



UNIVERSITÀ POLITECNICA DELLE MARCHE
Repository ISTITUZIONALE

Prediction of non-breaking wave induced scour depth at the trunk section of breakwaters using Genetic Programming and Artificial Neural Networks

This is the peer reviewed version of the following article:

Original

Prediction of non-breaking wave induced scour depth at the trunk section of breakwaters using Genetic Programming and Artificial Neural Networks / Pourzangbar, Ali; Losada, Miguel A.; Saber, Aniseh; Ahari, Lida Rasoul; Larroudé, Philippe; Vaezi, Mostafa; Brocchini, Maurizio. - In: COASTAL ENGINEERING. - ISSN 0378-3839. - STAMPA. - 121:(2017), pp. 107-118. [10.1016/j.coastaleng.2016.12.008]

Availability:

This version is available at: 11566/247011 since: 2022-05-23T17:15:54Z

Publisher:

Published

DOI:10.1016/j.coastaleng.2016.12.008

Terms of use:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. The use of copyrighted works requires the consent of the rights' holder (author or publisher). Works made available under a Creative Commons license or a Publisher's custom-made license can be used according to the terms and conditions contained therein. See editor's website for further information and terms and conditions.

This item was downloaded from IRIS Università Politecnica delle Marche (<https://iris.univpm.it>). When citing, please refer to the published version.

note finali coverage

(Article begins on next page)

1
2
3
4 **Prediction of non-breaking wave induced scour depth at the trunk section of**
5
6
7 **breakwaters using Genetic Programming and Artificial Neural Networks**
8

9
10 Ali Pourzangbar ^{a,1}; Miguel A. Losada ^b; Aniseh Saber ^c; Lida Rasoul Ahari ^d; Philippe
11 Larroude ^e; Mostafa Vaezi ^f; Maurizio Brocchini ^g
12
13

14 ^a School of Civil Engineering, Iran University of Science and Technology (IUST), Tehran, Iran.

15 ^b Dinamica de Flujos Ambientales, Instituto Interuniversitario de Investigación del Sistema Tierra en Andalucía,
16 Universidad de Granada, Granada, Spain.
17

18 ^c Architecture department, School of engineering, University College of Nabi Akram, Tabriz, Iran.

19 ^d School of Electrical, IT and Computer sciences, Islamic Azad University of Qazvin, Iran.

20 ^e Laboratoire des Ecoulements Géophysiques et industriels, UFR Phitem, University of Grenoble, Grenoble, France.

21 ^f Department of maritime engineering, AmirKabir University of Technology, Tehran, Iran.

22 ^g Department of Civil and Building Engineering and Architecture (DICEA), Università Politecnica delle Marche,
23 Ancona, Italy.
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58

59 ¹ Corresponding author. Tel.: +989143065207; E-mail address: A_pourzangbar@sut.ac.ir ,
60 Pourzangbar.ali5@gmail.com (A. Pourzangbar)
61
62
63
64
65

1
2
3
4 **Abstract:**
5

6 Scour may act as a threat to coastal structures stability and reduce their functionality. Thus,
7 protection against scour can guarantee these structures' intended performance, which can be
8 achieved by the accurate prediction of the maximum scour depth. Since the hydrodynamics of
9 scour is very complex, existing formulas cannot produce good predictions. Therefore, in this
10 paper, Genetic Programming (GP) and Artificial Neural Networks (ANNs) have been used to
11 predict the maximum scour depth at breakwaters due to non-breaking waves (S_{\max}/H_{nb}). The
12 models have been built using the relative water depth at the toe (h_{toe}/L_{nb}), the Shields parameter
13 (θ), the non-breaking wave steepness (H_{nb}/L_{nb}), and the reflection coefficient (Cr), **where in**
14 **the case of irregular waves, $H_{nb}=H_{rms}$, $T_{nb}=T_{peak}$ and L_{nb} is the wavelength associated with the**
15 **peak period ($L_{nb}=L_p$).** 95 experimental datasets gathered from published literature on small-scale
16 experiments have been used to develop the GP and ANNs models. The results indicate that the
17 developed models perform significantly better than the empirical formulas derived from the
18 mentioned experiments. The GP model is to be preferred, because it performed marginally better
19 than the ANNs model and also produced an accurate and physically-sound equation for the
20 prediction of the maximum scour depth. Furthermore, the average percentage change (APC) of
21 input parameters in the GP and ANNs models shows that the maximum scour depth dependence
22 on the reflection coefficient is larger than that of other input parameters.
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

40 **Keywords:**
41

42 Scour; Non-breaking waves; Genetic Programming (GP); Artificial Neural Networks
43 (ANNs); Breakwater; Uncertainty assessment
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1. Introduction

Coastal structures such as breakwaters are constructed to protect harbors and vessels from wave attacks. Proper and optimum initial design of these structures can eliminate the main construction problems, such as the structure instability, which could cause significant unforeseen expenditure. Therefore, optimizing the design of coastal structures is fundamental.

