Prediction of extreme and tolerable wave overtopping discharges through an advanced neural network

(Article begins on next page)

07 February 2025

PREDICTION OF EXTREME AND TOLERABLE WAVE OVERTOPPING DISCHARGES THROUGH AN ADVANCED NEURAL NETWORK

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

1) Associate Professor, PhD, University of Bologna,
DICAM, Viale Risorgimento 2, 40136 Bologna, Italy, barbara.zanuttigh@unibo.it

2) Research Fellow, PhD, University of Bologna,
DICAM, Viale Risorgimento 2, 40136 Bologna, Italy, saramizar.formentin2@unibo.it

3) Principal, Van der Meer Consulting bv, P.O. Box 11, Akkrum, 8490 AA,
The Netherlands, jm@vandermeerconsulting.nl;
Professor, UNESCO IHE, Westvest 7, Delft, 2611 AX, The Netherlands;
Professor, Delft University of Technology, Stevinweg 1, 2628 CN, Delft, The Netherlands

Abstract

This paper presents an Artificial Neural Network (ANN) to predict the wave overtopping discharge at coastal and harbour structures for a variety of wave conditions and complex geometries. The goal of this work is to provide a robust tool in both extreme and tolerable overtopping conditions, starting from the ANN recently developed by the authors for wave reflection, overtopping and transmission. Optimisation of the existing ANN is analysed: i) by training the ANN also on very low values of the overtopping discharge: i) by the set-up of an architecture consisting of a classifier-quantifier scheme; iii) i) through the modification of the weight factors included in the boot-strapping resampling technique. The accuracy of the optimised ANN is proved predicting new data and datasets.

Keywords: artificial neural network; wave overtopping; classifier-quantifier scheme; weight factors; training database

1. Introduction

Most coastal and harbour structures are constructed primarily to limit wave overtopping or prevent flooding. New challenges to the risk based design of these structures are posed by the ongoing effects of climate change, with sea level rise and increasing intensity and frequency of storms (THESEUS team, 2014). Therefore the accurate estimation of overtopping discharges and volumes, together with the characteristics of the overtopping flow over the structures, are extremely important to assess and assure the safety – or at least limit the exposure – of people, activities and goods.

Formulae and methods are available to predict overtopping at particular structures, often fairly simplified geometries, under given wave conditions and water levels (EurOtop, 2007; Van der Meer et al., 2009). Numerical models do exist that can simulate the wave-by-wave process and the details of 3D flows (Higuera et al., 2013), also with some simplification and in general with a significant effort for the preparation of the required data and the need for calibration. A good option to predict wave overtopping is to use physical model tests, but they are expensive and time consuming. They should certainly be considered for a final design, but are often a way too far in a preliminary design.

For conceptual design purposes, a simple and rapid approach is to use an Artificial Neural Network (ANN), which is particularly recommended in case of complicated structure geometries and variable wave conditions (EurOtop, 2007). This kind of predictive method requires however a homogeneous and "wide-enough" database to train the ANN, based on a number of parameters for total range of possible output values. There are specific cases where an ANN cannot deal with, such as very complex walls and double promenades (Van Doorslaer et al., 2015), see for details the methodology released within the PC-OVERTOPPING calculator (http://www.overtopping-manual.com/calculation_tool.html).

The ANN developed within the CLASH (2004) project and proposed by EurOtop (2007) for the prediction of the average overtopping discharge, q, is the ANN by Van Gent et al. (2007). Further analysis and other ANNs have been delivered during and after CLASH (2004): the ANN by Verhaeghe (2005) for q and the ANN developed by the authors (Zanuttigh et al., 2014). The last one predicts the main wave-structure interaction parameters: besides q, the wave reflection, K_r, and transmission, K_t coefficients. These ANNs showed a good performance when predicting the same database used for training but were not systematically tested against new data, i.e. data that were not already used for training.

The goal of this work is to provide coastal designers with a tested robust and accurate ANN able to represent extreme and tolerable wave overtopping discharges for a wide range of structure types under a variety of wave conditions. This work is based on the recent research carried out by the authors (Zanuttigh et al., 2014) and is going to answer key questions such as:
  - How the ANN can deal with zero measured values of overtopping?

- Does the implementation of a new classifier-quantifier scheme (based on the idea by Verhaeghe, 2005; Verhaeghe et al., 2008) allow a better prediction of the extreme values or is the effect of the error propagating from the first classification dominant?
- How should weight factors be introduced in the ANN training, to let the ANN learning from the more reliable data?
- How can the results of the ANN be used in practice, accounting for model and scale effects?

The paper structure is as follows. Section 2 describes the new extended database for wave overtopping, including the explanation of the differences and updates with respect to CLASH (2004). Section 3 focuses on the optimisation of the existing ANN (Sub-section 3.1) with respect to the training process and the representation of the extreme values. The analysis of extreme conditions accounts for i) the low values ($q<10^{-6}$ m^3/s/m), which are at present overestimated in previous works by Van Gent et al. (2007) as well as Verhaeghe et al. (2008); and ii) the high values ($q>10^{-3}$ m^3/s/m), whose accurate estimate is essential to assess the potential impacts of disasters. As for the first objective, the training of the ANN including all the non-zero values of the discharge is examined (Sub-section 3.2) and its capacity to deal specifically with zero values is discussed (in Sub-section 3.3). The definition of the weight factors affecting the training process is also analysed and a new way to evaluate them is proposed (Sub-section 3.4) to assure a more balanced assessment of the data reliability and complexity. As for the second objective, the architecture of the ANN is modified into a classifier-quantifier scheme, which is inspired by the work of Verhaeghe et al. (2008), but is very different both as purpose and as set-up (Sub-section 3.5). Section 4 provides the accuracy of the final ANN (Sub-section 4.1) when predicting either datasets excluded from the training database (Sub-section 4.2), or new data and datasets (Sub-section 4.3). The limitations of the optimised ANN with regard to the model (i.e. wind, currents) and scale (i.e. permeability) effects are also discussed (Sub-Section 4.4). Conclusions are finally drawn in Section 5.

2.   The new database: parameters and schematisation of the structure

The wave overtopping Data Base (DB hereafter) employed in this work is composed by more than 13,500 tests mainly derived from the CLASH DB (Van der Meer et al., 2009), which consists of more than 10,000 irregular tests on dikes, rubble mound breakwaters, berm breakwaters, caissons and combinations of these structures resulting in complicated geometries. The following datasets have been added to the existing CLASH DB:
- 170 tests on vertical walls (Oumeraci et al., 2007);
- 56 tests on rubble mound with cobs (Besley et al., 1993);
- 75 tests on smooth structures with berms (private communication);
- 103 tests on harbour caissons (private communication);
- 249 tests on reshaping berm breakwaters (of which 30 from Lykke Andersen et al., 2008 and the remaining 219 from private communication);

- 366 tests on smooth steep slopes by Victor and Troch (2012);
- 671 tests on smooth slopes in combination with walls by Van Dorslaer et al., (2015).

This extended DB is part of a DB recently assembled to gather all the available data on wave overtopping, reflection and transmission (Zanuttigh et al., 2014).

The DB set-up follows the original CLASH DB, by adopting the same schematization of the structures (see Fig. 1) and maintaining the same geometric and hydraulic parameters. In addition, the following parameters have been included

- K_r and K_t where available,
- the average unit size D representative of the structure elements around the water level. It could be the D_{n50} for rock armour, D_n for concrete armour, but it could also be the height of a step of a staircase geometry.

A new original procedure has been developed to evaluate few parameters (D, γ_f, $cot\alpha_{incl}$) in such a way to be consistent through the DB. D is calculated as the weighted average of the characteristic downslope D_d and upslope D_u sizes of the elements in the run-up/down area, i.e. within ±1.5 $H_{m0,t}$ above and below the still water level, following the formula:

$$D = \frac{D_d \cdot (h_{sub} - h_b) + D_u \cdot (h_b + h_{em})}{h_{sub} + h_{em}} \quad , \tag{1}$$

where $h_{sub} = \min(1.5 \cdot H_{m0,t}; h)$; $h_{em} = \min (1.5 \cdot H_{m0,t}; A_c)$.

Consistently, also the roughness factor γ_f and the average slope $cot\alpha_{incl}$ that is the average slope in the run-up/down area are now respectively evaluated as

$$\gamma_f = \frac{\gamma_{fd} \cdot (h_{sub} - h_b) + \gamma_{fu} \cdot (h_b + h_{em})}{h_{sub} + h_{em}} \quad , \tag{2}$$

$$cot\alpha_{incl} = \frac{cot\alpha_d \cdot (h_{sub} - h_b) + B + cot\alpha_u \cdot (h_b + h_{em})}{h_{sub} + h_{em}} \quad , \tag{3}$$

Eq. (3) is valid for $|h_b| < 1.5 \cdot H_{m0t}$; otherwise $cot\alpha_{incl} = cot\alpha_d$ ($h_b > 0$) or $cot\alpha_{incl} = cot\alpha_u$ ($h_b < 0$).
The ANN tool that has been prepared and is going to be delivered through the website requires the users to enter –through an interface - the correct values of $cot\alpha_d$ and $cot\alpha_u$, D_d and D_u, of γ_{fd}, γ_{fu}, as well as the other input parameters for the ANN to be described in Section 3. The tool then calculates the average values according to Eq.s (1)-(3).

In the CLASH database, each test is characterised by a reliability and a complexity factors, RF and CF, which describe respectively the reliability level of the measured or estimated structural and hydraulic parameters and how well a structure geometry could be described by the geometrical parameters. The values of RF and CF vary from 1 to 4, i.e. from the more reliable or simple to the more unreliable or complex situation. Sometimes interesting data were found during the CLASH project that could not be used for the training of the ANN, such as tests with a wind simulator. In the CLASH database these tests were present, but were given a reliability factor RF=4, indicating that the data should not be used, however in reality the data could be quite reliable. To overcome this problem, an extra column has been created in the new

database. This column indicates whether the test belongs to the "core"-data, which means that it can be considered as a case to be used for training of the ANN, or that it is outside this core-data but with a peculiar feature denoted by a letter, for instance: w=wind, p=prototype, c=current, b=bull nose, pc=perforated caisson.
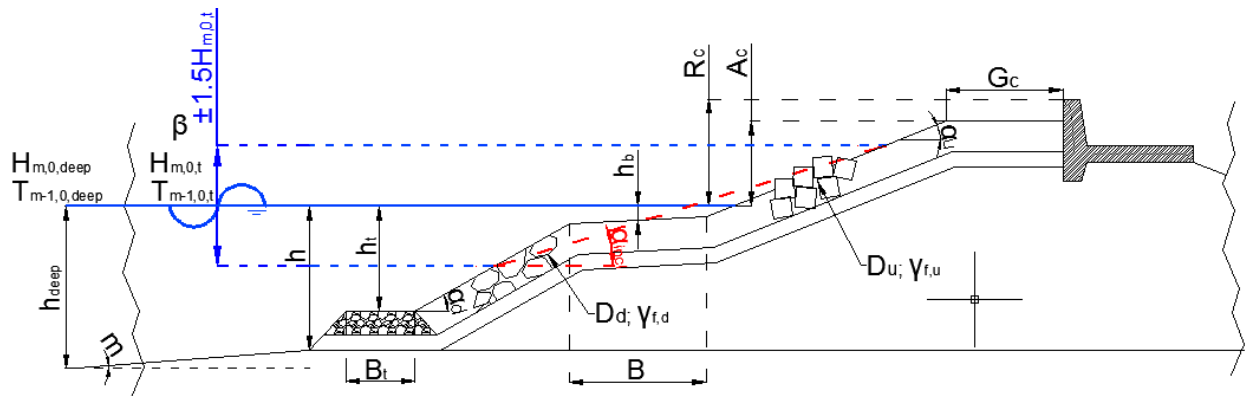

Figure 1 – Schematization of the structure based on CLASH, including some of the geometrical and hydraulic parameters.

The final DB includes therefore 8 parameters more than the CLASH DB, consisting of:

- 11 hydraulic parameters, characterizing the wave attack conditions (extension with the spreading parameter, as all the available laboratory reports include this value in the case of short-crested waves; long-crested waves will get the spreading=0);
- 21 structural parameters (extension with D_d, D_u, D, γ_{fd}, γ_{fu});
- 4 general parameters, the test label, the reliability RF and the complexity CF factors, the inclusion in the "core" data for the ANN training;
- 3 output parameters (extension with K_r and K_t).

Table 1 reports the type and the number of all the parameters included in the extended DB, in comparison with the original CLASH DB.

Figures 2 and 3 provide the comparison of the new DB to the original CLASH DB. In Figure 2 the values of the average overtopping discharge q are grouped in 9 classes according to the order of magnitude of q. One additional class collects the values of q identically equal to zero, which roughly represent 12% of the overall amount of data. In Figure 3 the data are divided on the basis of the structure type and/or wave attack, as well as in Zanuttigh and Van der Meer (2008). The DB therefore consists of 7 groups: rock permeable straight slopes (denoted by the label "A"), rock impermeable straight slopes ("B"), armour units straight slopes ("C"), smooth and straight slopes ("D"), structures with combined slopes and berms ("E"), vertical walls ("F") and oblique wave attacks ("G"). The G group indeed includes all the tests performed in 3D conditions: all the other groups A-F are related to different structures tested in wave flumes at different scales. As the presence of 3D effects is well known from the literature, the

5

identification of a specific group has been selected to ease the comparison among experimental conditions and the training of the ANN. From the diagrams of Figure 3, it can be observed that most of the tests is concentrated within groups E and F (reaching about the 42% of the total), while the remaining tests are almost evenly spread around.

