



COMMISSION
OF THE EUROPEAN
COMMUNITIES



FP5- EESD

CREST LEVEL ASSESSMENT OF
COASTAL STRUCTURES BY
FULL-SCALE MONITORING,
NEURAL NETWORK PREDICTION
AND HAZARD ANALYSIS
ON PERMISSIBLE WAVE OVERTOPPING

CLASH

EVK3-CT-2001-00058

R
E
P
O
R
T

D 42 Final report on generic prediction method

December 2005

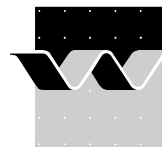
M.R.A. Van Gent
B. Pozueta
H.F.P. Van den Boogaard
J.R. Medina



WL | delft hydraulics



UNIVERSIDAD
POLITECNICA
DE VALENCIA



CLIENT:	Commission of the European Communities, CLASH (EVK3-2001-00058)
TITLE:	Prediction method Neural network modelling of wave overtopping at coastal structures
ABSTRACT:	<p>Within the framework of the European project CLASH phenomena related to wave overtopping are studied. In this report a method to predict wave overtopping at a wide range of coastal structure types is given. The method is based on Neural Network modelling of the mean wave overtopping discharge. The method can be applied for coastal structures such as dikes, rubble mound breakwaters and caisson structures.</p> <p>The predictions by the Neural Network go together with information on the reliability of these predictions (confidence levels). The quality of the predictions strongly depends on the quality of the database on which the Neural Network is based. The data base used here consists of about 8400 test conditions. These data mainly originate from small-scale tests from many different laboratories.</p> <p>The results show that rather accurate predictions can be obtained with the Neural Network.</p>

VER.	ORIGINATORS	DATE	REMARKS	REVIEW	APPROVED BY
3	M.R.A. van Gent B. Pozueta H. v/dn Boogaard	Dec. 2004	WP8-D42	J.R. Medina	W.M.K. Tilmans
KEYWORDS					STATUS
Wave overtopping Coastal Structures Neural network Prediction method Database			Breakwaters Dikes Caissons Reliability Uncertainty		<input type="checkbox"/> PRELIMINARY <input type="checkbox"/> DRAFT <input checked="" type="checkbox"/> FINAL
PROJECT IDENTIFICATION: H3969					

Contents

List of Figures

List of Tables

List of Symbols

1	Introduction.....	1—1
1.1	Background.....	1—1
1.2	Purpose of the present study	1—1
1.3	Outline of the present report	1—2
2	Description of database and parameters involved.....	2—1
2.1	Description of database.....	2—1
2.2	Parameters involved.....	2—1
3	Neural Network modelling.....	3—1
3.1	General.....	3—1
3.2	Multi-layer feed-forward neural networks.....	3—1
3.3	Applied Neural Network simulator.....	3—3
4	Preparation and validation of the NN models.....	4—1
4.1	Preparations of the database used for the NN configuration	4—1
4.1.1	Validation and reduction of database	4—1
4.1.2	Weight Factors	4—2
4.1.3	Froude’s similarity law	4—2
4.1.4	Data distribution	4—3
4.2	NN configuration	4—4
4.2.1	Training and testing	4—4
4.2.2	Training and testing sets based on resampling.....	4—5
4.2.3	Optimum number of neurons in the hidden layer	4—5
4.3	Uncertainty assessment.....	4—6

4.3.1	Bootstrap resampling	4—7
4.3.2	Weight Factors in the bootstrap resampling.....	4—7
4.3.3	Statistics of the predictions: Confidence intervals.....	4—8
4.4	Performance of the NN model	4—8
4.4.1	Classification of overtopping events.....	4—10
5	Sensitivity Analysis	5—1
6	Conclusions	6—1

Acknowledgements

References

List of Figures

In text:

- Figure 1 Parameters used for the NN modelling of wave overtopping discharge at coastal structures
- Figure 2 Example of the configuration of a three-layered NN
- Figure 3 NN performance for configurations with different number of neurons in the hidden layer
- Figure 4 Final NN configuration
- Figure 5 Observations versus NN predictions; application on data $q > 10^{-6} \text{ m}^3/\text{s}/\text{m}$
- Figure 6 Observations versus NN predictions for different Reliability Factors; application on data $q > 10^{-6} \text{ m}^3/\text{s}/\text{m}$
- Figure 7 Observations versus NN predictions
- Figure 8 Observations versus NN predictions for different Reliability Factors
- Figure 9 Observations versus NN predictions; in dimensionless form
- Figure 10 Observations versus NN predictions for different Reliability Factors; in dimensionless form
- Figure 11 Observations versus NN predictions; classification analysis
- Figure 12 Observations versus NN predictions; classification analysis in dimensionless form
- Figure 13 Sensitivity of wave overtopping to relative crest freeboard; Vertical structure
- Figure 14 Sensitivity of wave overtopping to relative armour width; Rubble mound structure
- Figure 15 Sensitivity of wave overtopping to relative crest freeboard; Smooth structure
- Figure 16 Sensitivity of wave overtopping to relative crest freeboard; Rubble mound structure with berm

List of Tables

In text:

- | | |
|---------|---|
| Table 1 | Values of the Weight Factor for different combinations of RF and CF in the database |
| Table 2 | Statistical parameters of the scaled input / output-patterns of the database used for calibration / training of the NN (scaled to $H_{m0} = 1$ m) |
| Table 3 | Results of the classification analysis (Total number of tests in the reduced database, $N_{total} = 8372$) |
| Table 4 | Results of the classification analysis in dimensionless form (Total number of tests in the reduced database, $N_{total} = 8372$) |

List of Symbols

A_c	:	armour crest freeboard of the structure (m)
B	:	width of the berm (m)
B_t	:	width of the toe berm (m)
CF	:	complexity factor (-)
G	:	gravitational acceleration, $g = 9.81 \text{ m}^2/\text{s}$
G_c	:	crest width of the structure (m)
h	:	water depth in front of the structure (m)
h_b	:	water depth on the berm (m)
h_t	:	water depth at the toe of the structure (m)
$H_{m0, \text{toe}}$:	significant wave height from spectral analysis at the toe of the structure, $H_{m0} = 4\sqrt{m_0}$ (m)
q	:	mean overtopping discharge ($\text{m}^3/\text{s}/\text{m}$)
q'	:	mean overtopping discharge scaled with Froude's law to $H_{m0} = 1 \text{ m}$ ($\text{m}^3/\text{s}/\text{m}$)
q_{obs}	:	mean overtopping discharge of the observations/measurements ($\text{m}^3/\text{s}/\text{m}$)
q_{NN}	:	mean overtopping discharge of the NN predictions ($\text{m}^3/\text{s}/\text{m}$)
Q	:	dimensionless overtopping discharge, $Q = \frac{q}{\sqrt{g \cdot H_{m0}^3}}$ (-)
R_c	:	structure crest freeboard (m)
RF	:	reliability factor (-)
$T_{m-1,0, \text{toe}}$:	mean wave period from spectral analysis at the toe of the structure (s)
WF	:	weight factor (-)

Greek letters:

α_d	:	angle of the slope of the structure downward of berm (-)
α_u	:	angle of the slope of the structure upward of berm (-)
α_b	:	angle of the slope of the berm (-)
β	:	direction of wave attack
γ_f	:	roughness/permeability coefficient (-)

I Introduction

The international CLASH project of the European Union (Crest Level Assessment of coastal Structures by full scale monitoring, neural network prediction and Hazard analysis on permissible wave overtopping, www.clash-eu.org) under contract no. EVK3-CT-2001-00058 focuses on the modelling and predictions of wave overtopping for a wide variety of coastal structures, both in prototype and laboratory conditions (De Rouck *et al.*, 2002). The main scientific objectives of CLASH are (i) to analyse scale effects in wave overtopping and (ii) to produce a prediction method based on Neural Network models for the estimation of wave overtopping.

I.1 Background

For the design, safety assessment and rehabilitation of coastal structures reliable predictions of wave overtopping are required. Several design formulae exist for dikes, rubble mound breakwaters and vertical breakwaters. Nevertheless, often no suitable prediction methods are available for structures with non-standard shapes.

In literature several formulae are proposed for wave overtopping and these formula are expressed in a wide variety of wave-structure parameters. Three are the most popular wave and structural characteristics present in all of them: wave height, wave period and crest freeboard. The overtopping highly non linear dependence of these and other wave and structural characteristics makes the problem extremely difficult to solve; small errors in measurements or slight differences in structural characteristics may have a considerable impact on the results.

There are a number of formulas to estimate overtopping depending of the structural typology and range of environmental variables; however, the reliability of most formulae is not defined, and big differences in predictions are found when using different formulas.

I.2 Purpose of the present study

One of the main objectives of the CLASH project is to obtain a prediction method for the estimation of wave overtopping and uncertainty assessment, applicable to a wide range of coastal structures.