Scour, which may act as a threat to the stability and functionality of **marine** structures, is one of the main reasons for the failure of coastal (Hughes and Fowler, 1991; Lillycrop and Hughes, 1993; Oumeraci, 1994; Whitehouse, 1998) and **offshore** (e.g. **Mattioli et al., 2012, 2013**) structures. Therefore, protecting structures against scour is critical in the construction of well-functioning man made harbors. To do this, the accurate prediction of maximum scour depth at coastal structures has inevitable importance. Although several studies have been conducted on scour at coastal structures, the complexity of onshore hydrodynamic and complex interaction between incoming waves, bed sediments and structure has impeded the accurate maximum scour depth prediction. Scour at breakwaters or seawalls (vertical or inclined) can be categorized into two main classes: scour at the head of coastal structures; and scour at the trunk section of coastal structures (due to breaking or non-breaking waves). Since the present paper focuses on predicting of the maximum scour depth at breakwaters due to non-breaking waves (hereafter S_{max}), only the non-breaking wave-induced scour depth at the trunk section of coastal structures has been discussed here. It is noted that S_{max} is the ultimate value of scour depth when the equilibrium bottom profile is reached and it is independent of time.

Scour at inclined and vertical breakwaters due to non-breaking waves was investigated in several studies based on small-scale experiments. Sawaragi (1966) and Baquerizo and Losada (1988) investigated the relation between the wave reflection and the equilibrium scour depth at a rubble-mound breakwater and suggested that the scour depth becomes larger with the increase of the reflection coefficient (Cr). Similarly, using small-scale experiments, Oumeraci (1994) studied the effect of breakwater slope on S_{max} and suggested that the maximum scour depth in front of a vertical breakwater is larger than that at sloped breakwaters. Furthermore, he indicated that the key mechanism for scour due to non-breaking waves is the action of standing waves (fully or partially), which leads to a steady streaming pattern. Carter et al. (1973) investigated the regular and irregular wave-induced scour depth at vertical breakwaters, and showed that the scour

1
2
3
4 and deposition pattern in front of the vertical breakwaters emerges in the form of alternating scour
5 and deposition developing parallel to the shoreline. This finding has also been obtained by
6 Baquerizo and Losada (1988).
7

8
9 Soft computing approaches like Artificial Neural Networks (ANNs) and Genetic
10 Programming (GP) have been successfully employed for the prediction of scour depth in various
11 fields of coastal engineering, such as the estimation of scour depth below free overfall spillways
12 (Samadi et al., 2015), the estimation of scour around submarine pipelines (Kızıloz et al, 2015),
13 the prediction of scour depth under live-bed conditions at river confluences (Balouchi et al.,
14 2015), the prediction of scour depth in bridges (Akib et al., 2014), the prediction of scour at a
15 bridge abutment (Azamathulla, 2012), the determination of the most important parameters on
16 scour at coastal structures (Yeganeh-Bakhtiary et al., 2012; Pourzangbar, 2012), the study of
17 scour below submerged pipeline (Azamathulla et al., 2011). Regarding the mentioned studies,
18 GP and ANNs can predict scour depth at coastal structures with high precision, and, to the best
19 knowledge of the authors, these approaches have not been implemented in the prediction of the
20 S_{max} . Therefore, ANNs and GP have been used in this study as robust and promising tools.
21
22 Furthermore, GP is capable of producing physically-sound and accurate solutions in the form of
23 mathematical equations. Using this capability of GP, a new formula was developed for the
24 prediction of S_{max} .
25
26
27
28
29
30
31
32
33
34
35
36
37

38 This study is structured as follows: Section 2 shows the overview of scour governing
39 parameters; section 3 presents ANNs and GP concepts. The modeling approach and the data at
40 the basis of the analyses are reported in section 3; the results and discussions are given in section
41 4; the sensitivity analysis is given in section 5 and finally section 6 contains this study summary
42 and the conclusion.
43
44
45
46
47