The two DBs (the original CLASH and the new extended one) show nearly the same distribution of the values of q (Fig. 2), while a more significant difference can be detected when comparing the structure types (Fig. 3). In particular, the new DB includes a wider number of vertical walls (group "F") and rock impermeable structures (group "B"), but smaller percentages of non-straight slopes (group "E") and smooth slopes (group "D"). Such different assortment of the new DB is not expected to affect the quality of the predictions, since the large amount of available tests ensures the representativeness of each structure type. The narrower group "B", representing the 7% of the whole DB, collects more than 900 data.

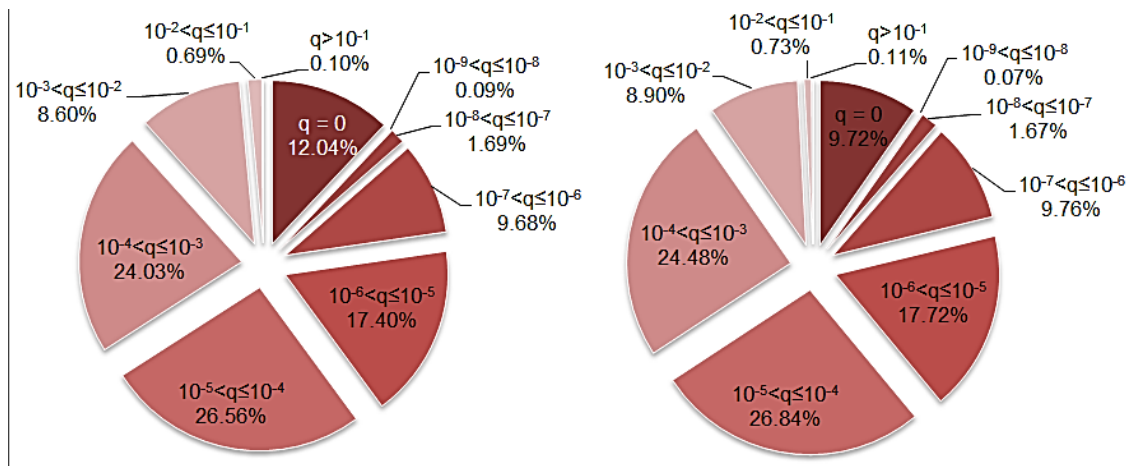Figure 2 – The DB distribution of the values of q by orders of magnitude. To the left, the complete DB; to the right, the original CLASH DB (Van der Meer et al., 2009).
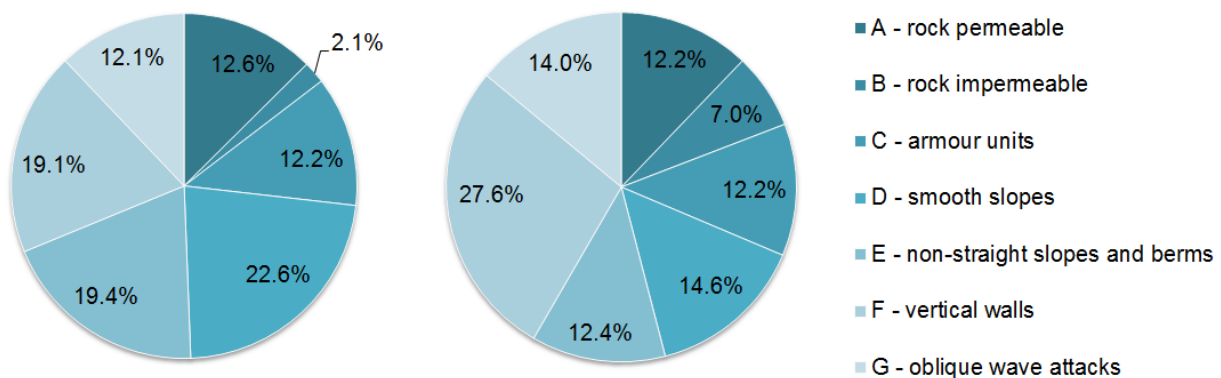

Figure 3 – The DB distribution of the values of q based on the type of the structures and/or the wave attack, following Zanuttigh and Van der Meer (2008). To the left, the complete DB; to the right, the original CLASH DB (Van der Meer et al., 2009).

Table 1 - Parameters included in the new extended database compared with the ones included in the original CLASH database.

#	Parameter	Unit	Type	CLASH	New	Definition of the parameter
1	Name	[-]	general	√	√	
2	$H_{m,0,deep}$	[m]	hydraulic	√	√	Off-shore significant wave height
3	$T_{p,deep}$	[s]	hydraulic	√	√	Off-shore peak wave period p
4	$T_{m,deep}$	[s]	hydraulic	√	√	Off-shore average wave period
5	$T_{m-1,deep}$	[s]	hydraulic	√	√	Off-shore spectral wave period
6	h_{deep}	[m]	structural	√	√	Off-shore water depth
7	m	[-]	structural	√	√	Foreshore slope
8	β	[°]	hydraulic	√	√	Wave obliquity
9	Spreading	[-]	hydraulic		√	Spreading
10	h	[m]	structural	√	√	Water depth at the structure toe
11	$H_{m,0,t}$	[m]	hydraulic	√	√	Significant wave height at the structure toe
12	$T_{p,t}$	[s]	hydraulic	√	√	Peak wave period at the structure toe
13	$T_{m,t}$	[s]	hydraulic	√	√	Average wave period at the structure toe
14	$T_{m-1,t}$	[s]	hydraulic	√	√	Spectral wave period at the structure toe
15	h_t	[m]	structural	√	√	Toe submergence
16	B_t	[m]	structural	√	√	Toe width
17	Type	[-]	structural	√	√	Type of structure and armour unit
18	$cot\alpha_d$	[-]	structural	√	√	Cotangent of the angle that the structure part below/above the berm makes with a horizontal
19	$cot\alpha_u$	[-]	structural	√	√	
20	$cot\alpha_{excl}$	[-]	structural	√	√	Cotangent of the mean angle that the structure makes with a horizontal, excluding/including the berm, in the run-up/run-down zone
21	$cot\alpha_{incl}$	[-]	structural	√	√	
22	γ_{fd}	[-]	structural		√	Roughness factor for $cot\alpha_d$
23	γ_{fu}	[-]	structural		√	Roughness factor for $cot\alpha_u$
24	γ_f	[-]	structural	√	√	Roughness factor (average in the run-up/down area in the new DB, Eq. 2)
25	D_d	[-]	structural		√	Size of the structure elements along $cot\alpha_d$
26	D_u	[-]	structural		√	Size of the structure elements along $cot\alpha_u$
27	D	[m]	structural		√	Average size of the structure elements in the run-up/down area (Eq. 3)
28	B	[m]	structural	√	√	Berm width
29	h_b	[m]	structural	√	√	Berm submergence
30	$tan\alpha_b$	[-]	structural	√	√	Berm slope
31	B_h	[m]	structural	√	√	Horizontal berm width
32	A_c	[m]	structural	√	√	Crest height with respect to swl
33	R_c	[m]	structural	√	√	Wall height with respect to swl
34	G_c	[m]	structural	√	√	Crest width
35	RF	[-]	general	√	√	Reliability Factor
36	CF	[-]	general	√	√	Complexity Factor
37	Pow	[-]	hydraulic	√	√	Overtopping probability
38	q	[m³/s/m]	output	√	√	Wave overtopping discharge per unit width
39	K_r	[-]	output		√	Wave reflection coefficient
40	K_t	[-]	output		√	Wave transmission coefficient

3. Optimisation of the existing ANN with focus on extremely high and low wave overtopping discharges

This Section aims at describing the research and the advances specifically implemented for the prediction of the extreme values of the overtopping discharge, starting from the ANN developed by the authors (short overview in Sub-section 3.1). The training DB of the ANN is extended to all the non-zero values of q (Sub-section 3.2), achieving a better representation of the low values of q that were overestimated by previous ANNs. The newly trained ANN shows also a good capability to deal with the measured zero values of q that are reproduced as the minimum values within each dataset (Sub-section 3.3). The definition of the Weight Factors to be included in the training process is also discussed (Sub-section 3.4). A classifier-quantifier scheme has been implemented for a better prediction of extreme values of q and compared with the improved ANN (Sub-section 3.5).

3.1 The existing ANN

This paper describes the development of the ANN presented by Zanuttigh et al. (2014) for the prediction of the wave overtopping discharge q, the wave reflection and the wave transmission coefficients K_r and K_t.

This ANN is a multilayer network, composed of 1 input layer of 15 dimensionless input parameters, 1 hidden layer of 20 hidden neurons and 1 bias, and 1 output layer of 1 output parameter. Moreover, the ANN
- is based on a "feed-forward back-propagation" learning algorithm,
- adopts the *Levenberg – Marquardt* (Marquardt, 1963; Hagan and Menhaj, 1994) training algorithm;
- implements the hyperbolic tangent sigmoid and the linear transfer functions as the hidden neurons and the output neuron transfer functions respectively;
- uses the *mse* (mean squared error) for estimating the error;
- adopts the *Bootstrap resampling* technique for the assessment of the ANN performance.

The list of the 15 dimensionless input parameters together with a synthetic description is given in Table 2, while the meaning of the symbols is explained in Figure 1. All the structure heights (h_t, h_b, R_c, A_c) are scaled with the significant wave height ($H_{m0,t}$) and, similarly, all the structure widths (B_t, B, G_c, A_c) are scaled with the wave length ($L_{m-1,0,t}$). The use of $H_{m0,t}$ as height-scaling parameter allows to describe the wave dissipation due to local breaking on the different parts of the structure (h_t, h_b) and the potential of wave overtopping and transmission (R_c, A_c). The use of $L_{m-1,0,t}$ as width-scaling parameter allows to account for the induced local reflection that might be in phase, or not, with the wave reflection from other parts of the structure slope (Numata, 1976). The two key processes, i.e. wave breaking by steepness $H_{m0,t}/L_{m-1,0,t}$ and shoaling $h/L_{m-1,0,t}$, are also accounted for.

Table 2 – Synthesis of the 15 selected dimensionless input parameters of the adopted ANN.

Parameter	Min	Max	Type	Representation of
$H_{m0,t}$	0.017	1.480	Wave attack	Significant wave height in front of the structure - Scale parameter
$L_{m-1,0,t}$	0.691	156.944	Wave attack	Wave length depending on the spectral wave period in front of the structure - Scale parameter
$H_{m0,t}/L_{m-1,0,t}$	0.002	0.084	Wave attack	Wave steepness (breaking)
B	0.000	80.000	Wave attack	Wave obliquity
$h/L_{m-1,0,t}$	0.003	0.942	Wave attack	Shoaling parameter
$h_t/H_{m-1,0,t}$	0.430	25.926	Geometry	Effect of the toe submergence
$B_t/L_{m-1,0,t}$	0.000	0.760	Geometry	Effect of the toe width
$h_b/H_{m0,t}$	-2.133	7.143	Geometry	Effect of the berm submergence
$B/L_{m-1,0,t}$	0.000	0.972	Geometry	Effect of the berm width
$A_c/H_{m0,t}$	-5.247	16.076	Geometry	Effect of the crest freeboard
$R_c/H_{m0,t}$	0.000	16.076	Geometry	Effect of the crest freeboard accounting for the presence of a crown wall
$G_c/L_{m-1,0,t}$	0.000	0.362	Geometry	Effect of the crest width
m	0.000	1000.0	Geometry	Effects of the foreshore slope
$cot\alpha_d$	0.000	7.000	Geometry	Lower slope
$cot\alpha_{incl}$	-1.347	12.820	Geometry	Average slope in the run-up/down area
γ_f	0.330	1.000	Structure characteristics	Roughness factor
$D/H_{m0,t}$	0.000	1.298	Structure characteristics	Indication of structure stability

3.2 Improving the prediction of tolerable wave overtopping discharges

The existing ANN was trained (Zanuttigh et al., 2014) against the data with $q>10^{-6}$ m^3/s/m only (for a total of 8,194 tests). The exclusion of the values of $q\leq10^{-6}$ from the training DB was based on the hypothesis that the low overtopping data are likely to be affected by greater measurement errors (Van Gent et al., 2007) and on the assumption that $q=10^{-6}$ m^3/s/m represents a threshold value to distinguish between "negligible" and "significant" overtopping, i.e., $q\leq10^{-6}$ and $q>10^{-6}$ m^3/s/m respectively (Verhaeghe et al., 2008).

The existing ANN tends to overestimate the low values of q ($q<10^{-5}$ m^3/s/m) as well as the CLASH ANN (Van Gent et al., 2007; Verhaeghe et al., 2008). The analysis presented here starts from the observation that this bias may be mainly induced by the elimination of the values of $q<10^{-6}$ m^3/s/m from the training process.

In order to minimize the bias, all the non-zero values of q have been included in the training DB. A discussion about the representation of the zero values is given in Sub-section 3.3. The introduction of all the values with $q>0$ determines an increase of the number of the available

tests from 8,194 up to 9,303 tests. These are 1,109 additional tests, representing more than the 13% of the total amount of tests, as indicated by Fig. 2. Following the previous works, the tests with factors RF or CF equal to 4 are discarded from the training.