A method is developed herein to provide a conceptual-design tool to estimate wave overtopping discharges for a very wide range of coastal structures. Only one schematisation is used for all types of coastal structures, where not only dikes, rubble mound breakwaters or vertical breakwaters are defined, but also other non-standard structures can be included. Additionally, not only is the effect of the most common parameters described before (wave height, wave period and crest freeboard) analysed herein, but also the effect of many other wave and structural characteristics is considered.

The prediction method that was developed and described in this report is based on the technique of Neural Network modelling. For this purpose use is made of a data set obtained from a large number of physical model tests. This data set includes more than 10,000 tests provided by a large number of laboratories. The partners within the European project [CLASH](#) provided the main part. The present investigation focuses on the development of a neural network for estimating mean wave overtopping discharges. Moreover, a method was developed to obtain the confidence intervals around these predictions. The latter is a new essential step since neural network modelling results in a tool that acts for users as a kind of black box. Consequently, predictions are extended with information regarding their reliability.

1.3 Outline of the present report

This report is organised in six chapters. Chapter 2 contains a brief description of the database that forms the basis for the development of an artificial neural network. The parameters involved in the analysis of wave overtopping are described as well. In Chapter 3 some general issues on neural network modelling are outlined, followed by a brief description of the neural network simulator presently used. Chapter 4 deals with the methodology followed in the construction of the prediction method; first the details of the neural network configuration, followed by the model results are presented; then, the technique used for the uncertainty assessment is described. In Chapter 5, a first sensitivity analysis, showing the influence of some parameters in the outcome of the neural network predictions, is briefly discussed. Finally, Chapter 6 provides an overview of the main conclusions of this study.

2 Description of database and parameters involved

2.1 Description of database

The database that was presently used for the set up of the neural network (hereafter “NN”), is the database created within the framework of the European project CLASH. This database includes tests collected from several laboratories and the results of new tests performed within the CLASH project to fill in “white spots” in the parameter domain. This database is described in detail in Verhaeghe *et al.* (2003) and Steendam *et al.* (2004).

About 10,000 tests are included in this database. Each of these tests is described by a number of parameters that represent hydraulic information (incident wave characteristics, measured overtopping volume) and structural information (characteristics of the test section). Moreover, each of the tests includes some general information regarding the estimated reliability of the test and the complexity of the structure.

The reliability of each test in the database was estimated and defined in terms of a *Reliability Factor (RF)*. The values of the *RF* ranged from $RF = 1$ to $RF = 4$. $RF = 1$ represents a ‘very reliable’ test (i.e. test for which all needed information was available, therefore no estimation of certain parameters was necessary, and for which the measurements were done in a reliable way). $RF = 4$ is assigned to a ‘not-reliable’ test (i.e. test for which the estimation of some parameters was not acceptable and for which the measurements included many uncertainties).

In the same way, the complexity of each test in the database was established on the basis of the complexity of the structure and was defined in terms of a *Complexity Factor (CF)*. The values of the *CF* ranged from $CF = 1$, for a test with a ‘very simple’ structure (i.e. test for which the parameters describe the cross-section exactly), to $CF = 4$, for a test with a ‘very complex’ structure (i.e. test for which the cross-section cannot accurately be described by the chosen parameters).

These *Reliability* and *Complexity Factors* played an important role in the configuration of the NN. With the purpose of giving more weight to tests with high reliability in the NN configuration, the *RF* and the *CF* were combined into a kind of *Weight Factor*. This *Weight Factor* was defined for each of the tests and was taken into account in the *training (calibration)* of the NN. The definition of the *Weight Factor* together with the details of the NN *training (calibration)* process are described in Section 4.2.

2.2 Parameters involved

Due to the large number of parameters involved in the process of wave overtopping, it is difficult to describe the influence of all of them. Clear evidence of this is the few structural

parameters considered in the formulae available in literature. The technique of NN modelling allows for the analysis with a larger amount of structure characteristics.

For the description of the wave field, the effects of three parameters have been considered here: the significant wave height from a spectral analysis at the toe of the structure (H_{m0}), the mean wave period from spectral analysis at the toe of the structure ($T_{m-1,0}$), and the direction of wave attack (β). For the description of the geometrical shape of the structure, the effects of twelve parameters has been considered: the water depth in front of the structure (h), the water depth at the toe of the structure (h_t), the width of the toe berm (B_t), the roughness/permeability of the structure (γ_f), the slope of the structure downward of the berm ($\cot \alpha_d$), the slope of the structure upward of the berm ($\cot \alpha_u$), the width of the berm (B), the water depth on the berm (h_b), the slope of the berm ($\tan \alpha_b$), the crest freeboard of the structure (R_c), the armour crest freeboard of the structure (A_c) and the crest width of the structure (G_c). Figure 1 shows the fifteen parameters used as input to the NN model.

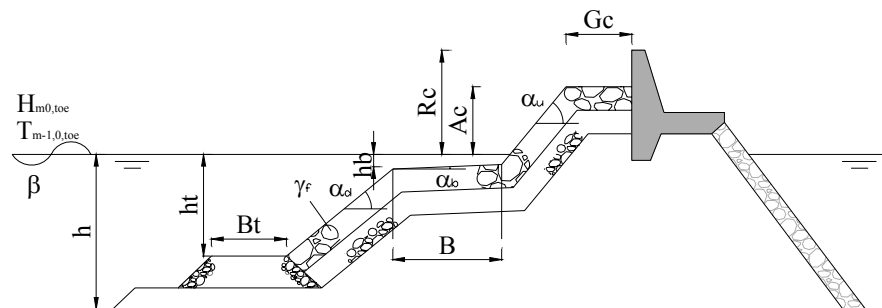


Figure 1 Parameters used for the NN modelling of wave overtopping discharge at coastal structures

Although 12 parameters are used herein to describe the shape of the structure, there remain also aspects of the structure that were not considered in this investigation as input parameters for the NN. Some of the parameters that were not directly taken into account are the size and permeability of the armour layer, filter layer and core material. However, to some extent the roughness/permeability factor (γ_f) incorporates some of these aspects.

For user applications, the determination of each of these wave and structural parameters is briefly described in the ‘Manual Neural Network: NN_OVERTOPPING 2.0’. A more detailed description can be found in Verhaeghe *et al.* (2003).

3 Neural Network modelling

3.1 General

Neural networks (NN) are data analyses or data driven modelling techniques commonly used in artificial intelligence. NN are often used as generalised regression techniques for the modelling of cause-effect relations. NN have proven to be very effective for solving difficult modelling problems in a variety of technical and scientific fields. NN are an alternative technique for the modelling of processes for which the physical interrelationship of parameters is unknown or too complex, while sufficient experimental data is available to identify the desired relation of the parameters.

In the past, NN have been also successfully applied in hydraulic engineering. Examples of NN modelling on coastal structures are: Mase *et al.* (1995) - stability analysis of rock slopes; Van Gent and Van den Boogaard (1998) - wave forces on vertical structures, including estimates for the reliability of these NN predictions; Medina *et al.* (1999, 2002, 2003) - wave run-up and wave overtopping predictions, and armour damage analysis; and Panizzo *et al.* (2003) – wave transmission at low-crested structures. In the present case NN are used for the prediction of wave overtopping discharges at coastal structures and a much larger database than those used in the other NN studies is available.

3.2 Multi-layer feed-forward neural networks

In the literature several types and/or architectures of NN have been proposed (see e.g. Haykin, 1994). In practice most applications deal with the modelling of input-output processes and in such cases the so called Multi-layer feed forward NN are most frequently used. This type of NN is also known as Multi-Layer Perceptrons (MLP), and here this architecture will be adopted as well.

The architecture of an MLP is organised in layers and within each layer there are one or more processing units, the so called ‘neurons’. In a standard multi-layer feed-forward NN the neurons are connected in one direction and from one layer to the next layer only. There are no connections between units of the same layer, nor connections in backward direction. The first layer is called the *input layer* and the number of neurons in this layer equals the number of input parameters (i.e. the dimension of the NN’s input). The last layer is called the *output layer* and the number of neurons in this layer is equal to the number of output parameters to be predicted. The layers between the input layer and output layer are called *hidden layers*. The number of neurons in these hidden layers must be defined in the preparation of the NN.

Information is propagated from left to right, from an input layer via the hidden layers to the output layer. Each neuron in each layer receives information from the preceding layer through the network’s connectivities. To each connectivity a weight is assigned, and these weights are used in the neuron’s integration of its inputs from (all the units of) the preceding

layer. Actually this integration consists of a linear weighted sum of the outputs of the preceding layer and the outcome of this weighted sum gives the *state* of the neuron. By means of an activation function, applied to this state, the neuron's output is produced. For the neurons in the hidden layer the activation function is always non-linear, and is often of a sigmoid form. However, for the output layer a linear activation function is commonly applied (Haykin, 1994). The output neuron(s) generates the final prediction of the NN.