48 **2. Scour governing variables and formulas**

49 Scour at the trunk section of breakwaters due to non-breaking waves depends on three
50 classes of parameters: the wave characteristics, the sediment properties and the breakwater
51 configuration. Several small-scale experimental studies are available that provide useful
52 information about the governing parameters of scour at breakwaters. Among the most important
53 experimental studies that also led to empirical formulas to predict scour depth we find the
54 following.
55
56
57
58
59
60
61
62
63
64
65

Xie (1981) examined the scouring profile of a bed consisting of fine and coarse sediments at a vertical breakwater. Xie (1981)'s experiments were performed using four different sediment diameters $d_{50}=0.106, 0.150, 0.200,$ and 0.780 mm, d_{50} being the mean diameter of bed sediments. The fall velocities (V_s) associated with $d_{50}=0.106, 0.150, 0.200,$ and 0.780 mm are $V_s= 0.7, 1.5, 2.2,$ and 11 cm/s., respectively. He showed that the scour profile is utterly different for fine and coarse sediments depending on the waves' characteristics. In the case of fine material (suspension mode of sand transport), the bed sediments move in suspension from the node towards the antinode, while in the case of relatively coarse sand (bedload inception), the sediments transport are governed by the bed shear and scour occurs halfway between the node and the antinode, and deposition at the node, this finding is in line with De best et al. (1971) results. Xie proposed the following threshold for bedload inception:

$$\frac{U_{\max} - U_{cr}}{V_s} \leq 16.5 \quad (1)$$

where U_{\max} is the maximum value of the orbital velocity at the bed, U_{cr} is the critical velocity for initiation of the bed sediments motion and V_s is the sand grain fall velocity. The bed sediments are transported in suspension mode (fine sediments) when $U_{\max} - U_{cr} / V_s > 16.5$.

Based on the results of mobile-bed flume experiments, Xie (1981) proposed Eq. (2) for the prediction of the maximum scour depth at vertical breakwaters:

$$\frac{S_{\max}}{H_{nb}} = \frac{C}{\left(\sinh \frac{2\pi h}{L_{nb}} \right)^{1.35}} \quad (2)$$

where S_{\max} is the maximum scour depth (the ultimate value of scour depth when the equilibrium bottom profile is reached), H_{nb} is the non-breaking wave height (both regular and irregular waves analysed), h is the still water depth in deepwater and L_{nb} is the non-breaking wave length (regular or irregular), and $C=0.3$ (suspension mode of sand transport) for fine sediments and $C=0.4$ (bedload mode of sand transport) for coarse sediments. In the following the subscript "nb" means "non-breaking" and it is employed to clarify that this study focuses on the scour induced by

non-breaking waves. In the case of random waves or irregular waves, $H_{nb}=H_{rms}$, $T_{nb}=T_{peak}$ and L_{nb} is the wavelength associated with the peak period ($L_{nb}=L_p$).

Eq. (2) was proposed to describe the action of fully standing waves condition, thus the effect of the structural configuration, such as the breakwater slope, was not accounted for. Although this formula is very limited in application, it was the basis for subsequent investigations. To amend Xie (1981)'s formula deficiencies, Sumer and Fredsøe (2000) conducted some wave flume small-scale experimental studies on scour at rubble-mound and vertical breakwaters. They concluded that the wave reflection is the most important phenomenon accounting for the effects of breakwater slope and structural configuration, and, thus, suggested the following empirical equation to predict the maximum scour depth at a vertical or a rubble-mound breakwater:

$$\frac{S_{\max}}{H_{nb}} = \frac{0.3 - 1.77 \exp\left(-\frac{\alpha}{15}\right)}{\left(\sinh \frac{2\pi h}{L_{nb}}\right)^{1.35}} \quad (3)$$

where α is the breakwater slope in the range of 30 - 90 degrees. Eq. (3) includes the effect of the breakwater slope, while the effects of structural configuration, breakwater submergence and permeability, sediments properties and bed slope were not accounted for.

Lee and Mizutani (2008) investigated the scour at vertical submerged breakwaters. In agreement with Sumer and Fredsøe (2000)'s findings, they introduced the reflection coefficient as the main parameter affecting the scour depth. Their proposed formula is as follows:

$$\frac{S_{\max}}{H_{nb}} = \frac{0.06}{(1 - Cr) \left(\sinh \frac{2\pi h}{L_{nb}}\right)^{2.04}} \quad (4)$$

where Cr is the reflection coefficient.

This equation accounts for the structural configuration, the relative water depth at the toe of the breakwater and the wave height. However, it cannot be used to predict the scour depth for fully

standing waves since it diverges when $Cr = 1$. Furthermore, it does not give account of the sediment-waves interactions, since sediment properties do not appear in (4).