Table 3 presents the quantitative results of the newly trained ANN in comparison with the original ANN (rows 1 and 2). The results are provided in terms of the root mean square error (rmse), the Willmott Index (WI), the coefficient of determination R^2 and the number of "large errors", as described more in detail in the previous work by the authors (Zanuttigh et al., 2014). The number of large errors correspond to the percentage of tests (with respect to the total number of tests) for which the ANN gives systematically (i.e., in more than the 50% of the predictions) an output value that differs more than 1.5 times from the experimental corresponding value. The values in Table 3 are the average indexes and the corresponding standard deviations derived after 500 bootstrapped resamples of the DB. The very low and almost constant values of the standard deviations - with the exception of the prediction of $q \leq 10^{-6}$ - for each tested case and for each index can be justified by the combination of the database size with the number of bootstraps. The standard deviations are therefore reported more for completeness rather than for their real meaningfulness, since the assessment of the ANN performance is based on the average values.

From Table 3 the ANN trained on the whole non-zero values ($q>0$) of the DB is characterized (rows 1, 3) by a lower value of the rmse and a significantly lower value of the percentage of large errors with respect to the original ANN trained on $q>10^{-6}$ m^3/s/m only (row 2), being the values of WI and R^2 almost constant. The evident improvement of both rmse and of the large errors when the ANN trained on $q>0$ is used to predict the cases with $q>10^{-6}$ m^3/s/m (compare rows 1 and 3) proves that the ANN performance decreases when predicting the smaller values of q (see also the discussion in Sub-section 3.4).

Figure 4 shows the distribution of the predicted values q_{ANN} versus the corresponding measurements q_S. The predictions of the ANN trained on $q_s>0$ (right plot) are far more symmetric than the original ones with the ANN trained on $q \geq 10^{-6}$ m^3/s/m only (left plot). The overestimation of q in the range $[10^{-6};10^{-5}]$ m^3/s/m is significantly reduced. Also the error bands, representing the 95% confidence interval are narrower, revealing that the predictions obtained training the ANN on the values of $q_s>0$ is not only less biased, but also more accurate.
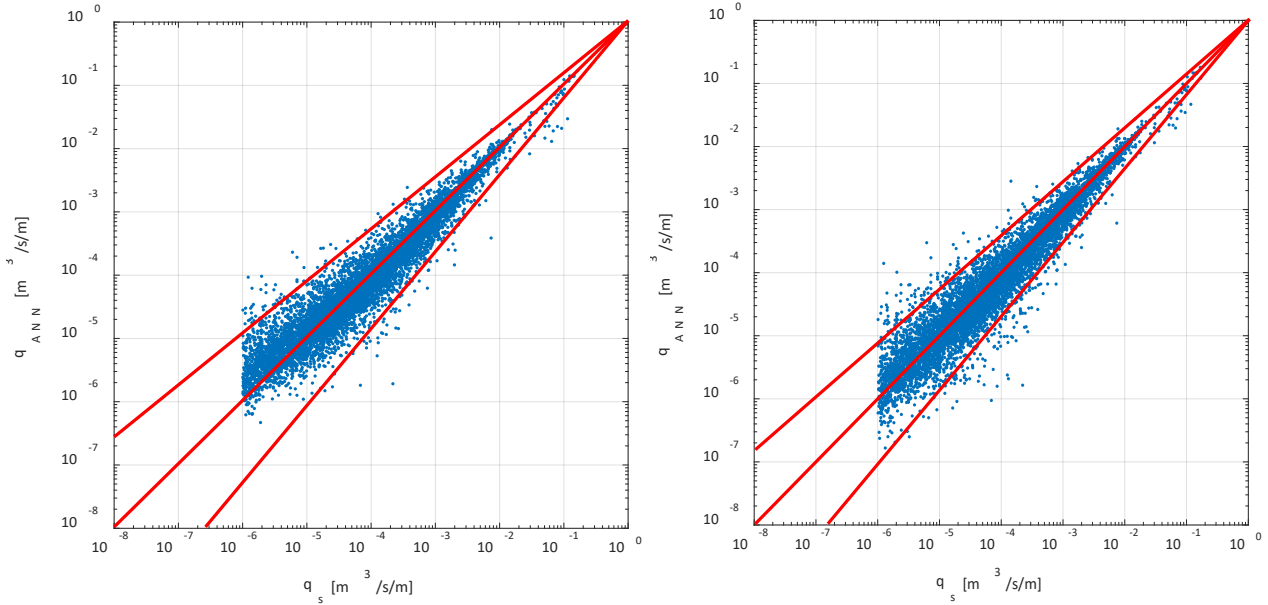
Figure 4 – Comparison of the predicted q_{ANN} and experimental q_s values, when $q_s \geq 10^{-6}$ m^3/s/m. To the left, the ANN is trained only on $q_s \geq 10^{-6}$ m^3/s/m; to the right, the ANN is trained on $q_s > 0$ m^3/s/m.

Table 3 – Quantitative performance of the ANN. Effects of training the ANN on $q \geq 10^{-6}$ m^3/s/m and on $q > 0$ m^3/s/m (rows 1-3). Effects of training the ANN with different Weight Factors WFs (rows 3 and 4). Use of a single ANN, i.e. the 'classifier', C scheme, and use of a combination of ANNs, i.e. the 'classifier-quantifiers, CQ scheme (rows 5-11).

	Scheme type	WF, Eq.#	Range of q (training) [m^3/s/m]	Range of q (prediction) [m^3/s/m]	Rmse [m^3/s/m]	WI [-]	R^2 [-]	Large errors [%]
1	C	4	$q > 10^{-6}$	$q > 10^{-6}$	0.052±0.005	0.974±0.006	0.90±0.02	7.0
2	C	4	$q > 0$	$q > 10^{-6}$	0.042±0.002	0.975±0.004	0.91±0.01	0.3
3	C	4	$q > 0$	$q > 0$	0.050±0.004	0.974±0.007	0.90±0.02	3.3
4	C	5	$q > 0$	$q > 0$	0.047±0.002	0.977±0.003	0.92±0.01	2.6
5	CQ	5	$q > 0$	$q > 0$	0.048±0.003	0.967±0.007	0.88±0.02	4.2
6	C	5	$q > 0$	$q \geq 10^{-3}$	0.026±0.004	0.96±0.02	0.90±0.1	0.0
7	CQ	5	$q \geq 10^{-4}$	$q \geq 10^{-3}$	0.023±0.006	0.97±0.02	0.91±0.1	0.0
8	C	5	$q > 0$	$10^{-6} < q < 10^{-3}$	0.044±0.002	0.958±0.005	0.84±0.02	0.4
9	CQ	5	$q > 0$	$10^{-6} < q < 10^{-3}$	0.045±0.003	0.955±0.007	0.83±0.02	0.4
10	C	5	$q > 0$	$q \leq 10^{-6}$	0.075±0.004	0.68±0.03	-	16
11	CQ	5	$q \leq 10^{-5}$	$q \leq 10^{-6}$	0.076±0.007	0.65±0.04	-	17

3.3 How to deal with "zero" measurements of the wave overtopping discharge?

Based on the re-training of the ANN performed in Sub-section 3.2, we have now an optimized ANN capable to deal with low values of q without bias or overestimation. The ANN is trained on tests where an overtopping discharge was measured. It cannot be trained on tests where the overtopping was zero, as the training results in a continuous relationship between input and output parameters and a zero-value gives a discontinuity in this process. This means also that the idea proposed in Verhaeghe et al. (2008), to build up a specific DB of zeros to better train the ANN, is not applicable.

Due to the presence of small overtopping discharges in the DB, the ANN is also able to predict (very) small q-values even for test conditions where the laboratory was not able to measure overtopping, maybe due to limitations in the measuring equipment.

Here the newly trained ANN is adopted to predict the zero values within the datasets that include these zeros and the predictions are then compared - dataset by dataset - with both the measured and predicted non-zero values. Few examples are given in Figure 5 for four different datasets as a function of the relative crest freeboard $R_c/H_{m0,t}$. The plots compare the measured values of q with the predicted values in case of overtopping and no overtopping. Looking at these points, the following observations can be made for each plot/dataset:

i. the predictions of the measured zeros are always lower than the predictions of the other non-zero values;

ii. the predictions of the zeros follow the same trend with $R_c/H_{m0,t}$ as the other predicted/measured non-zero values;

iii. overall the whole ANN predictions show a good degree of continuity.
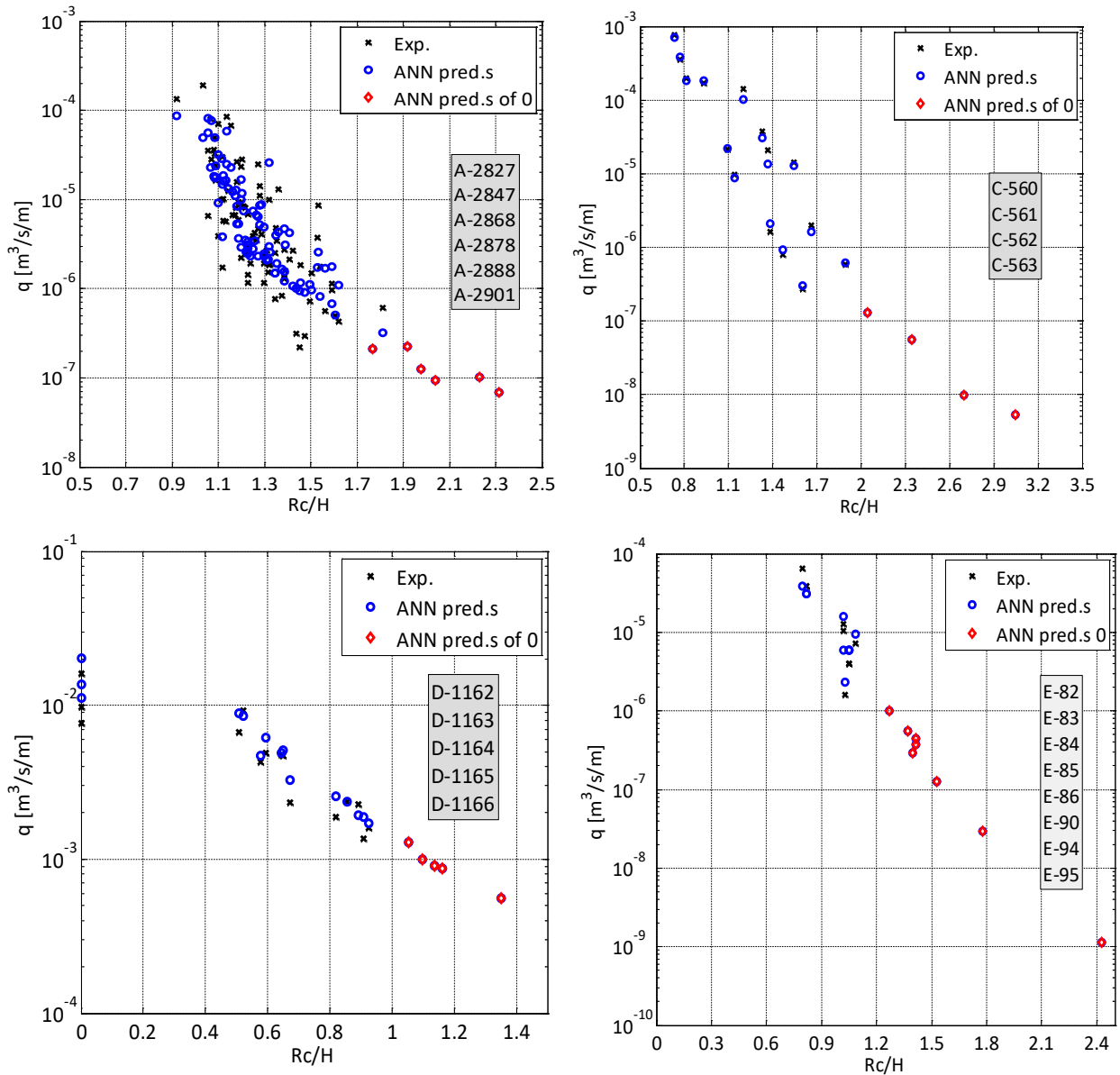
Figure 5 – Experimental (black crosses) and predicted values (blue void circles) that correspond to the experimental values and to the measured zeros (red void diamonds) within a given dataset. The values are shown as a function of the relative crest freeboard R_c/H_{m0t}. Each plot refer to a single dataset, from top to bottom and from left to right: rock permeable structures, rubble mound with armour units, vertical walls, and structures with berms.

The definition of zero overtopping depends on the set-up of measuring overtopping discharges. Limits to measure overtopping are not given in the DB other than that the laboratory has decided that in some tests "there was no overtopping". These limits are not consistent through the DB and this can also be concluded from Fig. 5. For example, dataset D in Fig. 5 shows that q-values smaller than 10^{-3} m³/s/m could not be measured, where dataset A shows that q-values up to 2.10^{-7} m³/s/m could be measured. This is a very significant difference, due to scale as well as measuring systems (weighting boxes, wave or pressure gauges in an overtopping box, or propellers at the crest of a structure).

The prediction by the ANN overcomes the inhomogeneity of the DB on zero-values of q. It always gives a value for q, it is then up to the user to decide whether this overtopping discharge is indeed significant, or it can be regarded as "no overtopping", or "below a tolerable limit". In case of a laboratory test set-up, the limit of the set-up determines what should be regarded as "no overtopping".

As a final comment, it should also be considered that both the measurements and the results of the ANN - as it is trained on the experimental DB - are affected by model and scale effects (De Rouck et al., 2005; Franco et al., 2009). This will be discussed in Sub-section 4.4.