The activation functions in this architecture are usually fixed, while the weights of the connectivities in the network are adjustable. In this way an MLP can well be seen as a parameterised non-linear black box model, with the NN's weights as the uncertain model parameters. The weights must be identified by a calibration of the model. This means that a data set of observed input-output patterns must be available and the weights are set to the values for which the NN's predictions agree as good as possible to the observed outputs (the so called targets). In NN modelling this tuning of weights is usually called 'training'. In this way 'training' is a form of (model) calibration, for which numerical techniques must be applied. Because of their computational efficiency gradient descent algorithms based on the "Error Back Propagation Rule" (Rumelhart, McClelland and Williams, 1986) are virtually always used. In this procedure an initial guess of the weights is iteratively updated. In each epoch where an input (or all the inputs) is presented to the network the MLP's prediction error is evaluated, and this error is propagated over the NN in backward direction. On the basis of these (back propagated) errors the weights can be adjusted in a way that in a next cycle the NN provides better predictions. This 'training' is repeated until no further improvements are obtained. From a mathematical viewpoint this training corresponds to a minimisation of an error (or cost) function.

For the present NN modelling of overtopping the details of the NN *calibration* or *training* process are described in Section 4.2.

It has been shown (Alexander and Morton, 1991) that in practice for most applications, no further accuracy is gained by using more than one hidden layer in the network. In this study a three-layered MLP (one hidden layer) is applied as well. An example of such a three-layered NN is given in Figure 2. In this configuration, the input and output vectors consist of three and one parameters, respectively. In this sample configuration, the hidden layer consists of two units. As a consequence the NN has eight links and therefore eight associated weights. There is no rule or general guideline for the choice of the number of units in the hidden layer. Several studies have been conducted with the aim of solving this matter, but no general conclusions have been reached so far. Usually, the optimal configuration is decided by evaluating several possible alternatives and comparing their performances for the given applications. This will be explained in more detail in Section 4.2.

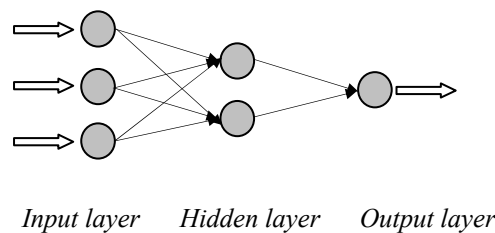


Figure 2 Example of the configuration of a three-layered NN

3.3 Applied Neural Network simulator

Several standard software programs are available for the configuration of an NN. They vary in flexibility, types of networks that are implemented, and form of user interface. In the current investigation, a neural network simulator MLP developed by WL | Delft Hydraulics was used.

The MLP simulator is dedicated to MLP applications, also known as feed-forward, error-backpropagation networks. It allows for any number of hidden layers, and an arbitrary number of processing units (neurons) in the configuration. The learning rule is based on the error-backpropagation rule in combination with an efficient quasi-Newton gradient descent algorithm.

The prediction tool NN_OVERTOPPING developed herein makes use of the MLP simulator program previously described. A userguide for the prediction method is described in the ‘Manual Neural Network: NN_OVERTOPPING 2.0’.

4 Preparation and validation of the NN models

This chapter summarises the methodology followed for the construction of the prediction method. The objective of the prediction method is to estimate mean overtopping discharges.

Two NN models were considered, and the database used in the configuration of each of them was the only difference between both NN models. Firstly, the preparations and analysis of the database used for the configuration of the NN are described. Secondly, the steps followed in the configuration of the NN followed by the model results are presented. Finally, the technique used for the assessment of the uncertainty of the predictions is briefly discussed.

4.1 Preparations of the database used for the NN configuration

4.1.1 Validation and reduction of database

Since the quality of the NN depends highly on the quality of the database (erroneous data can severely degrade the performance of the NN), the initial database of more than 10,000 tests was reduced by removing the data that was qualified as ‘non-reliable’ tests (e.g. tests with a ‘reliability factor’ $RF = 4$, as defined previously) or the data for which the cross-section was considered as ‘very complex’ (e.g. tests with a ‘complexity factor’ $CF = 4$, as defined previously).

Since most of the tests in the database originate from small-scale tests, very low overtopping discharges are likely to be less accurate due to measurement techniques/errors in these small-scale tests. Here data with a mean overtopping discharge smaller than $q = 10^{-6} \text{ m}^3/\text{s}/\text{m}$, (i.e. 1 ml/s/m) is considered as less accurate than larger overtopping discharges. Nevertheless, for practical applications it would be useful to predict also very low wave overtopping discharges, here characterised by the non-dimensional limit of $Q = 10^{-6}$. Note that this non-dimensional limit ($Q = 10^{-6}$) corresponds to lower overtopping discharges than the limit of the data with reduced accuracy ($q = 1 \text{ ml/s}/\text{m}$). Data with $q = 0$ has not been used since in different sub-sets of the database, different definitions of $q = 0$ have been used; sometimes “ $q = 0$ ” is related to all discharges smaller than $10^{-4} \text{ m}^3/\text{s}/\text{m}$, sometimes smaller than $10^{-6} \text{ m}^3/\text{s}/\text{m}$, and sometimes even lower values.

Only the tests related to registered overtopping events (i.e. $q \neq 0 \text{ m}^3/\text{s}/\text{m}$) were considered in the setup of the NN. Some data corresponding to overtopping events such that $q > 10^{-6} \text{ m}^3/\text{s}/\text{m}$ were rather randomly checked and clear inconsistencies within the database were eliminated. The resulting database used for the NN model consisted of 8,372 tests.

4.1.2 Weight Factors

As mentioned in Section 2.1, a *Weight Factor* (hereafter *WF*) was defined for each of the tests in the above-mentioned databases. The purpose of this *WF* was to give more importance to tests with high reliability. This *WF* was defined as a combination of the Reliability and Complexity Factors according to:

$$WF = (4 - RF) \cdot (4 - CF)$$

Table 1 shows the possible values of the *WF* used in the NN configuration for the different combinations of *RF* and *CF* in the database. The values of this *WF* ranged from $WF = 9$, for a ‘very reliable’ test with a ‘very simple’ structure ($RF = CF = 1$), to $WF = 1$, for a ‘low-reliable’ test with a ‘quite complex’ structure ($RF = CF = 3$). These weights were included in the NN’s *calibration* process. In fact, the *WF* assigns a multiplicity to the input-output patterns in the NN’s training (see Section 4.3.2).

<i>RF</i> \ <i>CF</i>	1	2	3
1	9	6	3
2	6	4	2
3	3	2	1

Table 1 Values of the Weight Factor for different combinations of *RF* and *CF* in the database

It should be noted that for the configuration of the NN, all the tests with observed overtopping discharges $q < 10^{-6}$ m³/s/m, were given a weight factor $WF = 1$ (i.e. $RF = CF = 3$), since the measurements of wave overtopping discharges in this range can be considered as relatively inaccurate due to limitations in the small-scale tests.

4.1.3 Froude’s similarity law

Since it was the aim of the NN to be applicable both for small-scale and prototype conditions, all the input and output parameters in the database were scaled to $H_{m0,toe} = 1$ m using Froude’s similarity law. The advantage of using Froude’s law, taking into account that the database was mainly based on small-scale tests, is that a better generalisation for large scale applications can be obtained. Since physical knowledge was then incorporated in the NN, this procedure allowed for predictions in different scales.

From the Froude law the following scaling relationships, expressed in terms of the length scale n_L , are derived. The length scale n_L is determined by the ratio of the wave height $H_{m0} = 1.0$ m to the measured wave height, $n_L = 1/H_{m0,toe (observed)}$.

- $H_{m0}, h, h_t, B_t, B, h_b, R_c, A_c, G_c$ (m) : $[n_H, n_h, n_{ht}, n_{B_t}, n_B, n_{hb}, n_{R_c}, n_{A_c}, n_{G_c}] = n_L$
- T (s) : $n_T = n_L^{0.5}$
- q (m³/s/m) : $n_q = n_L^{1.5}$

The corresponding parameters scaled with Froude’s law to $H_{m0} = 1.0$ m will be hereafter denoted as $[H_{m0}', T_{m-1,0}', \beta', h', h_t', B_t', \gamma_f', \cot \alpha_d', \cot \alpha_u', B', h_b', \tan \alpha_b', R_c', A_c', G_c', q']$.

For user applications, when a prediction of the wave overtopping discharge is required for a certain input-pattern $[H_{m0}, T_{m-1,0}, \beta, h, h_t, B_t, \gamma_f, \cot \alpha_d, \cot \alpha_u, B, h_b, \tan \alpha_b, R_c, A_c, G_c]$ this input-pattern is scaled according to Froude's similarity law to an input-pattern $[H_{m0}', T_{m-1,0}', \beta, h', h_t', B_t', \gamma_f', \cot \alpha_d', \cot \alpha_u', B', h_b', \tan \alpha_b', R_c', A_c', G_c']$ with a wave height on which the NN was trained ($H_{m0, toe} = 1$ m). The NN prediction (q') is then scaled back (q) to the original wave height using again Froude's law. The result of this approach is that the NN predictions do follow Froude's law. This means that for a small-scale condition (e.g. model scale) and an identical condition on a larger scale (e.g. prototype), the NN predictions are the same after scaling the output discharge to the same scale (e.g. prototype).