The aforementioned studies experimentally investigated the non-breaking wave-induced scour at breakwaters and developed regression-based empirical formulas to predict S_{max} . However, the proposed equations do not include the effects of all important parameters for scouring; also they do not have adequate accuracy and wide applicability in predicting S_{max} . Hence, developing accurate and robust models that describe the effects of all important parameters on the maximum scour depth can be very imperative.

Since the non-breaking wave-induced scour at breakwaters is the result of a complicated interaction between waves, structure, and bed sediments, the waves' characteristics, the bed sediments properties, and the structural configuration are of great importance in the accurate prediction of scour depth. Accordingly, the non-breaking wave-induced maximum scour depth (S_{max}), the ultimate value of scour depth when the equilibrium bottom profile is reached, can be expressed by the following functional relationship:

$$S_{max} = f(\alpha, h_c, B, D_{50}, P, H_{nb}, L_{nb}, T_{nb}, h_{toe}, \rho, \nu, \rho_s, d_{50}, \beta) \quad (5)$$

where α is the structure slope, h_c is the submerged depth of breakwater, B is the breakwater width, D_{50} is the median size of breakwater armor stones and P is the permeability of the breakwater. These parameters account for the effects of breakwater configuration on the scour depth. The non-breaking wave height (H_{nb}), the non-breaking wave length (L_{nb}), and the wave period (T_{nb}) represent the incoming waves' characteristics. H_{nb} gives the non-breaking wave height for both regular waves (H_0) and random waves (in the case of random waves $H_{nb} = H_{rms}$, $T_{nb} = T_{peak}$, $L_{nb} = L_{peak}$). The water depth at the toe (h_{toe}), the mass density of water (ρ), the kinematic viscosity (ν), the specific gravity of sediments (ρ_s), the mean diameter of bed sediments (d_{50}), and the bed slope (β) indicate the effects of bottom profile and the fluid and bed sediments properties on S_{max} . For a better understanding, the variables appearing in Eq. (5) are schematically shown in **Fig. 1**.

1
2
3
4 Sumer and Fredsøe (2000) attributed the difference between the maximum scour depth of
5 vertical and rubble-mound breakwaters to the wave reflection. Lee and Mizutani (2008)
6 indicated that the effects of submerged depth, breakwater width and structural configuration can
7 be effectively described by the reflection coefficient. Therefore, in this paper, the reflection
8 coefficient (Cr) accounts for the effects of structural configuration, such as the breakwater slope
9 (α), the submerged depth (h_c), the breakwater width (B), the size of armor stones (D_{50}) and the
10 permeability of the breakwater (P). Sumer and Fredsøe (2000) indicated that the empirical
11 formula of Losada and Gimenez-Curto (1981), Eq. (6), can be successfully employed in the
12 prediction of the reflection coefficient. **Therefore, in the present paper the reflection coefficients**
13 **of Xie (1981), and Sumer and Fredsøe (2000) datasets are calculated based on the following**
14 **equation:**
15
16
17
18
19
20
21
22
23
24
25
26

$$27 \quad Cr = 1.35 \left(1 - \exp\left(-0.071 \left(\frac{\tan \alpha}{\sqrt{2\pi H_{nb}/gT_{nb}^2}} \right) \right) \right) \quad (6)$$

28
29
30
31
32
33 in which g is gravity acceleration.

34
35 The key mechanism for scour at breakwaters due to non-breaking waves is the action of
36 standing waves (fully or partially), which leads to a steady streaming pattern and recirculating
37 cells (Xie, 1981, Sumer and Fredsøe, 2000). The water depth at the toe of the breakwater (h_{toe})
38 can adversely affect the steady streaming strength and the bed sediments transport capability, in
39 a way that an increase in the water depth will decrease the steady streaming and recirculating
40 cells strength, and so leads to the scour depth decrease. In this study, the relative water depth
41 normalized by the non-breaking wave length (h_{toe}/L_{nb}) has been used to account for the effects
42 of the water depth on the S_{max} . The Reynolds number (Re) and the Shields parameter (θ)
43 account for the wave-seabed interaction and the sediment transport, respectively. According to
44 Sumer and Fredsøe (2000), since the bed acts as a rough wall in most coastal engineering
45 problems ($Re \geq 3$), the effects of Re are taken to be unimportant and, thus, neglected in this
46 study. However, the effects of seabed and bed sediments properties on the maximum scour depth
47 (S_{max}) have been accounted for through the Shields parameter ($\theta \geq 0.07$). Hajivalie et al. (2008)
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 studied the effect of the bed slope on the pattern of standing waves and indicated that bed slope
5 changes the size of steady streaming and recirculating cells. Therefore, the bed slope can be one
6 of the fundamental parameters governing the scour at breakwaters due to non-breaking waves.
7
8 However, no experimental or field data that describes the effects of bed slope on scour process is
9 available, and, thus, it is impossible to actually describe its role in the present modeling.
10
11
12