3.4   The definition of the Weight Factors

This Sub-section focuses on the possibility to adopt different Weight Factors (WF) to drive the "random" selection of the training data in the BS towards those tests which are considered "more reliable". By definition the introduction of the WF will worsen the prediction of the ANN when used for predicting the same data used for the training. In fact, an ANN trained on a dataset where all the data are equally important will predict them with greater accuracy, but it will also provide a prediction of new data biased by the less reliable data included in the training DB. In essence one can say that a system of WF has to be applied, but it should not over estimate reliable or less reliable data. It is therefore difficult to find a good balance.

Based on the definition of the factors RF and CF (see Section 2), Van Gent et al. (2007) and Verhaeghe et al. (2008) adopted the following WF in the bootstrapping

WF=(4-RF)·(4-CF)   (4).

The more reliable the test the lower the RF and/or the better the structure schematisation the lower the CF and therefore the higher the values of the WF. A very reliable and simple structure with WF=9 has a 9 time higher probability of being selected for training than a less reliable test on a complex structure with WF=1. Since the definition of WF by means of Eq. (4) was given within the CLASH project, the corresponding WF are named as WF(CLASH) hereafter. Starting from the definition of WF(CLASH), the existing ANN has been trained on 100 boot-strap resamplings of the complete DB by using WF(CLASH).

The distribution of the values of the WF(CLASH) across the DB is given in Figure 6. Nearly 20% of the DB belongs to WF(CLASH)=0, and is discarded from the training. The 2 classes of WF(CLASH)= 6 and 9 include approximately 50% of the data, while the remaining 4 ones, WF(CLASH)=1-4, represent altogether 30% of the DB only. It can be concluded that the definition of WF(CLASH) as in Eq. (4) is such that the WF vary within a quite wide range of values (from 1 to 9), and this can significantly affect the training process. Since most of the DB is associated to high values of WF(CLASH)=6 and 9, the distribution of the WF induces a systematic training of the ANN towards these most numerous data, while the other data characterised by WF(CLASH)=1-4 may be under-represented. Therefore the combination of

the available data with the corresponding values of WF(CLASH) tends to reduce the effectiveness in the training process due to the wide distribution of the WF(CLASH).

The purpose of the analysis now is to define and test a new more balanced formulation for the WF than the WF(CLASH):

$$WF(a) = 6/(RF+CF), WF(a)=0 \text{ if RF or CF}=4 \tag{5}$$

where RF and CF are still the factors defined in CLASH, and the value of 6 stands for the maximum value of their sum. Eq. (5) is meant to be applied only to tests characterised by RF or CF different from 4, as these corresponding tests are automatically discarded from the training. Eq. (5) gives values of WF in the range 1-3 and therefore has a narrower distribution, and an expected smoother effect on the training process than the WF(CLASH). The definition of WF(a) as in Eq. (5) aims to equally weigh the contribution of RF and CF, so that a test that is characterised by a simple geometry (CF=1), but is less reliable as for the measurement techniques or the absence of active wave absorption (RF=3), has the same WF(a) of a test that is considered very reliable (RF=1), but has a quite complex geometry (CF=3).

The distribution of the values of WF(a) across the DB is also given in Figure 6. Of course, also in this case nearly 20% of the data belong to WF=0 (i.e. RF and/or CF equal to 4), as well as for WF(CLASH). WF(a) defines 5 classes of values (just 1 class less than WF(CLASH)), that vary with continuity and correspond to data percentages, ranging from about 7% of the data associated to WF(a)=1 to the maximum of about 26% associated to WF(a)=2.

Table 3 gives the quantitative results of the ANN trained by implementing WF(CLASH) as in Eq. (4) or WF(a) according to Eq.(5). The rmse is slightly greater, as well as WI and R^2 are a little lower for WF(CLASH) than for WF(a), being the main difference in the number of the large errors (see rows 3 and 4 of the table).

Figure 7 shows the comparison among the predicted and measured values of q derived from the ANN trained with WF(CLASH), Eq. (4), and WF(a), Eq.(5). The values of q associated to the different values of the WF are marked by the corresponding different colours. The graph shows that the highest scatter is not necessarily correlated to the lowest values of WF, independently from the definition of WF. This may be explained by the possibility that the RF and CF do not always reflect the actual accuracy of each test or measurements and that their definition is to some extent affected by subjectivity. For each class of WF, the percentage of values exceeding the 95% confidence bands and the large errors are reported in Table 4. In both cases of WF(CLASH) and WF(a) there is no evidence that the accuracy of the predictions is increasing with increasing the values of WF, accounting also for the number of tests included in each class. The error distribution with the values of the WFs is therefore non-monotone, with the only exception of WF=1 where most of the errors are concentrated in. Overall the outliers and the large errors are significantly lower in case of the ANN including WF(a) instead of WF(CLASH). The combined results of Table 3 and 4 suggest that a better ANN performance is achieved when using WF(a).
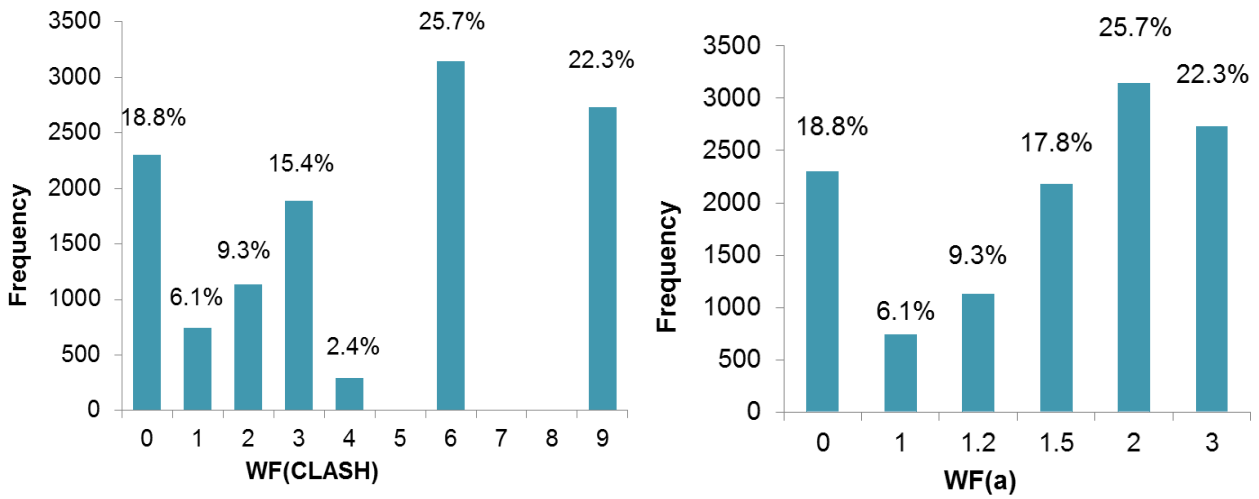
Figure 6 – Distribution of the WF across the DB. To the left: WF(CLASH), Eq. (4). To the right, WF(a), Eq. (5). The case WF=0 corresponds to the data not used for training.
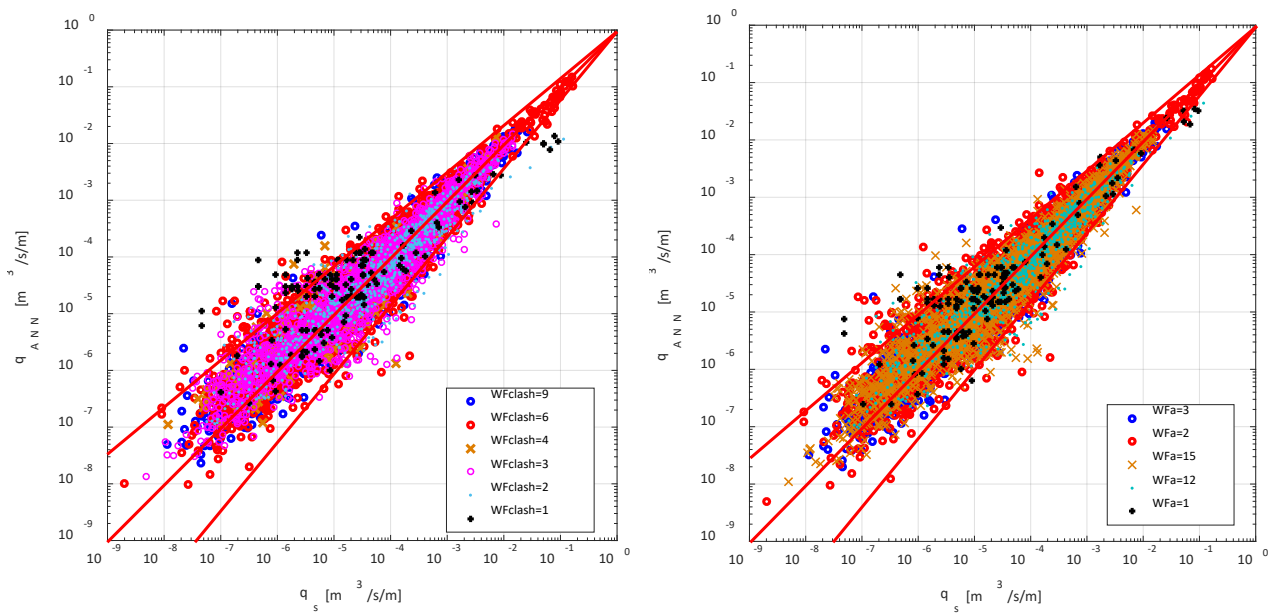

Figure 7 – Comparison of the predicted q_{ANN} and experimental q_s values with WF(CLASH) as defined in Eq. (4), to the left, and WF(a) as defined in Eq. (5), to the right. The corresponding 95% confidence bands are also shown.

Table 4 – Predictions exceeding the 95% confidence bands and belonging to large errors for each class of the WFs evaluated following Eq. (4) or Eq. (5).

WF(CLASH), Eq. (4)				WF(a), Eq. (5)			
CLASS		Exceedance of the 95% confidence bands	Large errors	CLASS		Exceedance of the 95% confidence bands	Large errors
Value	#			Value	#		
9		7.1%	21.9%	3		14.8%	23.9%
6		16.7%	35.0%	2		29.6%	33.3%
4		7.1%	2.9%	1.5		25.9%	22.6%
3		19.0%	20.3%	1.2		11.1%	3.3%
2		14.3%	3.6%	1		18.5%	1.6%
1		35.7%	1.0%				
TOT #		42 (=100%)	306 (=100%)			27 (=100%)	243 (=100%)

3.5   A classifier-quantifiers scheme

To further increase the ANN accuracy for extreme values of q, the idea was to develop a classifier–quantifier scheme, following the work proposed by Verhaeghe et al. (2008). The use of a classifier-quantifier means that the solution is reached into two steps. A first ANN, the classifier, provides the preliminary prediction of the data and then the data are either discarded or predicted again by another ANN, the quantifier.

In the approach proposed by Verhaeghe et al. (2008), the goal was to identify and discard from predictions the values of $q<10^{-6}$ m^3/s/m. The classifier consisted of an ANN trained on the complete CLASH DB and producing a logical value, being 0 in case of $q<10^{-6}$ m^3/s/m and 1 otherwise. In case the results of the classifier were equal to 0, a warning was issued to the user that the case corresponded to no overtopping. In the cases of $q>10^{-6}$ m^3/s/m, the data were then quantitatively predicted by the ANN. Such a Boolean network of course cannot not lead to a continuous prediction.

The scheme proposed in this paper is novel and consists of a quantitative classifier and three quantifiers. The conceptual scheme is given in Figures 8 and 9. The classifier is the ANN trained on the whole DB for $q>0$, and therefore it is used to provide a quantitative value instead of a «logical» value. Based on the output value, the case is then processed by one of the three quantifiers, which are similar ANNs, but trained on different datasets of the same DB. The reason to have more than one quantifier was to use ANNs specifically trained on extremely high and low values and on average values, in order to see whether a better performance would be achieved. The definition of the Classifier - Quantifiers, the CQ scheme hereafter, is given in Figure 8, while the process of estimation of the discharge is shown in Figure 9.

Figure 10 combines the results of the three ANNs ('the quantifiers'), showing the overall performance of the CQ scheme against the total DB. The results of the three ANNs, trained on the specific datasets when predicting again these datasets, are given individually in Fig. 11. Table 3 reports the results of the optimised ANN ('the classifier - C') and of the CQ scheme. Rows 4 and 5 compare the overall performance of the C and CQ schemes, while Rows 6-11 compare the performance of the two schemes for each specific dataset. The main conclusion from Table 3 is that the performance indices show better results for the (single) ANN (i.e., the C scheme) trained on the total DB with the exception of the higher values of q (i.e. the indices for the CQ scheme in row 7 give greater accuracy than the C scheme in row 6). The table also shows that the overall performance of both the C and CQ schemes is affected by the predictions of the lower values of q: the rmse values of 0.075 and 0.076, associated to the dataset of $q \leq 10^{-6}$, are sensible greater than the rmse values of 0.047 and 0.048, associated to $q > 0$. The main improvement derived from the use of the CQ is detected for the dataset of $q \geq 10^{-3}$, especially in terms of rmse. As for the remaining two datasets, there is no significant difference between the two schemes.