This approach required somehow a less complex configuration of the NN, since the number of input-patterns used reduced from 15 to 14 (once all input-patterns were scaled to $H_{m0, toe} = 1$ m, this parameter, H_{m0}' , is constant and was not used anymore as a separate input pattern to the NN, although the user of the tool provides this wave height to scale the input and the output; i.e. pre- and postprocessing of the input-output pattern).

In the NN-modelling Froude's scaling law was used to extrapolate the information from small scale tests to prototype conditions. However, the small scale tests may, to some extent, be affected by scale effects. Since there is no clear quantitative information on the magnitude of scale effects for each test condition (type of structure, wave conditions, etc) the NN predictions do not incorporate the influence of scale effects. Once there is clear information on scale effects for specific applications, the NN output may have to be corrected for scale effects. For instance, if for a rubble mound structure it is considered appropriate to account for a 20% increase in mean overtopping discharge to compensate for the influence of scale effects, in the comparison of a small-scale situation with $H_{m0} = 0.1$ m (order of magnitude of wave heights in the database) and a prototype situation with $H_{m0} = 4$ m, the output of the NN should be increased with a factor 1.2 for such application. However, quantitative information on scale effects is not yet available.

4.1.4 Data distribution

Using Froude scaling as described above, the database used in the NN configuration can be statistically summarised as shown in Table 2. It should be taken into account that the input components of the database are not uniformly distributed within the ranges of the parameters given in Table 2.

<i>Parameter</i>	<i>Mean</i>	<i>Standard Deviation</i>	<i>Maximum</i>	<i>Minimum</i>
H_{m0}' (m)	1.0	0.0	1.0	1.0
$T_{m-1,0}'$ (s)	5.0	2.8	69.5	2.7
β'	3.4	11.3	80.0	0.0
h' (m)	3.7	2.8	32.2	0.9
h_i' (m)	3.3	2.7	25.9	0.5
B_i' (m)	0.7	1.5	19.0	0.0
γ_f'	0.7	0.2	1.0	0.3
$\cot \alpha_d'$	1.9	1.4	7.0	0.0
$\cot \alpha_u'$	1.8	1.7	9.7	-5.0
B' (m)	0.7	2.1	38.5	0.0
h_b' (m)	0.1	0.4	7.9	-1.2
$\tan \alpha_b'$	0.0	0.01	0.13	0.0
R_c' (m)	1.4	0.6	6.4	0.0
A_c' (m)	1.3	0.6	6.2	0.0
G_c' (m)	0.9	1.3	13.2	0.0
q' (m ³ /s/m)	$8 \cdot 10^{-3}$	$2 \cdot 10^{-2}$	$4 \cdot 10^{-1}$	$1 \cdot 10^{-7}$

Table 2 Statistical parameters of the scaled input / output-patterns of the database used for calibration / training of the NN (scaled to $H_{m0}=1$ m)

4.2 NN configuration

As mentioned in Section 3.2, the NN model used in this investigation was a Multi-layer (three-layered) feed-forward network trained on the basis of the standard error-backpropagation learning rule and gradient descent algorithms.

In the following, some aspects that were taken into account in the training of the NN models are described.

4.2.1 Training and testing

The preparation of the NN model was performed in two phases, the *Training/Learning* phase and the *Testing/Validation* phase.

The process of calibration or *training* phase of the NN involves the adjustment of its configuration (calibration of the NN's weights) based on the performance of standard operations that allows the NN to *learn* from the input-output relations for each of the parameters included in the tests selected as *training set*. Starting with an initial guess for the

weights, the inputs of this *training set* are fed to the NN and the NN outputs are compared to the observed/measured outputs. On the basis of the differences between both, the weights are adjusted in such a way that when in the next phase the inputs are again fed to the NN, a better NN prediction is found. The procedure is repeated until no further improvements can be made. This iterative adjustment of the weights or *training* of the NN is performed by minimisation of some *cost function* (error function) that quantifies the differences between the predicted outputs and the desired measured/observed outputs, often called *targets*. A common form of the *cost function* is a superposition of the squared differences. As mentioned in the previous section, for the minimisation of the *cost function*, gradient based methods turned out to be the most efficient. For the computation of the gradient of the cost function, the well known error-backpropagation rule was used. It should be noted that the training of the NN was performed considering the logarithm of the observed overtopping discharges scaled with Froude's law ($\log q_{obs}'$) as the targets. The root-mean-square error used herein is defined as follows:

$$RMS_{train} = \sqrt{\frac{1}{N_{train}} \sum_1^{N_{train}} ((\log q_{obs}')_n - (\log q_{NN}')_n)^2}$$

where N_{train} is the number of tests considered in the NN training. Once trained (or during training), the correct performance of the resulting model is evaluated with a *testing set*, i.e. a set of input-output combinations not used before for training. This step is called *validation* or *testing* phase of the NN.

4.2.2 Training and testing sets based on resampling

It should be noted, that the actual partition of the data over the training and testing sets may significantly affect the outcome of the NN. Therefore, special attention must be given to this aspect. The present strategy was to train many NN's each time with different data in the training and testing phase. *Resampling Techniques* (Van den Boogaard *et al.*, 2000) were used for the construction of such training and testing sets. An important advantage of using Resampling Techniques is that also estimates for the uncertainty (or reliability) of a NN prediction can be obtained. Details of the Resampling Technique used in the present investigation are described later in Section 4.3.

4.2.3 Optimum number of neurons in the hidden layer

An important step in the configuration of the NN is to find the optimal number of neurons in the hidden layer. By increasing the number of neurons in the hidden layer, the differences between the NN output and the desired (observed/measured) output of the data used for calibration/training will decrease because more hidden neurons lead to more degrees of freedom (more adjustable parameters in the NN model). However, after a certain number of hidden neurons, the NN starts to model *noisy fluctuations* in the dataset which is unfavourable for the accuracy of the real predictions; the performance of the NN for the training set increases (root-mean square error decreases), while that of the testing set decreases (root-mean square error increases). At that moment the NN is said to be *overtrained*. To avoid *overtraining* of the NN, an *early stopping* (Heskes, 1997) criterion is used in the NN training process. With this technique, the training process of the NN is stopped when the performance of the NN in the testing set starts to decrease.

The optimal number of hidden neurons can be found by training the NN several times for a range of number of neurons in the hidden layer, and comparing each time the performance of the NN (root-mean square error) on the training and testing sets. Figure 3 shows the performance of the NN for different number of neurons in the hidden layer. From this figure it can be observed, that the root-mean-square error decreases rapidly for a small number of hidden neurons, while starts becoming more stable for larger number of hidden neurons.

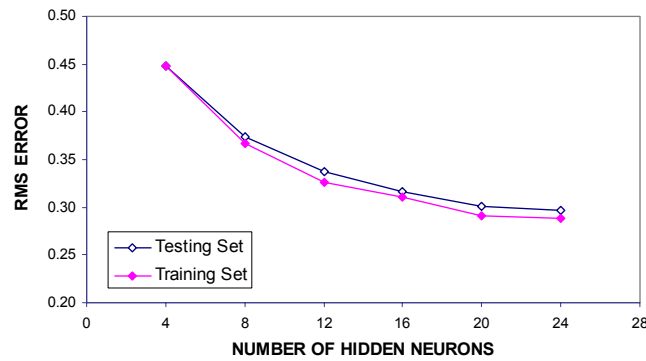


Figure 3 NN performance for configurations with different number of neurons in the hidden layer

Based on the results of these experiments, the optimum number of hidden neurons was chosen as 20, since the use of more hidden neurons did not really increase the accuracy of the NN while it increased the complexity of its architecture. Figure 4 shows the final configuration of the NN, with 15 neurons in the input layer (input parameters, wave/structure), 20 neurons in the hidden layer, and one neuron in the output layer (output parameter, the actual overtopping discharge).

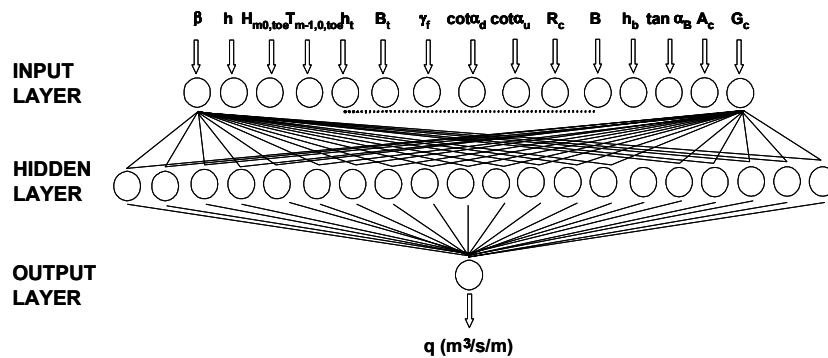


Figure 4 Final NN configuration

4.3 Uncertainty assessment

After obtaining the optimal NN configuration, predictions of mean overtopping discharges could be made; i.e. for a set of input parameters, a set of output parameters (overtopping discharge) could be obtained. Van Gent and Van den Boogaard (1998) developed a method to add information on the reliability of NN predictions. The method to add information on the *reliability* of NN-predictions has been developed further within this investigation, and makes use of how the available data are spread over the entire domain of applications. With

the double purpose of solving the matter of choosing the data to be used in the training and testing phases, and assessing the uncertainty of the NN predictions, Resampling techniques were used.