13 Results of various studies showed that an increase in the wave height intensifies the
14 sediment transport capability and ultimately leads to the large scour depths (Sumer and Fredsøe,
15 2000, Lee and Mizutani, 2008). Here, the relative non-breaking wave steepness (H_{nb}/L_{nb})
16 accounts for the wave characteristics effects on S_{max} . According to the above, Eq. (5) can be cast
17 in dimensionless form as follows:
18
19
20
21
22
23

$$24 \frac{S_{max}}{H_0} = f \left[Cr, \frac{H_{nb}}{L_{nb}}, \frac{h_{toe}}{L_{nb}}, \theta \right] \quad (7)$$

25
26
27
28
29 where

$$30 \theta = \frac{U_{fm}^2}{g \left(\frac{\rho_s}{\rho} - 1 \right) d_{50}} \quad (8)$$

$$31 U_{fm} = \frac{H_{nb}}{2} \cdot \frac{1}{\sinh \frac{2\pi h_{toe}}{L_{nb}}} \cdot \frac{2\pi}{T} \sqrt{\frac{f_w}{2}} \quad (9)$$

32
33
34
35
36
37
38
39
40
41
42
43 where S_{max}/H_{nb} is the maximum scour depth normalized by the non-breaking wave height, U_{fm}
44 is the maximum shear velocity, U_{max} is the maximum horizontal velocity of water particles just
45 above the wave bottom boundary layer, f_w is the wave friction coefficient, and ω is the wave
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65 angular frequency.

3. Soft computing approaches

3.1. Genetic Programming (GP)

GP is borrowed from the process of evolution occurring in the nature (survival of the fittest). GP employs a “parse tree” structure for the search of its solutions, which are continually evolving and never fixed. Unlike the most soft computing approaches, like ANNs, GP solutions are in the form of tree structure, mathematical equations or computer programs (see Fig. 2). Furthermore, there is no assumption made on the structure of the relationship between the independent and dependent variables, and, therefore, GP can produce very accurate solutions.

In order to evolve a model with GP, five preparatory steps must be taken, as follows:

Step 1: determining the terminal set, including the independent and the dependent variables, random coefficients and constant values;

Step 2: choosing the appropriate functional set (i.e. arithmetic operators (+, -, \times , \div) and mathematical functions (such as sin, tan, Exp, Sqrt)). This step is challenging, in a way that inappropriate functional set selection may result in programs that are not physically sound. In this paper the functional set are selected based on the trial and error procedure. This may lead to subjective overlearning, which can be minimized on the basis of an appropriate uncertainty assessment analysis (see section 6).

Step 3: selecting the fitness criteria to evaluate the individuals' accuracy. This criterion determines the individuals to be selected as parents and survived into the next generation. Here, the Root Mean Square Error (RMSE) has been used as fitness criterion.

Step 4: choosing the controlling parameters like the gene linking function, the genetic operators' rate, and chromosomes. These parameters can control the models' size and accuracy. One of the problems regarding GP models is the Bloat phenomenon, for which the program size increases without any corresponding improvement in the models' fitness. This problem results in nested models that are hard to interpret and computationally expensive (Poli and McPhee, 2008). The limitation of parse tree depth by applying of Parsimony Pressure may be regarded as a proper method as described by Poli and McPhee (2008). Parsimony pressure is a general family of methods that considers the size as part of the selection process and can be used to control the sizes of evolved models in genetic programming. This method effectively treats the minimization of size as a soft constraint and attempts to enforce this constraint using the penalty method, i.e.,

1
2
3
4 by decreasing the fitness of programs by an amount that depends on their size. The penalty is,
5 typically, simply proportional to the program size. The intensity with which bloat is controlled is,
6 therefore, determined by one parameter called the “parsimony coefficient” (Poli and McPhee,
7 2008). Here, the parsimony pressure has been implemented to the fitness function (Root Mean
8 Square Error (RMSE)). In other words, in this study the selection criteria are based on the fitness
9 function and individuals’ size.