One ANN gives a continuous prediction of q if a parameter is changed over a certain range. The use of more than one ANN in the CQ scheme raises the issue whether a continuous prediction is obtained from one ANN to another. Indeed in Figure 10 it can be observed some scatter of the predictions next to the transition points, $q = 10^{-3}$ and 10^{-6} m³/s/m (see Figure 8). To the purpose of a continuity check, Figure 12 compares the predictions from both C and CQ schemes for the same experiment with varying relative crest freeboard $R_c/H_{m0,t}$, and by keeping all the other geometrical and hydraulic conditions constant. A discontinuity in the prediction is clearly visible with the CQ scheme.

The overall conclusion is that the CQ scheme does not really improve the ANN performance, while the complexity of the ANN architecture increases and undesirable discontinuities in predictions are also obtained.
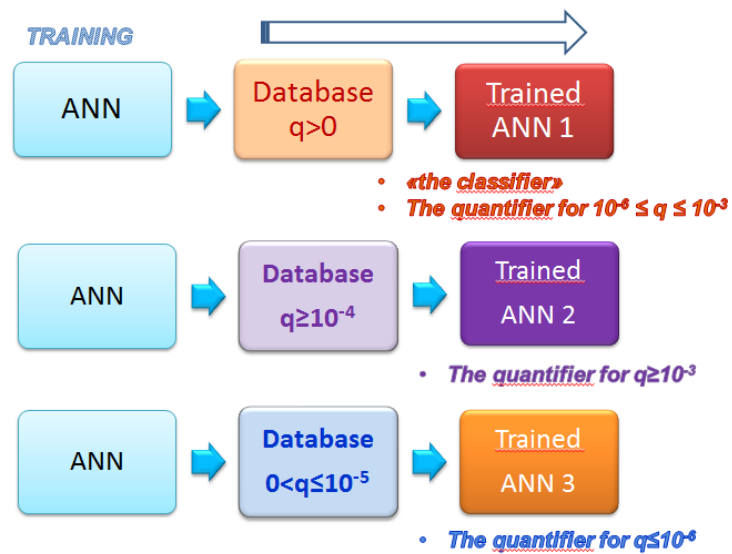
Figure 8 – Structure of the CQ scheme. The architecture of each ANN is the same, while the training datasets are different, leading to three different ANNs. The Trained ANN 1 is both the classifier and the quantifier for the non-extreme values of the discharge.
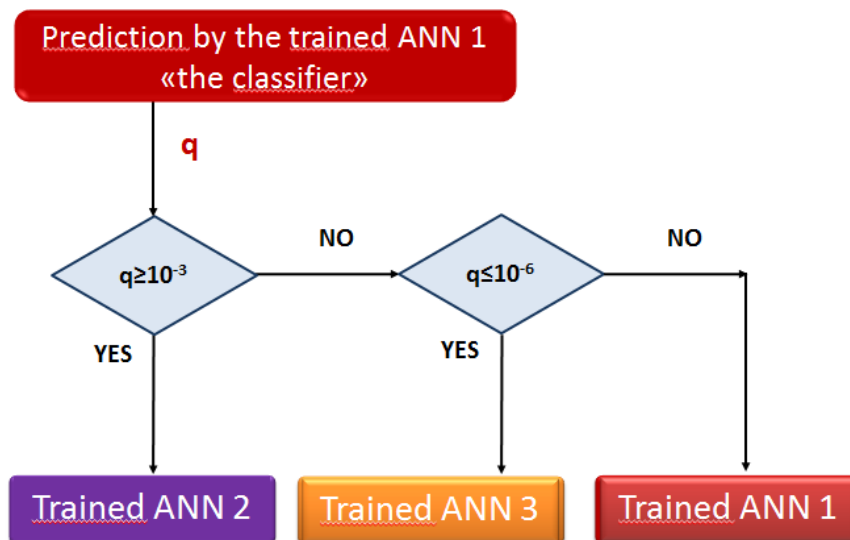

Figure 9 – How the CQ scheme works. The classifier provides a first prediction of the discharge, based on which the test is then processed by one of the three trained ANNs for the final output prediction.
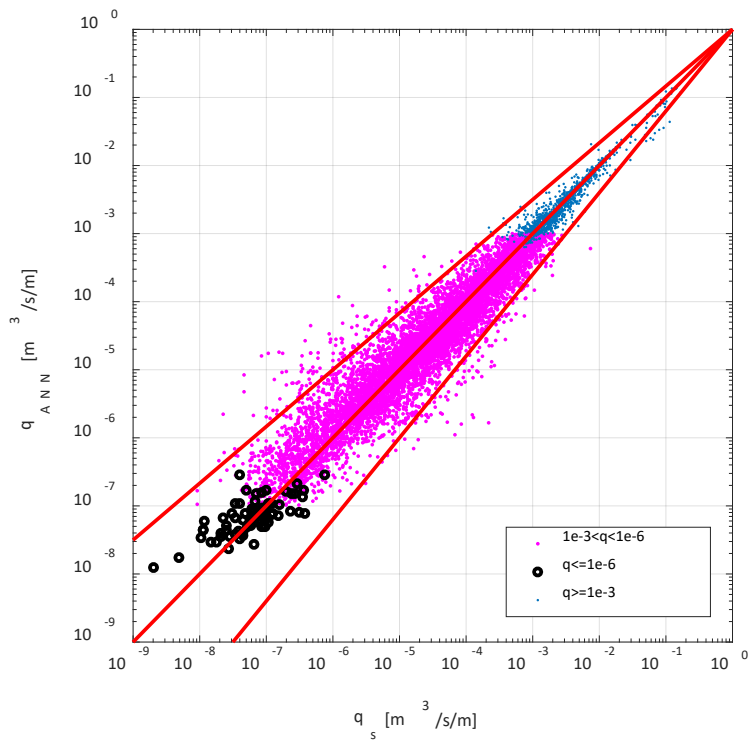
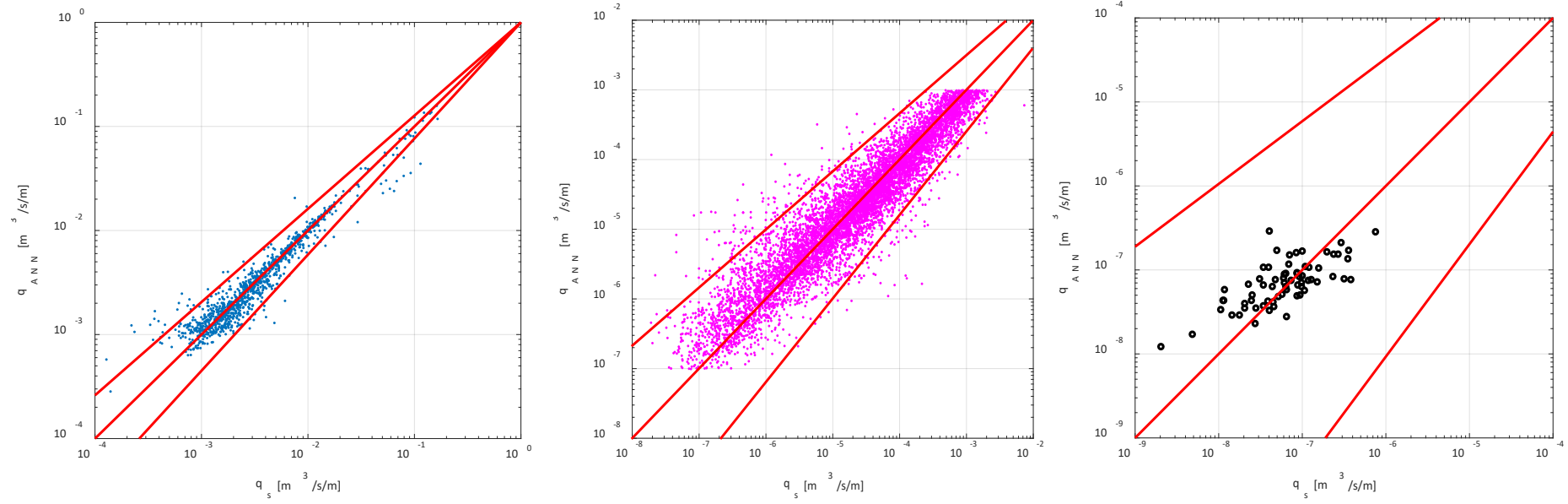Figure 10 – Performance of the CQ scheme (different colours), predicting the total DB.

Figure 11 – Performance of the three quantifiers against the same datasets they have been trained on: Trained ANN 2, to the left; Trained ANN 1 ('the classifier), at the centre; Trained ANN 3, to the right.
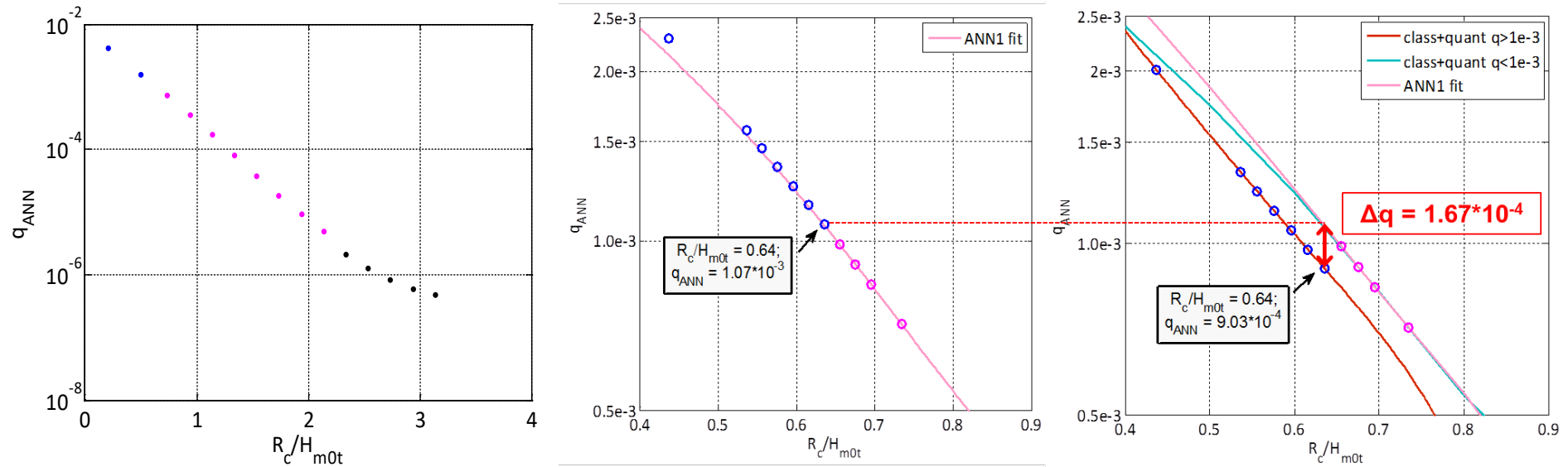
Figure 12 – Variation of the predicted discharge q_{ANN} as a function only of the relative crest freeboard R_c/H_{m0t}, keeping constant all the other geometrical and hydraulic conditions. To the left, the predictions of the CQ for the whole range of R_c/H_{m0t}; to the centre and to the right, zoomed plot of the predictions provided by the Trained ANN 1 only (i.e. the classifier) and by the CQ respectively for $R_c/H_{m0t}=0.64$, i.e. at the boundary between the Trained ANN 1 and 2 ($q=10^{-3}$). The indication Δq gives the difference of the prediction of q obtained with the C and with the CQ schemes. The experimental starting test is a breakwater with large rocks under oblique (45°) attack, Lykke Andersen and Burcharth (2004).

4. Test of the prediction capability of the final ANN

This Section aims at testing the final ANN (whose characteristics are synthesised in Sub-Section 4.1) when predicting new data and datasets (Sub-Sections 4.2 and 4.3) and at discussing the existing limitations of the results (Sub-Section 4.4).

4.1 Main characteristics and performance of the final ANN against the training database

The analyses carried out in Sections 3 demonstrated that the best performance with the ANN is achieved (see Tab. 3) if it is:

- characterized by the simple architecture (the C instead of the CQ scheme);
- trained on the whole values of $q>0$
- trained including WF(a) – defined in Eq. (5).