Resampling techniques are generic devices used for uncertainty analysis in statistics and model calibration. The use of these techniques involves the development of a set of NNs (*resamples*) based on the original database. This implies firstly that the training and testing processes are redone many times solving the problem of representativeness of the training and testing sets; and secondly, that the set of NNs developed results in a *set of predictions* (a so called *ensemble*) of overtopping discharge, allowing the estimation of the reliability of the predictions (i.e. standard deviation or 95% confidence intervals). As a result, the NN does not only give a prediction of the wave overtopping discharge but also a measure for the reliability (uncertainty) of the prediction.

4.3.1 Bootstrap resampling

The two most commonly applied forms of resampling are *jackknifing* and *bootstrapping* (Tichelaar and Ruff, 1989; Efron, 1982; Wu, 1986). In the current investigation, the technique of *bootstrap resampling* was used.

A *bootstrap* resample is a random selection of N data out of the N original data. The N individual draws within one such resample are independent but with replacement so that every time there is a probability of $1/N$ that a particular sample (input-output combination) of the original set is selected. At the end some samples are then selected once or more than once while other samples are absent in a resample. The samples selected in each resampling correspond to the tests that are used for the training of the NN, while the ones not selected correspond to the tests that are used in the testing/validation of the NN. The probability that an original sample is not present in a resample is $(1-1/N)^N$ which for large N is close to $1/e$. Therefore, within each resample, the probability that a test was selected (used for training) was 63%, while the probability that a test was not selected (used for testing) was 37%. In this way a set of L of such resamples is generated. This L should be sufficiently large and in practice it is typically of the order of a hundred or a few hundreds, somewhat depending on the statistics to be computed. For each resample of the data set (and the corresponding division into a training and testing set) the NN is trained.

A set of about 500 NNs or resamples was performed herein. As a result, the bootstrap yielded 500 estimates of the wave overtopping discharge (q_1', \dots, q_{500}'). The estimate of the model (NN output, q_{NN}') is given by the mean of all these predictions:

$$\log \mu = \log q_{NN}' = \frac{1}{L} \sum_{i=1}^L \log q_i'$$

where $L = 700$, in this case.

4.3.2 Weight Factors in the bootstrap resampling

In the bootstrap resampling the weights are included as follows. If a weight WF is assigned to a particular input-output pattern, i , and within a resampling this pattern is selected N

times, the total weight of the pattern is set to $WF_i \cdot N$. The contribution of this pattern to the cost function is then,

$$WF_i \cdot N \cdot \left((\log q_{obs})_i - (\log q_{NN})_i \right)^2$$

The total cost function during training in a bootstrap resample is the superposition of such terms for all the input-output patterns that are actually selected.

4.3.3 Statistics of the predictions: Confidence intervals

The set of 700 NNs constructed is further used to provide the uncertainty of the model with respect to the accuracy of its predictions. This uncertainty can be quantified by a standard deviation, or the variance, or a confidence interval. In this respect, the prediction method developed herein provides several statistics. Besides the mean, μ , of the predictions which is used as the model prediction ($\mu = q_{NN}$), the output of the prediction method includes quantiles of several orders, $q_{2.5\%}$, $q_{5\%}$, $q_{25\%}$, $q_{50\%}$, $q_{75\%}$, $q_{95\%}$ and $q_{97.5\%}$. The 95% confidence interval is, for instance, given by the quantiles $q_{2.5\%}$ and $q_{97.5\%}$.

It should be noted that the uncertainty assessment is based on how data is spread over the entire domain of application; it still assumes that the database is *correct*. The uncertainty levels do not account for systematic error or inaccuracies in the database (for instance caused by model effects, scale effects or measurement equipment).

4.4 Performance of the NN model

As mentioned at the beginning of this chapter, a NN was constructed on the basis of the data corresponding to registered overtopping events ($q \neq 0 \text{ m}^3/\text{s}/\text{m}$). In the following, the performance of the NN to be used in the prediction method is presented.

In a first stage, the NN was first applied to the subset of the data for which $q > 10^{-6} \text{ m}^3/\text{s}/\text{m}$. The RMS error of the differences between the observations and the NN predictions was $\text{RMSE} = 0.29$.

Figure 5 shows the observed wave overtopping discharge versus the predicted wave overtopping discharge for the range of overtopping discharges $q > 10^{-6} \text{ m}^3/\text{s}/\text{m}$. As described in the previous section, the predicted overtopping discharge (model prediction) corresponds to the mean of the set of 700 NNs. It can be observed that the predictions of the NN are reasonably accurate, especially in the range of high overtopping discharges. Additionally, to give a first idea of the amount of spread in the NN predictions, this figure shows an indicative 95% confidence band, where the amount of points in the outer regions of these bands represent only 5% of the total number of points (or NN output-observation combinations). Note that the confidence band is not the result of the uncertainty analysis as described in the previous section. Figure 5 indicates that there are almost no data with predictions a factor of 10 larger/smaller than observations. In this respect, it should be noted that the repetition of a certain test in the laboratory often gives a factor 5 difference. The indicative 95% confidence band can be expressed as follows:

$$\log q_{NN-95\%} = \begin{cases} 0.86 \cdot \log q_{obs} \\ 1.16 \cdot \log q_{obs} \end{cases}$$

Figure 6 shows the observations/targets versus the NN predictions (for $q > 10^{-6} \text{ m}^3/\text{s/m}$), where the (q_{obs}, q_{NN}) samples are plotted with a colour representing the Reliability Factor (RF) of the observed overtopping discharge.

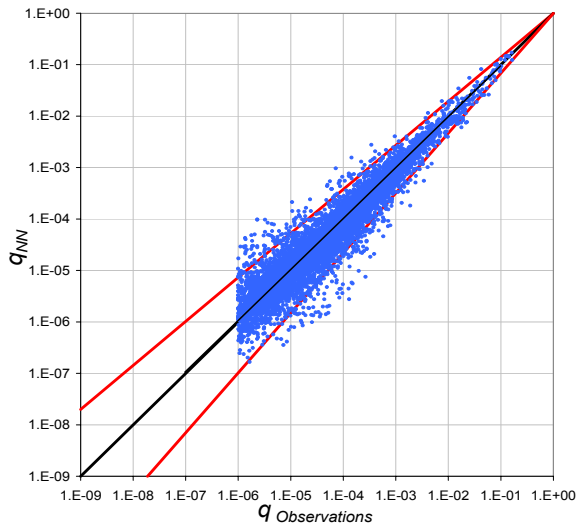


Figure 5 Observations versus NN predictions;
application on data $q > 10^{-6} \text{ m}^3/\text{s/m}$
(● NN predictions; — 95% confidence interval)

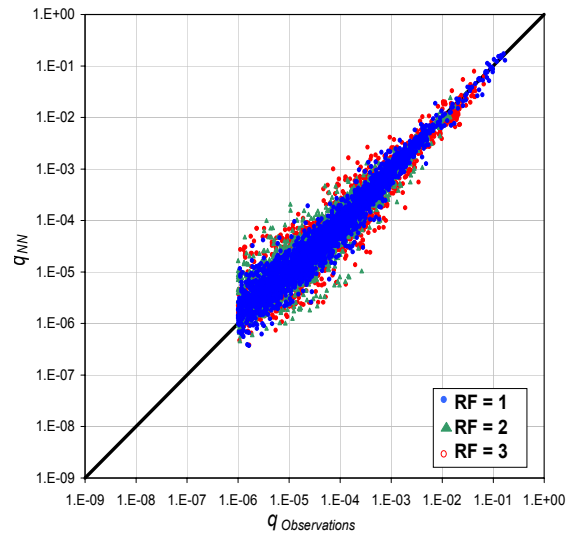


Figure 6 Observations versus NN predictions for
different Reliability Factors; application on
data $q > 10^{-6} \text{ m}^3/\text{s/m}$

To illustrate the performance of the NN for the complete range of overtopping discharges used for its configuration, Figure 7 shows in the same way as Figure 5, the observed wave overtopping discharge versus the predicted wave overtopping discharge (mean of the 500 NNs). This figure also shows the 95% confidence band, which also indicates that there are almost no data with predictions a factor of 10 larger/smaller than observations. Figure 10 shows again the observations versus the NN predictions (for $q > 10^{-6} \text{ m}^3/\text{s/m}$), where each test has been plotted with its corresponding Reliability Factor (RF), as given in the database.