10
11
12
13
14
15 Step 5: determining the termination condition. This parameter can be a specific number of
16 generations or a specific accuracy. Table 1 indicates the functional set, controlling parameters
17 and termination condition used in this study.

18
19
20
21 After accomplishing the preparatory steps, GP observes the following step-by-step procedure
22 to produce the final result, called solution. As indicated in Fig. 3, in the first place, GP creates
23 the initial models (population) using a random combination of the terminal and functional sets.
24 Each model or individual of this initial generation is evaluated based on the statistical criterion
25 such as RMSE to be selected as a parent.

26
27
28
29
30 Then, GP utilizes selection methodologies, such as ranking, in which individuals are ranked
31 and selected out according to their fitness value, or tournament, in which all of the individuals
32 regarded as a “gene pool”, and the fittest models are selected out by randomly picking up a
33 certain number of individuals for several times, to select the parents. Applying genetic operators
34 like crossover, mutation or reproduction, new individuals (offspring) are produced (Figs. 4a and
35 b). Comparing the created models accuracy or generations number with the termination
36 condition, the GP states when to terminate the modeling process. If the termination criterion,
37 such as the maximum number of generation or the fitness function is not satisfied, the modeling
38 process continues by selecting new parents, creating new generation, and comparing the created
39 generation with terminal condition. Otherwise, the best individual is known as the best solution.

3.2. Artificial Neural Networks (ANNs)

40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
ANNs provide a random mapping between an input vector and an output one. An ANN consists of a set of neurons, the fundamental processing element of a neural network, arranged in input, hidden, and output layers. In order to produce output vector, each neuron combines inputs and estimates their weights. Then they produce an output bypassing the summed values of all

1
2
3
4 nodes through transfer functions. The number of input layer neurons equals the input parameters,
5 and the number of neurons in the output layer equals the number of output parameters. However,
6 determining the number of neurons in hidden layers that controls the accuracy of prediction is a
7 challenging issue. Rogers and Dowla (1994) suggested the following criteria for the
8 determination of hidden layer neurons number:
9

$$14 \quad N^H \leq 2N^L + 1 \quad (10a)$$

$$17 \quad N^H \leq \frac{N^{TR}}{N^L + 1} \quad (10b)$$

20
21 where N^H stands for the number of hidden layer neurons, N^L is the number of input parameters
22 (here $N^L = 4$) and N^{TR} stands for the number of training datasets (here $N^{TR} = 68$). According to
23 Eqs. (10a and b), the number of hidden layer neurons must be less than 9 ($N^H \leq 9$) for this study.
24 A trial and error procedure is utilized to achieve the best configuration of neural network
25 employed. Accordingly, a three-layer feed forward network with Levenberg-Marquardt back
26 propagation training algorithm, which is the most commonly used neural network in many
27 studies (Jain and Deo, 2007), is utilized for the prediction of S_{max} . The learning rate and the
28 iteration, resulted from the trial and error process, were 0.01 and 1000, respectively. The Log-
29 Sigmoid function was also employed in the optimum network as a transfer function. Fig. 5 shows
30 the structure of artificial neural network which has been used in this study with 4 input
31 parameters in input layer, 5 neurons in hidden layer, and 1 output parameter. Further information
32 on ANNs can be found in published literatures.
33
34
35
36
37
38
39
40
41
42
43
44

46 **4. Datasets and modeling**

47 *4.1. Datasets description*

48
49
50 To evolve the GP model, a combination of Xie (1981), Sumer and Fredsøe (1981, 2000) and
51 Lee and Mizutani (2008) published datasets was used. All such data (95 data points), as well as
52 the related finding come from small-scale flume experiments. It is clear that all findings derived
53 from the mentioned datasets, as it happens for all studies based on small-scale experiments, can
54 be influenced by scale effects and their use for practical applications should be made with some
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 caution. Further, being based on flume experiment (2D flow) all results are only adequate to
5 describe dynamics evolving along vertical transects (typically normal to shore).
6

7
8 The data in use refer to both regular and irregular non-breaking wave-induced scour at
9 vertical and rubble-mound breakwaters, and the maximum scour depth is the ultimate value of
10 scour depth when the equilibrium bottom profile is reached. The dataset from Xie (1981)
11 contains 39 data records. 23 experimental data records were selected from the Sumer and
12 Fredsøe (2000)'s work and 33 data records were collected from the Lee and Mizutani (2008)'s
13 laboratory study.
14
15
16
17

18
19 Of the 95 experimental datasets 70% (65 data points) was used for the model's training and
20 30% for the model's testing. Table 2 shows the range of datasets used in this paper. Details of the
21 mentioned experimental studies are given in the following.
22
23
24
25
26
27