The quantitative performance of this final ANN has been already provided in terms of error indices in Table 3 (row 4). Figure 13 illustrates the distribution of the predicted values of q in comparison to the corresponding measured ones. In this Figure the values of the WF are not highlighted (as instead in Fig. 7-right), in order to ease the focus on the higher degree of symmetry of the distribution and the reduced scatter achieved by the new ANN with respect to the original one (see Fig. 4-left). Figure 14 shows the average error distribution after 500 BS of the final ANN and its approximation by the normal Probability Density Function (pdf), as suggested by Van Gent et al. (2007). The histogram is overall following the normal distribution, with zero average, indicating that the predictions are symmetrical, i.e. not biased. There are evident deviations in the representation of the peak value (corresponding to zero error) and of the tails that tend to zero more slowly than the pdf. The normal pdf provides a cautious estimate of the values of q within an interval of approximately 2 standard deviations around the average (indicated as 2σ in the Figure) and non-conservative estimates in correspondence of the extreme percentiles (i.e. beyond 2σ). In the application of the predicted results, provided by the ANN-tool, the 2.5% percentiles computed directly from the ANN error distribution are used to give the 95% confidence band around the mean prediction.
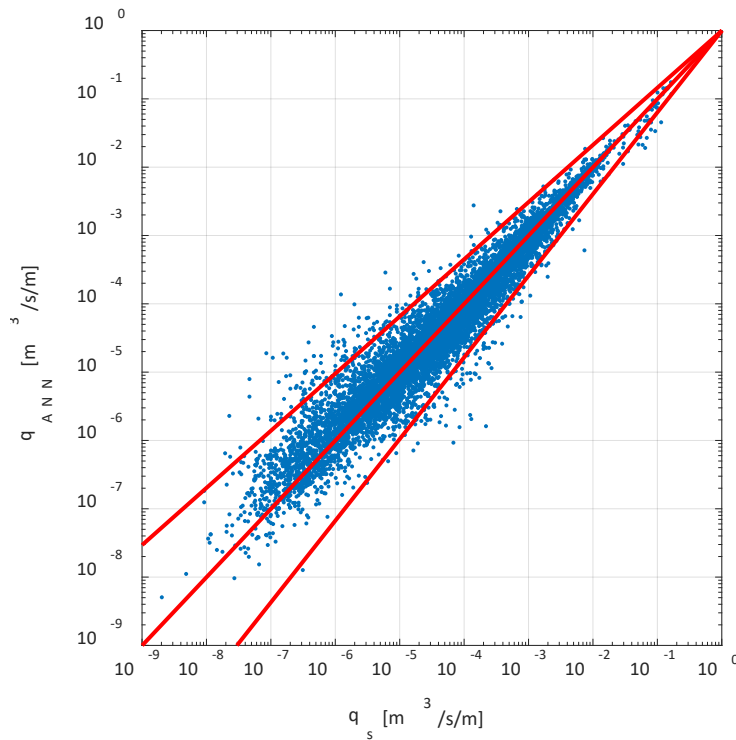
Figure 13 – Comparison of the experimental q_s and the predicted q_{ANN} from the final ANN, C scheme, WF(a) as defined in Eq. (5). The corresponding 95% confidence bands are also shown.
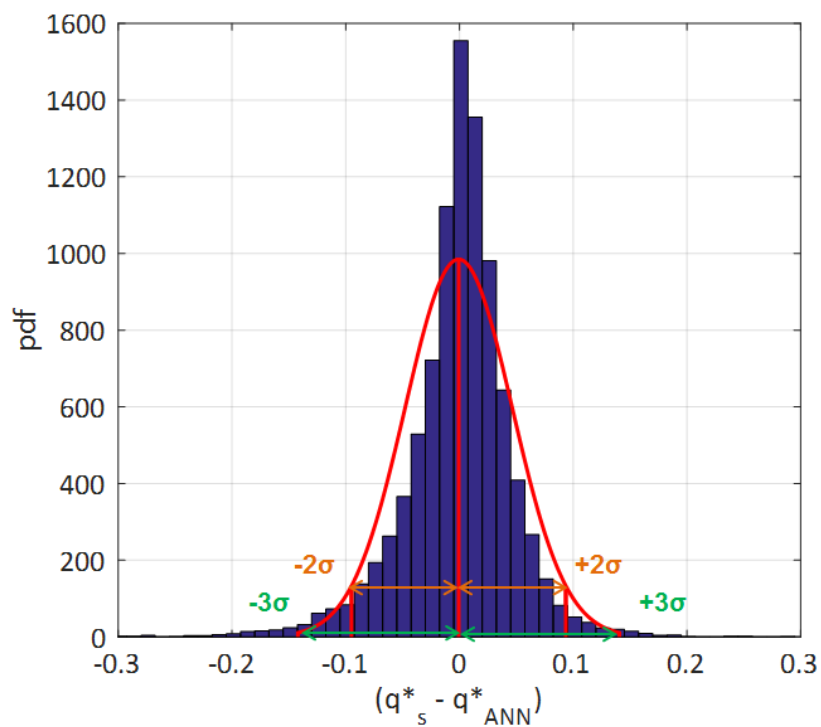

Figure 14 – Frequency error distribution associated to the average predictions provided by the final ANN for the same total DB used for training. The data are stored in 50 equally spaced bars. The continuous red line represents the approximating normal pdf curve.

4.2 Prediction of datasets excluded from the training database

Till now the selection of data for training was based on the total of available tests. The database makes distinction between datasets and tests. A dataset is a number of tests that were performed in the same testing facility and usually on similar structures under similar wave attacks. If only part of the data included in a dataset are eliminated, it may be assumed that still each dataset was present to some extent in training. But what will happen if a complete dataset is excluded from training?

For this purpose, some datasets have been selected from the total DB and excluded from the training process. They have been used exclusively in prediction. The selection process considered to check all the types of structures in the database (i.e. by selecting datasets from all the groups A-F) and to include typical cross-sections for each group. These datasets are listed in the following:

- 166 tests on rock permeable structures (73 from Pearson et al., 2004 and 93 from Helgason et al., 2000; group "A");
- 245 tests on permeable rubble mound structures, composed by different armour units (26 accropods, 29 antifers, 29 tetrapods, 25 Xblocs, 23 corelocs, 113 cubes, from Pearson et al., 2004; group "C");
- 166 tests on smooth dikes (18 from Pearson et al., 2004 and 148 from confidential sources; group "D");
- 25 tests on composite dikes with toe protection of rocks (confidential source; group "E");
- 34 tests on antifer berms (confidential source; group "E");
- 34 tests on vertical walls with recurved crown parapet and rock revetment (Owen and Steel, 1991; group "F").

The number of tests selected within "E" and "F" groups is relatively limited as most of the datasets within these groups are typically large (i.e. hundreds of data) and their removal might have affected the diversity of the training database. In total 670 tests were excluded from training, covering 8 full datasets. This means that the ANN was trained on 8,633 tests instead of the total DB of 9,303 tests.

Of course, every time the ANN is trained on another database it will become a new ANN. Therefore the results presented in this Sub-section do not represent, strictly speaking, the test of the accuracy of the same final ANN presented in Sub-section 4.1. Indeed this is a new ANN, with the same architecture as the final ANN but trained on a database that is the 7% narrower than the one used for the final ANN.

Figure 15 shows the measured and predicted values of q for these datasets (highlighted with different colours). Most of the predictions fall within the 95% confidence bands (i.e. the confidence bands defined based on the final ANN trained on the total DB). The datasets in group E with Antifer berms tend to be systematically overestimated. A few outliers can be observed in the prediction of the dataset in group D with smooth dikes.

The performance indexes associated to the ANN predictions of the datasets are collected in Table 5. With the exception of the four datasets belonging to groups "E" and "F" and to the smooth dikes "D" (rows 9, 11, 13, 14), the accuracy of the predictions is fair, since the values of the performance indexes (rmse ≈ [0.038;0.056], IW ≈ [0.91;0.97], R^2≈[0.75;0.93]) are comparable to the ones characterizing the average performance of the "final" ANN (Tab. 3, row 4). As expected, the ANN better represents those datasets (e.g., the tetrapods from Pearson et al., 2004) that are similar to other tests included in the training DB. It can be concluded that 4 of the 8 datasets are very well predicted, while all the 8 datasets are predicted mostly within the 95% confidence band.

Table 5 – Quantitative performance of the "final" ANN when predicting datasets excluded from training. The error index is not provided when the computed standard deviation is of the same order of magnitude of the average value of the index itself.

|	Dataset	Armour type (DB group)	Rmse [m³/s/m]	WI [-]	R^2 [-]	Large errors [%]
1	Helgason et al. (2000)	Rocks (A)	0.056±0.05	0.93±0.04	0.75±0.07	0
2	Pearson et al. (2004)	Rocks (A)	0.040±0.05	0.94±0.09	0.89±0.1	0
3	Pearson et al. (2004)	Accropods (C)	0.044±0.02	0.97±0.04	0.86±0.2	0
4	Pearson et al. (2004)	Antifers (C)	0.049±0.02	0.96±0.04	0.85±0.2	0
5	Pearson et al. (2004)	Core-locs (C)	0.054±0.02	0.91±0.05	0.79±0.2	0
6	Pearson et al. (2004)	Cubes (C)	0.045±0.02	0.97±0.05	0.88±0.2	1.7
7	Pearson et al. (2004)	Tetrapods (C)	0.050±0.04	0.92±0.1	0.80±0.6	0
8	Pearson et al. (2004)	X-blocs (C)	0.038±0.06	0.95±0.1	0.93±0.5	0
9	-	Smooth dike (D)	0.059±0.05	0.86±0.04	0.66±0.2	2.3
10	Pearson et al. (2004)	Smooth dike (D)	0.041±0.02	0.96±0.09	0.87±0.3	0
11	-	Dike with rock protection (E)	0.08±0.1	0.69±0.20	-	0
12	-	Antifer berm (E)	0.10±0.05	0.59±0.10	-	0
13	Owen & Steel (1991)	Wall with rock protection (F)	0.071±0.040	0.81±0.03	0.85±0.08	0
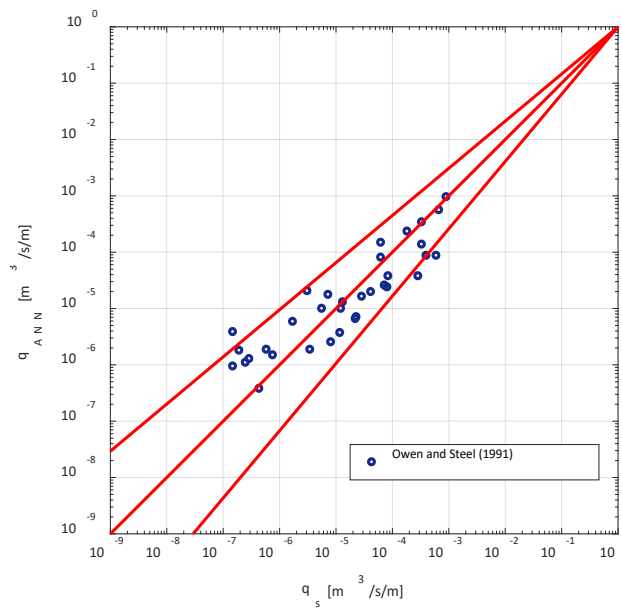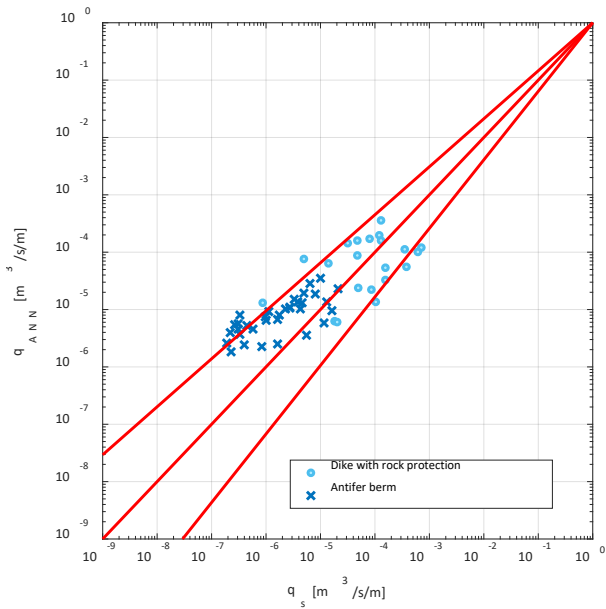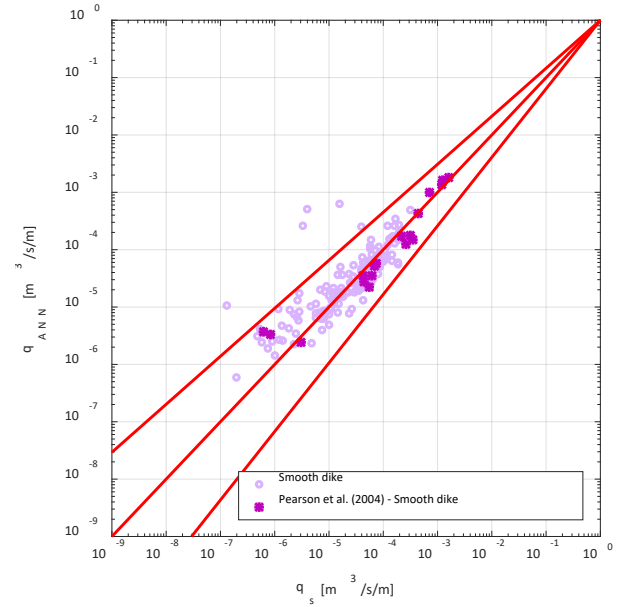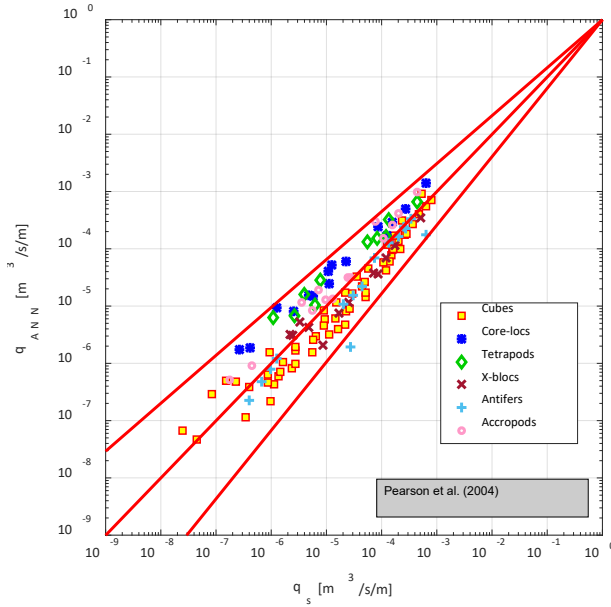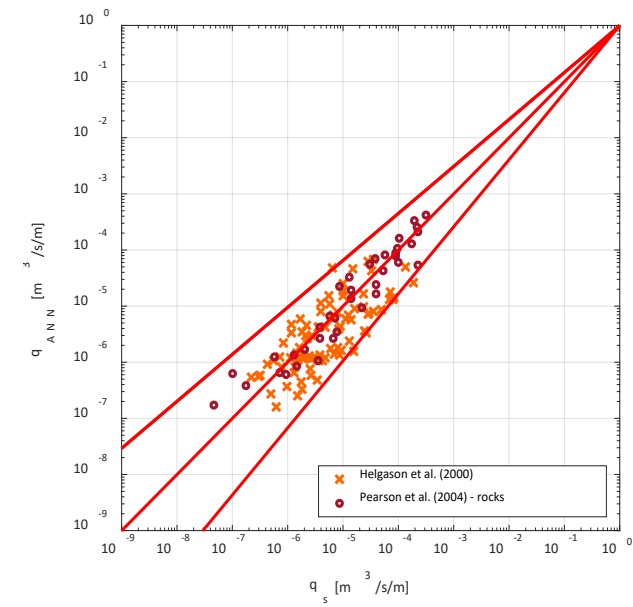
Figure 15 – Predictions of selected datasets that have been excluded from the total DB in the training process.