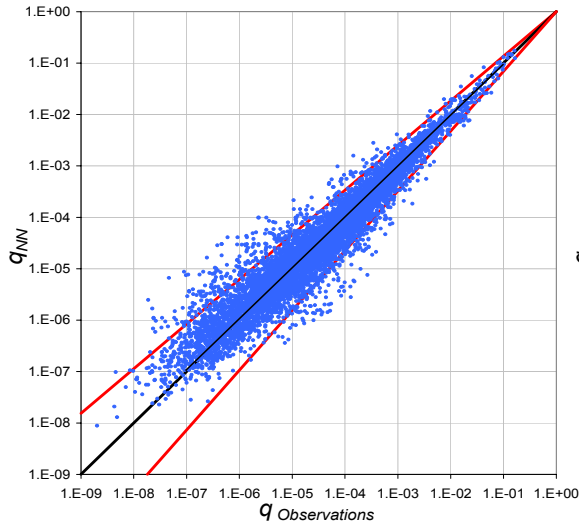


Figure 7 Observations versus NN predictions
(● NN predictions; — 95% confidence interval)

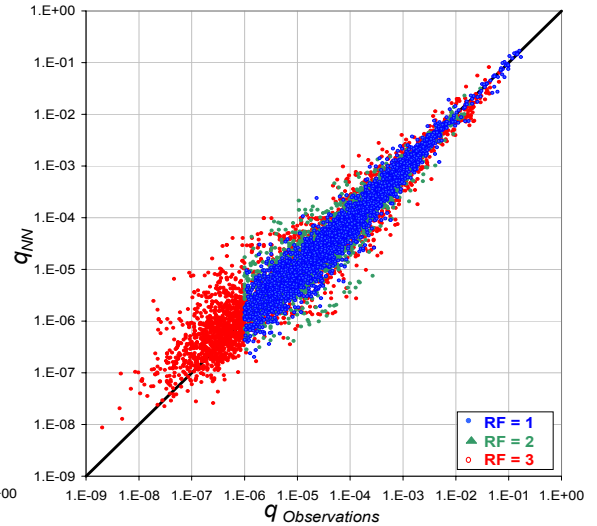


Figure 8 Observations versus NN predictions for different Reliability Factors

Because applications are envisaged for conditions with a non-dimensional overtopping discharge of $Q > 10^{-6}$, Figures 9 and 10 show the same results as Figures 7 and 8 but now with the axes in non-dimensional scale.

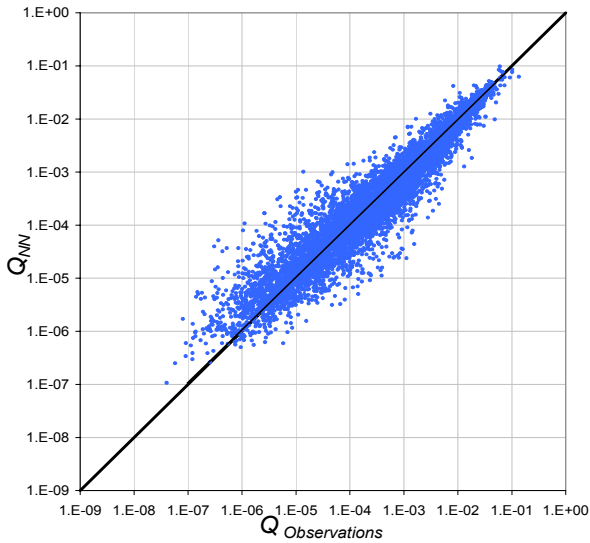


Figure 9 Observations versus NN predictions in dimensionless form (● NN predictions)

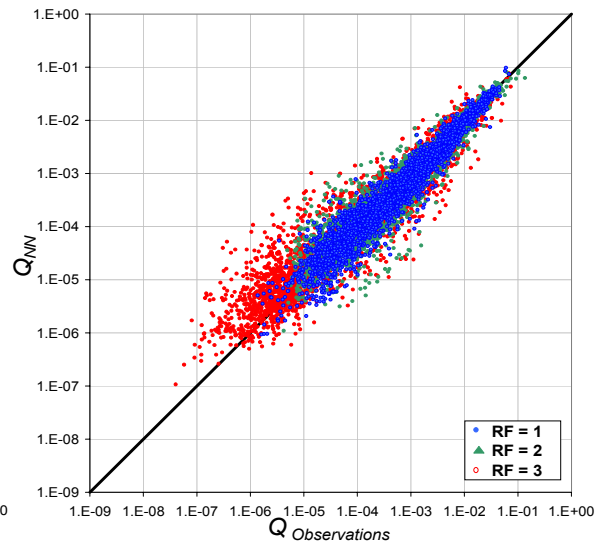


Figure 10 Observations versus NN predictions for different Reliability Factors in dimensionless form

4.4.1 Classification of overtopping events

As an additional step, it was verified whether or not the present NN could be used to classify events with very low overtopping discharges and events with higher overtopping discharges. This was done first to distinguish between events with $q > 10^{-6} \text{ m}^3/\text{s/m}$ and events with $q < 10^{-6} \text{ m}^3/\text{s/m}$.

On this assessment of the NN's potential use for classification, four possibilities were distinguished given a particular input pattern:

1. The NN predicts significant overtopping ($q > 10^{-6} \text{ m}^3/\text{s/m}$) and the observation also represents significant overtopping: *correctly classified* (S1).
2. The NN predicts insignificant overtopping ($q < 10^{-6} \text{ m}^3/\text{s/m}$) and the observation also represents insignificant overtopping: *correctly classified* (S2).
3. The NN predicts significant overtopping ($q > 10^{-6} \text{ m}^3/\text{s/m}$) while the observation represents insignificant overtopping ($q < 10^{-6} \text{ m}^3/\text{s/m}$): *wrongly classified* (S3).
4. The NN predicts insignificant overtopping ($q < 10^{-6} \text{ m}^3/\text{s/m}$) while the observation represents significant overtopping ($q > 10^{-6} \text{ m}^3/\text{s/m}$): *wrongly classified* (S4).

Figure 11 shows the same results as shown before in Figure 7, but now, the figure has been divided into four quadrants corresponding to the four situations (S1, S2, S3 and S4) previously described. In order to determine the performance of the NN as classifier, the number of tests lying in each of the quadrant was counted. Table 1 shows the results of this classification, where “1” refers to situations with overtopping events $q > 10^{-6} \text{ m}^3/\text{s/m}$, and “0” refers to situations with small overtopping discharges ($q < 10^{-6} \text{ m}^3/\text{s/m}$).

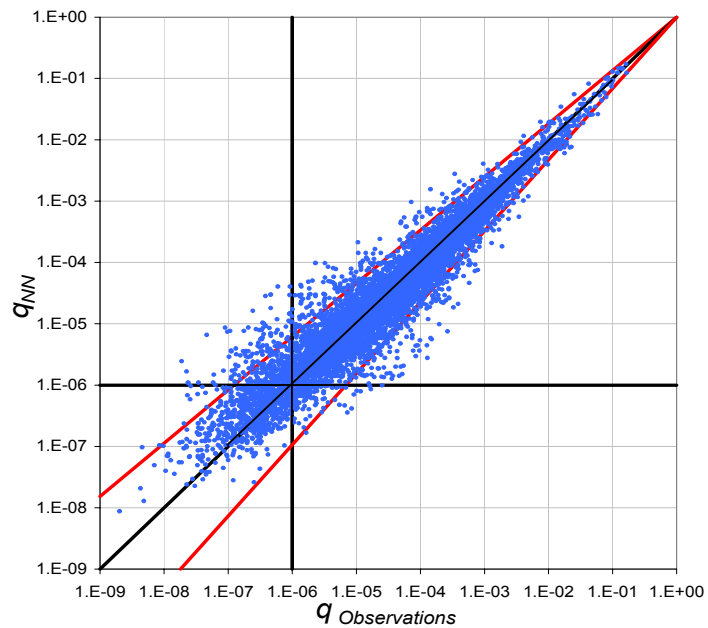


Figure 11 Observations versus NN predictions; classification analysis

(● NN predictions; — 95% confidence interval)

		OBSERVATIONS		
		1	0	
NN PREDICTIONS	1	S1 = 7160	S3 = 333	7493
	0	S4 = 134	S2 = 745	879
		7294	1078	8372

Table 3 Results of the classification analysis (Total number of tests in the reduced database, $N_{total} = 8372$)

This table indicates that the percentage of tests correctly classified $[(S1 + S2)/N_{total}]$ is 94%, and the percentage of tests wrongly classified $[(S3 + S4)/N_{total}]$ is 6%. This allows the conclusions that this NN can rather accurately distinguish between overtopping events with $q > 10^{-6} \text{ m}^3/\text{s}/\text{m}$ and events with $q < 10^{-6} \text{ m}^3/\text{s}/\text{m}$, and this be suited for classification.

A second classification of wave overtopping events was performed, but now not for events with wave overtopping discharges smaller or larger than $q = 10^{-6} \text{ m}^3/\text{s}/\text{m}$ but for events with wave overtopping discharges smaller or larger than the non-dimensional wave overtopping discharge $Q = 10^{-6}$.

