumes. PhD dissertation, *Ghent University*, Belgium

Wilmott, C.J. 1981. On the validation of models. *Physical Geography* 2, 184-194.

Zanuttigh, B. and J.W., Van der Meer, 2008. Wave reflection from coastal structures in design conditions. *Coastal Engineering*, 55, 771-779.

Zanuttigh, B., Formentin, S.M. and Briganti R. 2013. A Neural Network for the prediction of wave reflection from coastal and harbor structures. *Coastal Engineering* 80, 49-67.