4.3   Prediction of new data and datasets

The previous Sub-section 4.2 showed that datasets not used in training may be predicted quite well if the tested conditions and structure geometry are to some extent familiar to the DB. Other datasets that are less represented may show quite large deviations in predictions. One disadvantage of the procedure in the previous sub-section is that by excluding datasets, the resulting ANN is trained on less data.

To overcome this disadvantage, the final ANN, trained on the total DB, has been used to predict additional data that were not yet included in the new DB.
These data/datasets have been used here exclusively to validate the accuracy and uncertainty of the final ANN when used for prediction of completely new datasets. These data are listed below and consist of 957 tests.

- New data: 118 tests on smooth slopes with crown walls of different heights, placed at different distance from the offshore edge of the structure crest. These data belong to a larger dataset of 645 tests (Oumeraci et al., 2004), including also double crown walls, crown walls characterized by recurve or oblique parapets and/or still water basins. Some of these geometries cannot be schematized for the DB.
- New data: 613 tests on reshaping berm breakwaters (Lykke Andersen, 2006). These data also belong to a larger dataset of 695 tests, which is not completely new to the database and ANN as 82 tests were already in the CLASH DB and are included in the new DB (group "E"). However these data have been excluded from the training DB as they derive all from the same laboratory and therefore they might significantly affect (and maybe bias) the ANN predictions for berm breakwaters.
- New datasets: 140 tests from different projects on hardly and fully reshaping berm breakwaters, under perpendicular and oblique waves (private communication).

The results of the predictions of the new datasets and data are reported in Tab. 6 in terms of performance indexes, and are shown in Figure 16 in comparison to the experimental measurements. In most cases, the additional data by Lykke Andersen (2006) and by Oumeraci et al. (2004) are well predicted by the ANN, falling within the same confidence bands as the ANN predictions of the total DB. In the case of Lykke Andersen (2006), the outliers correspond to emerged berms (h_b/H_{m0t}<-0.5), which indeed are under-represented in the training DB as they do correspond either to rock revetments in front of vertical walls or to smooth dikes, not to berm breakwaters. The new hardly and fully reshaping berm breakwater dataset is accurately predicted by the ANN. Only in case of one dataset (named "Project X" in Fig. 16) the values of q_{ANN} tend to sistematically overestimate the corresponding q_s. In this project, the structure has a working road behind the permeable crest and the measurement was taken at the end of this working road, directly to the transition of the rear slope. It is a very wide and permeable structure. Datasets with such a wide structure with water percolation along the rubble mound berm, crest and working road are not present in the total DB. Therefore the real overtopping values result in much lower values of q_s than expected on the basis of the structures included

in the training DB. The overestimation of the ANN can be explained, but the prediction were far from reality. Moreover, all these data are characterized by high emerged hardly reshaping berm (-2.33<h_b/H_{m0t}<-0.51) and this may also have decreased the overtopping discharge.

As expected, the overall conclusion is that if wave conditions and structure geometry are to some extent in the training DB, a new dataset will be predicted quite well by the ANN. If this is not the case and the structure geometry and/or wave conditions are really out of the experience of the very large DB, the prediction might also be quite far off the reality.
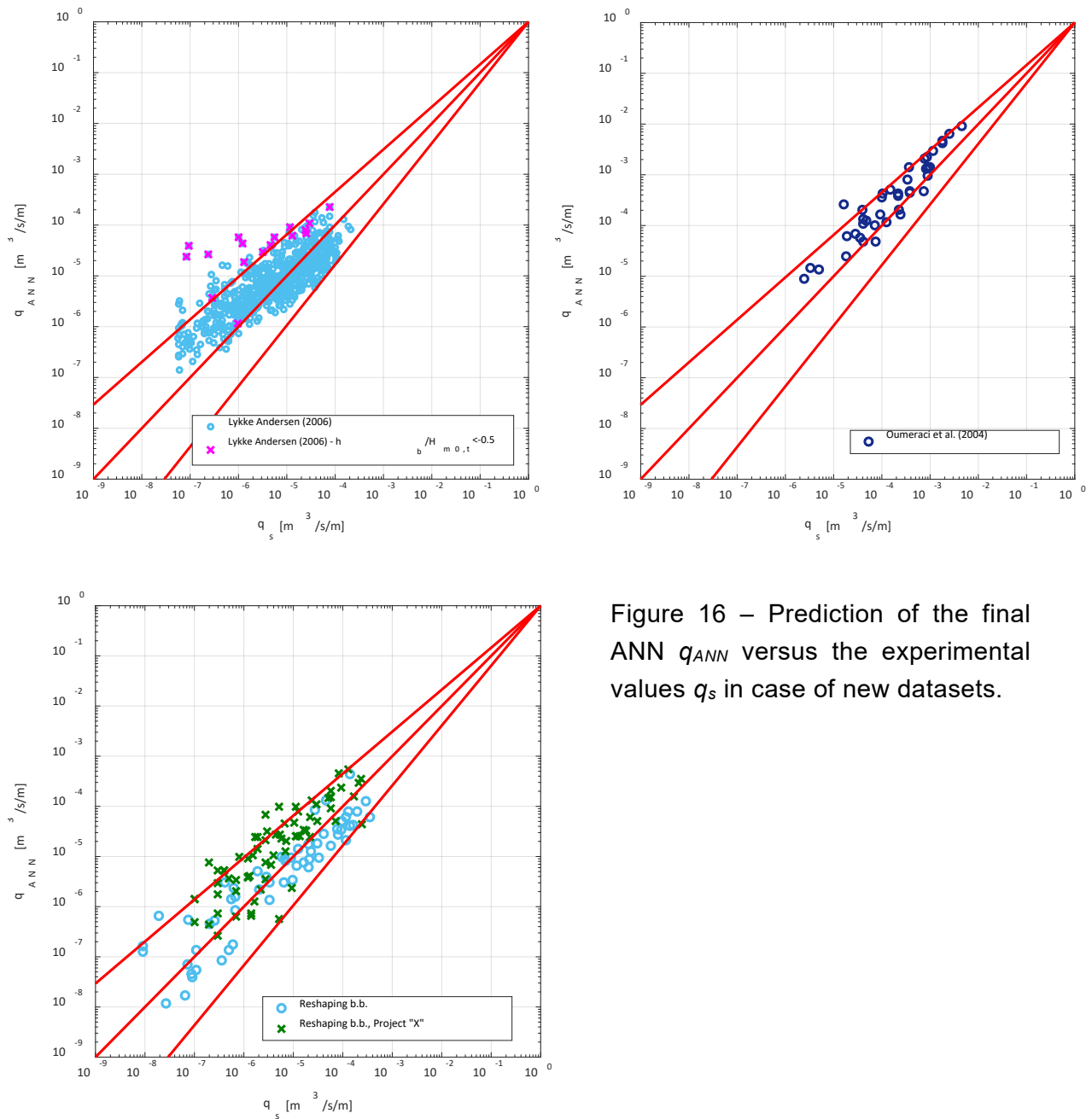

Figure 16 – Prediction of the final ANN q_{ANN} versus the experimental values q_s in case of new datasets.

Table 6 – Quantitative performance of the "final" ANN when predicting new data not included in the training DB. The error index is not provided when the computed standard deviation is of the same order of magnitude of the average value of the index itself.

Dataset	Armour type (DB group)	Rmse [m³/s/m]	WI [-]	R² [-]	Large errors [%]
Oumeraci et al. (2004)	Composite smooth slopes (D)	0.057±0.007	0.93±0.02	0.72±0.08	0
Lykke Andersen (2006)	Reshaping BB (E)	0.074±0.06	0.82±0.1	0.52	28
-	Reshaping BB, 3D (G)	0.069±0.01	0.94±0.02	0.81±0.07	0
"Project X"	Reshaping BB, 3D (G)	0.09±0.04	0.80±0.10	-	0

4.4   Scale and model effects

The designers need to have estimates of the structure performance at prototype scale. However, the training DB of the ANN includes only data derived from laboratory tests, both in wave flumes and tanks, performed at very different scales. A few (around 100) prototype data for smooth and permeable structures (Yamamoto and Horikawa, 1992; Pullen et al., 2004; Briganti et al., 2005; De Rouck et al., 2005) are available but they have not been included in the training DB as De Rouck et al. (2005) and Franco et al. (2009) have already proved scale and model effects inherent to the laboratory conditions. These effects are mainly related to the structure properties (porosity, permeability, roughness) and to the hydraulic loads (wind, spray, currents). Based on EurOtop (2007), permeable structures are affected by both model and scale effects, resulting in wave overtopping discharges that are lower in model tests than in prototype. Impermeable structures instead are affected by model effects only, giving wave overtopping discharges in model tests that are generally higher than in prototype. Therefore the maximum deviation of the ANN predictions is expected in case of rubble mound structures armoured by rocks or units. However also in case of smooth structures the difference of the hydraulic loads in the lab and in the prototype holds.

Figure 17 shows the comparison of the ANN predictions for 3 cases available in both laboratory and prototype scales: Samphire Hoe (Pullen et al., 2004), a smooth vertical wall; Zeebrugge (De Rouck et al., 2005), a rubble mound structure with Antifer cubes; Ostia (Briganti et al., 2005), a rock permeable breakwater. In agreement with the indications already provided by EurOtop (2007), the values of q are underestimated for the rock permeable breakwater whereas tend to be overestimated for the impermeable wall. The ANN systematically underestimates the overtopping discharge for Ostia, while the predictions for Zeebrugge still fall within the confidence bands. However also the two different structure slopes should be taken into account, being $cota_d$=1.4 for Zeebrugge and $cota_d$=4 for Ostia. Even if these effects are well known for experimental cases it was not obvious that the ANN performed similarly.

Eurotop (2007) proposed an empirically based approach for the re-scaling of the experimental values to prototype conditions, distinguishing between permeable and impermeable structures (see Sub-sections 6.3.6 and 7.3.6 of the manual respectively). This method consists of first re-scaling to prototype the model data (for permeable structures only) and then of applying different correction factors derived from a fitting process and accounting for the effects of wind, of the structure slope α_d and of the roughness γ_f. The corrections are applied only when the discharge exceeds 1 l/s/m.