Again four possibilities were distinguished given a particular input pattern:

1. The NN predicts significant overtopping ($Q > 10^{-6}$) and the observation also represents significant overtopping: *correctly classified* (S1).
2. The NN predicts insignificant overtopping ($Q < 10^{-6}$) and the observation also represents insignificant overtopping: *correctly classified* (S2).
3. The NN predicts significant overtopping ($Q > 10^{-6}$) while the observation represents insignificant overtopping ($Q < 10^{-6}$): *wrongly classified* (S3).
4. The NN predicts insignificant overtopping ($Q < 10^{-6}$) while the observation represents significant overtopping ($Q > 10^{-6}$): *wrongly classified* (S4).

Figure 12 shows the same results as shown before in Figure 11, but now with the axes in non-dimensional scale. Similarly, the figure has been divided into four quadrants corresponding to the four situations (S1, S2, S3, S4) previously described. In order to determine the performance of the NN as classifier, the number of tests lying in each of the quadrant was counted. Table 4 shows the results of this classification, where “1” refers to situations with overtopping events $Q > 10^{-6}$, and “0” refers to situations with insignificant overtopping events ($Q < 10^{-6}$).

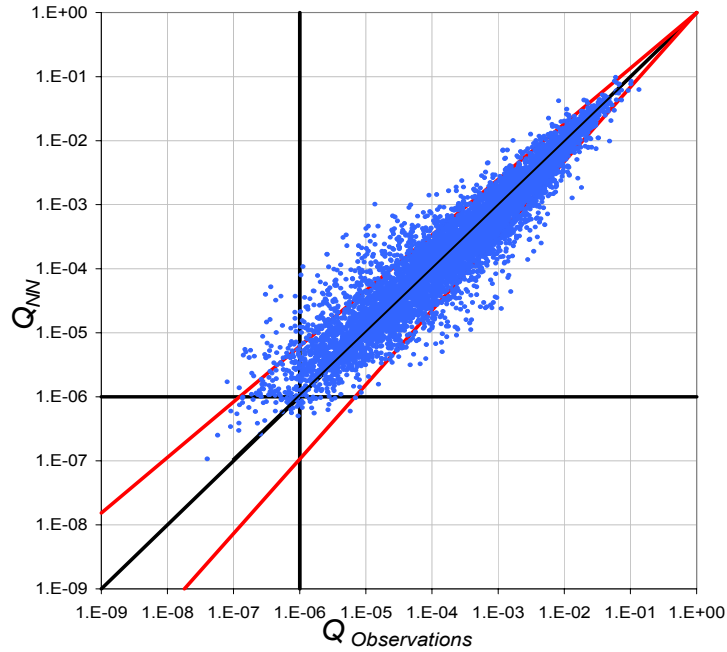


Figure 12 Observations versus NN predictions; classification analysis in dimensionless form (● NN predictions; — 95% confidence interval)

		OBSERVATIONS		
		1	0	
NN PREDICTIONS	1	S1 = 8164	S3 = 127	8291
	0	S4 = 42	S2 = 39	81
		8206	166	8372

Table 4 Results of the classification analysis in dimensionless form (Total number of tests the reduced database, $N_{total} = 8372$)

This table indicates that the percentage of tests correctly classified $[(S1 + S2)/N_{total}]$ is 98%, and the percentage of tests wrongly classified $[(S3 + S4)/N_{total}]$ is 2%. Although these results suggest that this NN can rather accurately distinguish between overtopping events with $Q > 10^{-6}$ and events with $Q < 10^{-6}$, it should be noted that there is only a small percentage of tests with observed overtopping events with $Q < 10^{-6}$ (2%) and that it means that many of the conditions with observed overtopping ($S2$ and $S3$) smaller than $Q = 10^{-6}$ is classified incorrectly: $S3/(S2+S3) = 76\%$.

This analysis shows that the NN is relatively accurate for $q > 1$ ml/s/m. Since the data on which this NN is based contains mainly wave heights in the order of magnitude of 0.1 m, 1 ml/s/m corresponds roughly with a non-dimensional discharge of $Q = 10^{-5}$. Therefore, it is recommended not to use the prediction method for $Q < 10^{-6}$ and to use the results with extra care for $Q < 10^{-5}$ due to the expected relatively large inaccuracies in the data with small overtopping discharges on which the NN predictions are based.

5 Sensitivity Analysis

A sensitivity analysis can demonstrate the effect of the input parameters on NN's prediction and the uncertainty of this prediction. In the following, four examples of the application of the NN for different types of structures are given.

Figure 13 shows the influence of an increase of the relative crest freeboard on the predictions of overtopping discharge for a vertical structure. As expected, an increase of the relative crest freeboard results in a decrease of the wave overtopping discharge predicted by the NN model. The reliability of these predictions, given by the 95% confidence band, indicates a maximum difference of a factor 10. Additionally, the influence of the crest freeboard was also studied with the empirical formula proposed in Allsop *et al.* (1995). For structures with a low freeboard $R_c/H_{m0} < 1$, the dimensionless overtopping discharge predicted with the NN is slightly higher than that predicted by Allsop *et al.* (1995). For structures with relatively higher freeboard $1 < R_c/H_{m0} < 3$ (where the range covered by the empirical formula is $0.03 < R_c/H_{m0} < 3.2$), and therefore relatively lower overtopping discharges, the NN predictions of wave overtopping are very similar to those by Allsop *et al.* (1995).

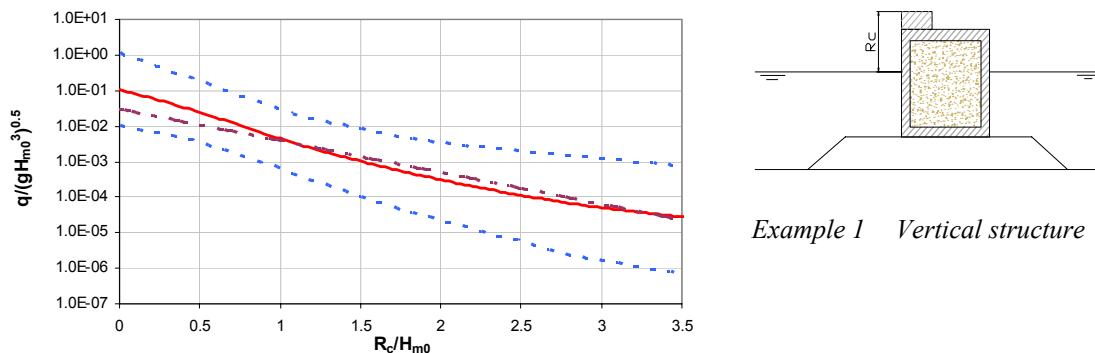
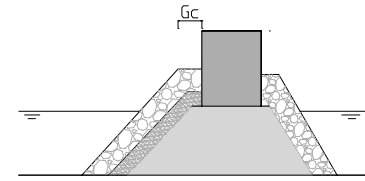
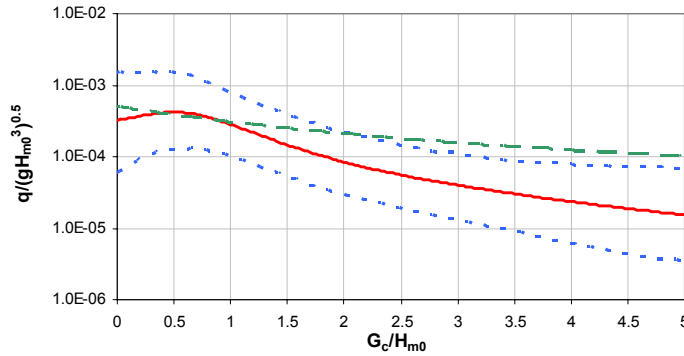


Figure 13 Sensitivity of wave overtopping to relative crest freeboard; Vertical structure

(— NN predictions; - - - 95% confidence interval; - - - Allsop *et al.* 1995, predictions)

Figure 14 shows the influence of an increase of the relative armour width in front of the crest wall on the predictions of overtopping discharge for a conventional rubble mound structure. As expected, the general trend indicates a decrease of the dimensionless wave overtopping discharge predicted by the NN model with increasing armour crest widths. For structures with a crest wall directly next to the end of the armour slope ($G_c/H_{m0} = 0$) the overtopping discharge is similar as for structures with a relative armour width of $G_c/H_{m0} = 0.5$. The influence of the armour width was also studied with the TAW guidelines for the design of dikes (TAW, 2002). In general, for structures with relative armour widths $G_c/H_{m0} > 1$, the NN predictions are lower than those provided by the TAW guidelines.