28 *4.1.1 Xie (1981) experimental study*

29

30
31 Xie (1981) studied experimentally the scour pattern in front of a vertical breakwater under
32 the action of standing waves. His experiments were conducted using two wave flumes: 1) a small
33 flume - 38 m long, 0.8 m wide, and 0.6 m deep - and 2) a large flume - 46 m long, 0.8 m wide,
34 and 1 m deep. The sandy bed, made of four different types of sand sizes (0.106, 0.15, 0.20 and
35 0.78 mm), was 0.15 m thick. The wave paddle was located at a distance of 32.9 m from the
36 vertical breakwater. Most runs were made with regular waves and only some with random
37 waves. Incident wave heights ranged from 4.5cm to 11 cm, wave periods were in the 1.1 to 3.56
38 s range and water depths were of 30, 40 and 50 cm. From this experiment, 39 data records were
39 selected to evolve the models.
40
41
42
43
44
45
46
47

48 *4.1.2 Sumer and Fredsøe (2000) experiment*

49

50
51 Sumer and Fredsøe (2000) studied the non-breaking wave induced scour along the trunk section
52 of vertical and rubble-mound breakwaters under regular and irregular wave conditions. The
53 experiments were conducted in a wave flume of 0.6 m width, 0.8 m depth and 28 m length.
54 Waves were generated by a piston-type wave generator located at a distance of 22 m from the
55 model breakwater. The water depth was of 0.31 m, and the sand size was of 0.2 mm. Two
56
57
58
59
60
61
62
63
64
65

1
2
3
4 breakwater slopes were implemented 1:1.2 ($\alpha = 40^\circ$) and 1:1.75 ($\alpha = 30^\circ$) and the results were
5 compared with those for the vertical breakwater made of a plywood plate. The surface elevation
6 was measured by a wave gauge, located at 18 m from the wave generator; and the scour
7 development was monitored by visual observations accompanied by scour measurements
8 undertaken at certain points and small time intervals.
9

10 11 12 13 14 15 *4.1.3 Lee and Mizutani (2008) experiment*

16
17 Lee and Mizutani (2008) studied experimentally the scour due to regular waves at a vertical
18 impermeable submerged breakwater that induced partial reflection of the incident waves and, as
19 a consequence, a field of partially standing waves. They conducted their experiments in a flume
20 with 0.7 m width, 0.9 m, depth and 30 m length. A flap-type wave generator and a wave
21 absorption layer were located at the ends of the wave flume. The breakwater was placed at 16.5
22 m from the wave paddle, and was seated on the bottom of the wave flume. The sandy bed was 2
23 m long, 0.2 m thick and 0.7 m wide, and the median sand diameter was of 0.2mm. The wave
24 steepness (H_{nb}/L_{nb}) ranged from 0.019 to 0.052, and the wave period (T_{nb}) varied in the range
25 1.1s to 1.9 s. The still water depth above the movable sand (h_{toe}) was of 20 and 30 cm.
26
27 Suspended load transport was found to be negligible.
28
29
30
31
32
33
34
35
36
37

38 *4.2. Statistical error parameters*

39
40 The performance of the GP and ANNs models in the prediction of S_{max}/H_{nb} for all the
41 datasets of interest has been evaluated in terms of statistical parameters like the Correlation
42 Coefficient (CC), the Root Mean Square Error ($RMSE$), the Scatter Index (SI) and the $BIAS$, as
43 given in Eqs. (11)-(14). The parsimony pressure is a selecting criterion, but the CC , SI and $BIAS$
44 have been used to evaluate the accuracy of the final and selected models. To have an
45 understandable evaluation between the GP and ANN models and the existing equations, SI and
46 $BIAS$ are of great importance. However, the parsimony pressure is only applicable to GP models
47 in the stage of selection.
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

$$CC = \frac{\sum_{i=1}^N (O_i - \overline{O_m})(P_i - \overline{P_m})}{\sqrt{\sum_{i=1}^N (O_i - \overline{O_m})^2 \times \sum_{i=1}^N (P_i - \overline{P_m})^2}} \quad (11)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (P_i - O_i)^2}{N}} \quad (12)$$

$$SI = \frac{\sqrt{\frac{\sum_{i=1}^N (P_i - O_i)^2}{N}}}{\overline{O_m}} \times 100\% \quad (13)$$

$$BIAS = \frac{\sum_{i=1}^N (P_i - O_i)}{N} \quad (14)$$

Here O_i and P_i denote the observed and predicted values, respectively. N is the number of observed data and $\overline{P_m}, \overline{O_m}$ are the corresponding mean values of the predicted and observed parameters, respectively.