Figure 18 shows the comparison among the ANN predictions, the ANN predictions corrected following Eurotop (2007) and the experimental values. The ANN predictions have not been previously re-scaled since the ANN is trained on data characterised by a variety of scales, i.e. the predictions are assumed not to be scale-dependent. In case of rubble mound structures, the inclusion of the effect of wind (since it is a constant factor) contributes to a systematic increase of the predictions, while in case of smooth structures significantly improves the predictions of the low values of q. Overall the re-scaled predictions based on Eurotop lead to a cautious design of coastal and harbour structures.
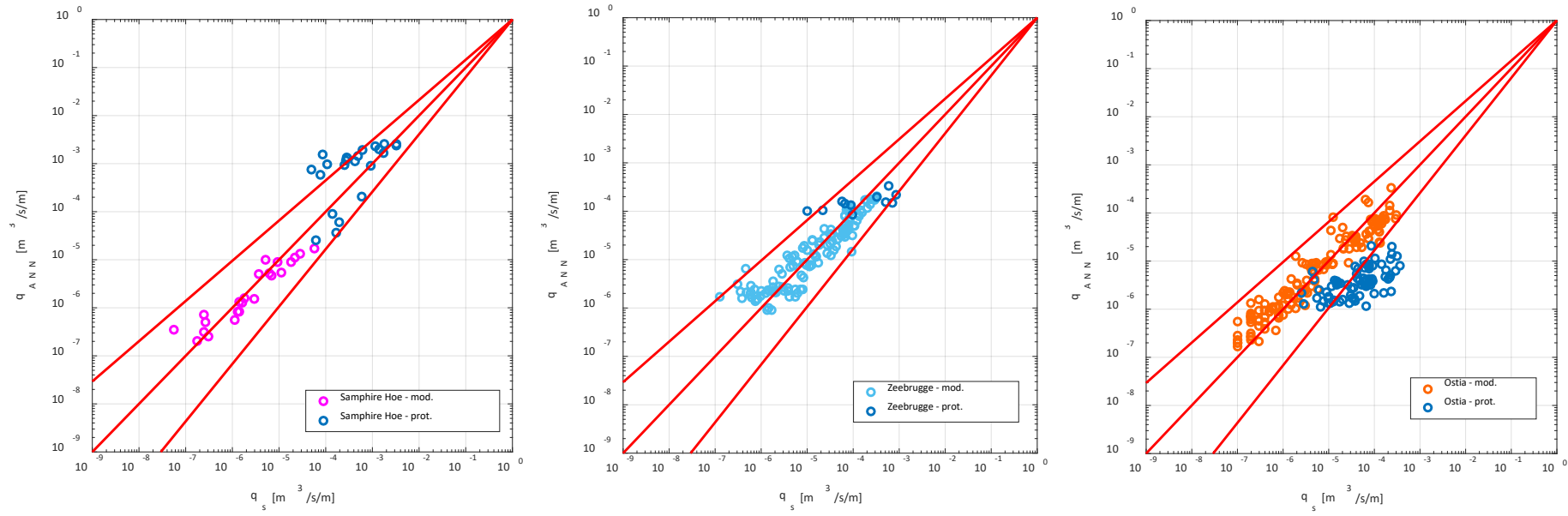
Figure 17 – Comparison of the predicted q_{ANN} and experimental q_S values. From left to right: Samphire Hoe, smooth impermeable structure; Zeebrugge, rubble mound with cubes; Ostia, rubble mound breakwater.
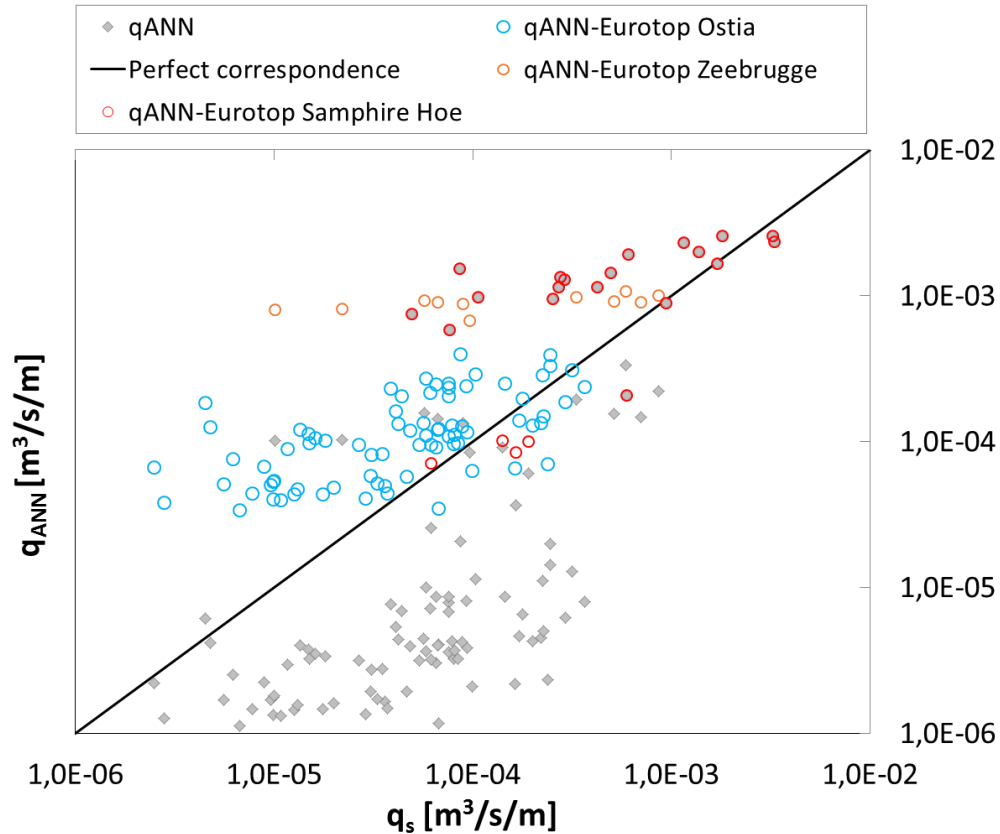
Figure 18. Outputs of the ANN (grey diamonds) and re-scaled outputs of the ANN including the EurOtop (2007) corrections without wind (void circles) versus the measured prototype values.

5. Conclusions

This paper presented an optimized ANN for the prediction of the overtopping discharge q under a variety of complex structure geometries and wave conditions. The optimization process of the ANN started from the existing ANN developed by the authors for the prediction of the wave overtopping discharge, and of the wave reflection and transmission coefficients. It included the new extended database, based on CLASH (2004) and consisting now of more than 13,000 data for wave overtopping only.

The optimized ANN consists of a multi-layer network, with 15 dimensionless input parameters, 20 hidden neurons and 1 output parameter, the average overtopping discharge q. It is trained on the extended database set-up by the authors (Zanuttigh et al., 2014), by including all the laboratory-scale non-zero values of q. This is different from previous work (Van Gent et al., 2007; Verhaeghe et al., 2008), which limited the training of ANN to $q>10^{-6}$ $m^3/s/m$, with a significant bias of the ANN predictions for small overtopping. This choice led to a significant reduction of the bias of the ANN predictions in case of low values of q, i.e. the overestimation for $q \leq 10^{-6}$ $m^3/s/m$.

The ANN cannot predict the measured zero values of q as exact zeros, i.e. the same fixed output for different input conditions. It provides non-zero predictions that dataset by dataset are generally in agreement with the expectations, i.e. always lower than the predictions of the other non-zero values and following a similar trend with the most relevant input parameters. Zero overtopping in laboratory may be strongly affected by the system adopted to measure the overtopping (e.g., a wave gauge in a box, a weighting system or propellers on top of the crest).

A different formulation of the Weight Factors to be used in the training process has been proposed here and tested, starting from the definition suggested by CLASH (2004). The alternative definition of the Weight Factors provided in Eq. (5) aims at equally considering both the Reliability RF and Complexity Factors CF and leads to a better ANN overall performance.

Specifically to improve the representation of the extreme values of q, a classifier-quantifiers scheme was implemented and verified. The scheme consists of a quantitative classifier, which is the ANN trained on the total database, and three classifiers, which consist of the same ANN architecture, but trained on three datasets from the total database, separating the prediction of high, average and low values of q. Based on the first prediction from the classifier, the input is processed by one of the three quantifiers, giving the final prediction. While the prediction of the high values of q slightly improves, the overall ANN performance decreases due to a worsening of the prediction of the low values of q, leading to the conclusion to keep the ANN architecture simple, i.e. by using the classifier only.

The performance of the optimised ANN is satisfactorily accurate (being the average values of the rmse=0.047, IW=0.977, and R^2=0.92). The ANN error distribution cannot really be approximated by a normal pdf and therefore all the output percentiles of q and the 95% confidence bands should be derived from the actual ANN error distribution (EurOtop, 2007).

The accuracy of the ANN has been tested by predicting data and datasets, either excluded from the total training database or completely new, i.e. not included yet in the existing wave overtopping database. In most cases the predictions fall within the 95% confidence bands and show the same dispersion as the predictions of the training database.

Finally, It is worthy to remark that the training of the ANN was performed on model tests only, disregarding the few available prototype data. Therefore the ANN predictions of real prototype cases are non-conservative for permeable structures. However, cautious estimates can be obtained by applying to the ANN predictions the corrections factors by Eurotop (2007).

Acknowledgments

THESEUS project ("Innovative technologies for safer European coasts in a changing climate"), www.theseusproject.eu.

References

Allsop, W., Bruce T., Pullen T. and van der Meer, J., 2008. Direct hazards from wave overtopping – the forgotten aspect of coastal flood risk assessment?, 43rd Defra Flood and Coastal Management Conference, 1-3 July 2008, Manchester University.

Besley P., Reeves M. and Allsop N.W.H., 1993. Random wave physical model tests: overtopping and reflection performance, Report IT 384, HR Wallingford.

Briganti, R., Bellotti, G., Franco L., De Rouck, J. and Geeraerts, J., 2005. Field measurements of wave overtopping at the rubble mound breakwater of Rome-Ostia yacht harbour, Coastal Engineering 52, 1155-1174.

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.

De Rouck J., Geeraerts J., Troch P., Kortenhaus A., Pullen T. and L. Franco, 2005. New Results on Scale Effects for Wave Overtopping at Coastal Structures, Proc. Of the International Conference on Coastlines, Structures and Breakwaters 2005, 29-44.

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.

Franco, L.,Geeraerts, J.,Briganti, R.,Willems, M.,Bellotti, G. and De Rouck, J. (2009). "Prototype measurements and small-scale model tests of wave overtopping at shallow rubble mound breakwaters: the Ostia-Rome yacht harbour case". Coastal Engineering, 56, 154-165.

Hagan, M.T., and Menhaj M., 1994. Training feed-forward networks with the Marquardt algorithm, IEEE Transactions on Neural Networks, 5(6), 989-993.

Heimbaugh, M. S., Grace, P. J., Ahrens, J. P., and Davidson, D. D. 1988 (Mar). "Coastal Engineering Studies in Support of Virginia Beach, Virginia, Beach Erosion Control And Hurricane Protection Project; Report No. 1: Physical Model Tests Of Irregular Wave Overtopping And Pressure Measurements," Technical Report CERC-88-1, US Army Engineer Waterways Experiment Station, Vicksburg, MS.

Helgason, E., Burcharth, H., and Beck, J. (2001) Stability of Rubble Mound Breakwaters Using High Density Rock. Coastal Engineering 2000: pp. 1935-1945. doi: 10.1061/40549(276)151

Higuera P., Lara J. L. and I. J. Losada. 2013. Simulating coastal engineering processes with OpenFOAM (R), Coastal Engineering; 71,119-134.

Lykke Andersen, T. and Burcharth, H.F., 2004. CLASH Work package 4.4 - D24 Report on additional tests. Part D Berm breakwater tests. On-line available.

Lykke Andersen, T. 2006. Hydraulic Response of Rubble Mound Breakwaters. Scale Effects – Berm Breakwaters. PhD Thesis, Series Paper No. 27, Dept. of Civil Eng., Aalborg University.

Lykke Andersen, T., Skals, K. T. and Burcharth, H. F. 2008. Comparison of homogenous and multi-layered berm breakwaters with respect to overtopping and front slope stability. ASCE, proc. ICCE 2008.

Marquardt, D., 1963. An algorithm for least-squares estimation of nonlinear parameters, SIAM Journal on Applied Mathematics 11 (2): 431–441. doi:10.1137/0111030.

Numata, A., 1976. Laboratory formulation for transmission and reflection at a permeable breakwater of artificial blocks. Coastal Engineering in Japan 19, 47–58.

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.

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.

Owen, M.W., and Steele, A.A.J., 1991. Effectiveness of Recurved Wave Return Walls, HR Wallingford, Report SR 26 1.

Pearson, J., Bruce, T., Franco, L. And Van der Meer, J.W, 2004. Report on additional tests, part B, CLASH WP4 report, University of Edinburgh, Edinburgh, UK.

Pullen, T., Allsop, N.W.H., Bruce, T., Pearson, J. and Geeraerts, J., 2004. Violent wave overtopping at Samphire Hoe: field and laboratory measurements. Proc. 29th Int. Conf. Coastal Engineering. pp.4379–4390 (ASCE) World Scientific, Singapore, ISBN 981-256-298-2.

THESEUS team, 2014. Coastal risk management in a changing climate, Zanuttigh B., Nicholls R., Vanderlinden J. P., Burcharth H. F. & R. C. Thompson editors, Elsevier, 671 pp., ISBN: 978-0-12-397310-8.

Van der Meer, J.W., Briganti, R., Zanuttigh, B. and B. Wang, 2005. Wave transmission and reflection at low crested structures: design formulae, oblique wave attack and spectral change, *Coastal Eng.*, 52 (10-11), 915-929.

Van der Meer, J.W., T. Pullen, N.W.H. Allsop, T. Bruce, H. Schüttrumpf and A. Kortenhaus, 2009. Prediction of overtopping. Chapter 14 in Handbook of Coastal and Ocean Engineering; Ed. Young C. Kim. World Scientific, 341-382.

Van Doorslaer, K., De Rouck, J., Audenaert, S. and Duquet, V., 2015. Crest modifications to reduce wave overtopping of non-breaking waves over a smooth dike slope, Coastal Engineering, 101, 69-88. doi:10.1016/j.coastaleng.2015.02.004

Verhaeghe, H., 2005. Neural network prediction of wave overtopping at coastal structures, PhD thesis, Universiteit Gent, Gent, BE.

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. and Troch, P., 2012. Wave overtopping at smooth impermeable steep slopes with low crest freeboards, J. Waterway, Port, Coastal, Ocean Engineering, 10.1061/(ASCE)WW.1943-5460.0000141, 372-385.

Yamamoto, Y., and Horikawa, K. 1992. New methods to evaluate run-up height and wave overtopping rate, Proc. 23rd Coastal Engr. Conf, Venice, Italy, ASCE.

Zanuttigh, B. & J. W. van der Meer, 2008. Wave reflection from coastal structures in design conditions, Coastal Engineering, 55 (10), 771-779, Elsevier.

Zanuttigh, B., Formentin, S., & Van der Meer, J. 2014. Advances in modelling wave-structure interaction through artificial neural networks. Coastal Engineering Proceedings, 1(34), structures.69. doi: http://dx.doi.org/10.9753/icce.v34.structures.69