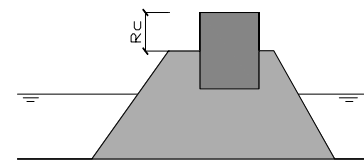
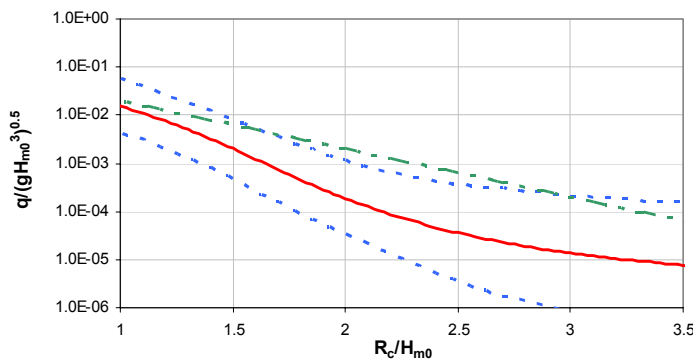


Example 2 Rubble-mound structure

Figure 14 Sensitivity of wave overtopping to relative armour width; Rubble mound structure

(— NN predictions; 95% confidence interval; - - TAW predictions)

Figure 15 shows the influence of an increase of the relative crest freeboard on the predictions of overtopping discharge for a smooth structure. In the same way as in Figure 8, an increase on the relative crest freeboard results in a decrease of the dimensionless wave overtopping discharge predicted by the NN model. The influence of the relative crest freeboard was also studied with the TAW guidelines for the design of dikes (TAW, 2002). For structures with a relative crest freeboard of $R_c/H_{m0} \sim 1$, the NN predictions are similar to those provided by the TAW guidelines. For structures with a relatively high crest freeboard $R_c/H_{m0} > 1$, the NN predictions are lower than those provided by the TAW guidelines.

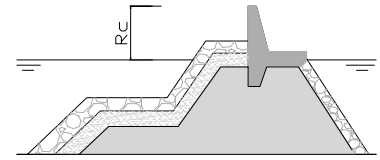
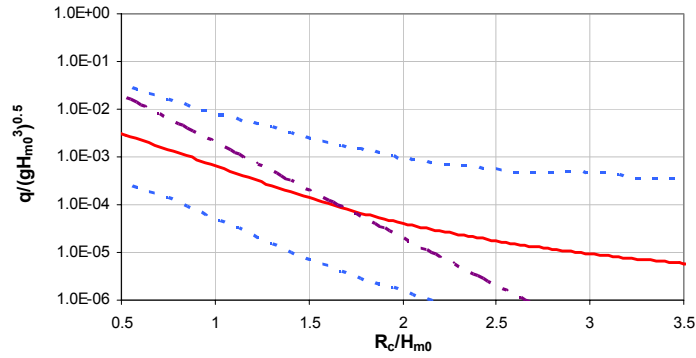


Example 3 Smooth structure

Figure 15 Sensitivity of wave overtopping to relative crest freeboard; Smooth structure

(— NN predictions; 95% confidence interval; - - TAW predictions)

Finally, Figure 16 shows the influence of an increase of the relative crest freeboard on the predictions of overtopping discharge for a rubble-mound structure with a berm and a crest element. The influence of the relative crest freeboard was also studied with the TAW guidelines for the design of dikes (TAW, 2002) and the influence of a berm.



Example 4 Rubble-mound structure with berm

Figure 16 Sensitivity of wave overtopping to relative crest freeboard; Rubble mound structure with berm

(— NN predictions; 95% confidence interval; - - TAW predictions)

These examples demonstrate that the NN is capable to detect from the data-set the influence of particular parameters for a wide range of structure types, together with an indication of the reliability of the NN predictions.

However, it should be noted that the quality of the predictions is, to a large extent, affected by the quality of the present database on which the NN is based.

6 Conclusions

A prediction method has been developed for the estimation of the mean wave overtopping discharge for many types of coastal structures, and the assessment of the uncertainties of these predictions.

The presented results show that Neural Networks can successfully be used to model the relationship between the input parameters involved in the process of wave overtopping and the mean overtopping discharge at coastal structures. The neural network model constructed herein is based on the database created within the framework of the European project CLASH, which includes tests collected from different universities and research institutes and new tests performed within the CLASH project to fill in white spots in the parameter domain. As for any NN, also the quality of the present NN is largely determined by the quantity and quality of the present database.

Resampling techniques have been used for the estimation of the uncertainties of the NN predictions. In general, it has been shown that the agreement between the predicted overtopping discharge and the measured overtopping discharge is good; the predictions are rather accurate compared to the observed overtopping discharges.

Acknowledgements

This study was co-sponsored by the Commission of the European Communities within the framework of the CLASH project ('Crest Level Assessment of Coastal Structures by full scale monitoring, neural network prediction and Hazard analysis on permissible wave overtopping', contract EVK3-2001-00058).

References

- Alexander, I. and Morton, H. (1991), *An introduction to neural network computing*. Chapman & Hall, London.
- Allsop, N.W.H., P. Besley and L. Madurini (1995). *Overtopping performance of vertical and composite breakwaters, seawalls and low reflection alternatives*. Paper 4.6 in Final Proceedings of MCS-project, MAS2-CT92-0047.
- De Rouck, J., Van der Meer, J.W., Allsop, N.W.H., Franco L. and Verhaeghe, H. (2002). *Wave overtopping at coastal structures: development of a database towards upgraded prediction methods*. ASCE, proc. 28th ICCE, Cardiff, UK, Vol. 2, pp. 2140-2152.
- Efron, B. (1982), *The Jackknife, the Bootstrap and Other Resampling Plans*, 92pp., SIAM, Philadelphia, PA.
- Efron, B., and Tibshirani, R. (1993), *An introduction to the Bootstrap*, London: Chapman & Hall.
- Franco L., de Gerloni M. and Van der Meer J.W. (1994). *Wave overtopping on vertical and composite breakwaters*. ASCE, proc. 24th ICCE, Kobe, Japan, pp. 1030-1045.
- Haykin, S. (1994), *Neural networks: a comprehensive foundation*. Macmillan College Publishing Company, Inc.
- Heskes, T. (1997), *Practical confidence and prediction intervals*. In Mozer, Jordan and Petsche (eds.), *Advances in neural information processing systems 9*, Cambridge, MIT Press.
- Mase, H., M. Sakamoto and T. Sakai (1995), *Neural network for stability analysis of rubble-mound breakwaters*. J. of Waterway, Port, Coastal and Ocean Engrg, ASCE, nov/dec 1995, pp. 294-299.
- Medina, J.R. (1999), *Neural network modelling of runup and overtopping*. ASCE, Proc. Coastal Structures 1999, Santander, Vol. 1, pp 421-429.
- Medina, J.R, J.A. González-Escrivá, J. Garrido and J. De Rouck (2002), *Overtopping analysis using neural networks*. ASCE, Proc. ICCE 2002.
- Medina, J.R, Garrido, J., Gómez-Martín, E. and Vidal, C. (2003). *Armour damage analysis using neural networks*. ASCE, Proc. Coastal Structures 2003, Portland (in press).
- Panizzo, A., R. Briganti, J.W. van der Meer and L. Franco (2003), *Analysis of wave transmission behind low-crested structures using neural networks*. ASCE, Proc. Coastal Structures 2003, Portland.
- Rumelhart, D.E. and McClelland, J.L. (1986), *Parallel distributed processing: Explorations in the microstructure of cognition*, Vol. 1, Cambridge, MIT Bradford Press.
- Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986a), *Learning representations by back-propagation errors*. Nature (London), Vol. 323, pp 533-536.
- Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986b), *Learning internal representations by error propagation*. In *Parallel distributed processing: Explorations in the microstructure of cognition* (Rumelhart, D.E. and McClelland, J.L., eds), Vol. 1, Chapter 8, Cambridge, MIT Bradford Press
- Stendam, G.J., J.W. van der Meer, H. Verhaeghe, P. Besley, L. Franco and M.R.A. van Gent (2004). *The international database on wave overtopping*. Proc. ASCE, 29th ICCE, Lisbon, Portugal.
- TAW (2002). *Technisch rapport golfoploop en golfoverslag bij dijken* (Technical report on wave run-up and wave overtopping at dikes - in Dutch). Technical Advisory Committee on Water Defences.

- Tichelaar, B.W. and Ruff, L.J. (1989), *How Good Are Our Best Models? Jackknifing, Bootstrapping, and Earthquake Depth*, EOS, Transactions, American Geophysical Union, Vol. 70: 593-606.
- Van den Boogaard, H., Mynett, A.E. and Heskes, T. (2000), *Resampling techniques for the assessment of uncertainties in parameters and predictions of calibrated models*. Proc. Hydroinformatics Conference, IOWA.
- Van Gent, M.R.A and H.F.P. van den Boogaard (1998), *Neural network modelling of forces on vertical structures*. ASCE, Proc. ICCE 1998, Vol. 2, pp. 2096-2109.
- Verhaeghe, H., J.W. van der Meer, G.-J. Steendam, P. Besley, L. Franco and M.R.A. van Gent (2003), *Wave overtopping database as the starting point for a neural network prediction method*. ASCE, Proc. Coastal Structures 2003, Portland.
- Wu, C.F.J. (1986), *Jackknife, bootstrap, and other resampling methods in regression analysis*, Ann. Stat., 14, 1261.