5. Results and discussion

5.1. Evaluation of the existing formulas

Best known empirical formulas for the prediction of the maximum scour depth are those of Xie (1981), Sumer and Fredsøe (2000), and Lee and Mizutani (2008) (Section 1). In this section the performance of these formulas in the prediction of the maximum scour depth for the different datasets (Section. 3.1) has been investigated on the basis of statistical error parameters. Fig. 6a, b, and c show that the performance of the empirical formulas of Xie (1981), Sumer and Fredsøe (2000), and Lee and Mizutani (2008) is fair in predicting the maximum scour depth for their own dataset. However, Fig. 6d illustrates the performances of the formulas of Xie (1981) and Sumer and Fredsøe (2000) in the prediction of the scour depth using all the data available (95 data

1
2
3
4 points). As indicated, such empirical equations do not have adequate accuracy for such a global
5 dataset. Furthermore, Xie (1981)'s formula (Eq. 2) predictions in some cases are larger than 1,
6 while the maximum scour depth is by definition smaller than 1. Although Lee and Mizutani
7 (2008) formula leads to completely scattered values for the maximum scour depth as function of
8 the reflection coefficient (e.g. under the same waves characteristics for $Cr=0.9$ and $Cr=0.8$,
9 the maximum scour depth coefficient changes from 10 to 5, respectively) and it cannot be
10 implemented in predicting the scour depth for all data set, since it diverges when $Cr=1$, it
11 shows less scatter and high precision in predicting the maximum scour depth for its own
12 experimental data set. The higher performance of Lee and Mizutani (2008) formula is mainly
13 due to the use of the reflection coefficient (Cr) as an input parameter.
14
15
16
17
18
19
20
21
22

23 Table 3 shows the error measures of the existing formulas in predicting the maximum scour
24 depth for various datasets. It can be seen that the CC , $RMSE$, SI , and $BIAS$ of Lee and
25 Mizutani (2008) formula - 0.952, 0.069, 23.67% and -0.054, respectively - are less than those of
26 other empirical formulas in the prediction of their own datasets. However, considering accuracy
27 and width of applicability, the empirical equations cannot be implemented for the prediction of
28 the maximum scour depth for all the datasets used in this paper. For instance, the prediction of
29 the maximum scour depth of all the available data by means of Xie (1981)'s formula (Eq. 2) is
30 not fair, the CC , $RMSE$, SI , and $BIAS$ being 0.528, 0.286, 77.9% and 0.154, respectively. To
31 improve the accuracy and width of applicability, a set of new models has been developed by
32 means of the ANNs and GP approaches (Section 5).
33
34
35
36
37
38
39
40
41
42
43

44 *5.2. Development of the ANNs and GP models*

45
46 A total of 95 regular and irregular non-breaking wave dataset has been used for the
47 prediction of the maximum scour depth using the ANNs and GP models. Different combinations
48 of dimensionless parameters have been considered in developing the GP and ANNs models to
49 achieve the most accurate and parsimonious models. Three different structures (Table 4) have
50 been considered to develop the models with the most important input parameters.
51
52
53
54
55

56 The first step in developing a model with ANNs is to determine the architecture and
57 configuration of the used ANNs model, these being discussed in Section 2. By increasing the
58 number of hidden neurons the performance of ANNs improves when using training data.
59
60
61
62
63
64
65

1
2
3
4 However, at certain number of hidden neurons ANNs produce noisy fluctuations with testing
5 data and “overtraining” occurs (Van Ghent et al., 2007). To avoid this problem, the number of
6 neurons in the hidden layer has been determined based on training the ANNs with a varying
7 number of neurons in the hidden layer and each time comparing the ANN’s performance on the
8 training and testing data sets. To achieve the highest possible accuracy and avoid the
9 overtraining problem, the optimum number of hidden layer neurons has been determined by a
10 trial and error approach (see Table 5).
11
12
13
14
15
16

17 Table 5 shows that the various ANNs models which have been developed based on the
18 different input parameters (S1, S2, and S3 model structures). As indicated, the ANNs model with
19 5 neurons in its hidden layer and the S3 model structure input parameters, including the wave
20 reflection (Cr), the non-breaking wave steepness (H_{nb}/L_{nb}), the relative water depth at the toe (h_{toe}/L_{nb}), and the Shields parameter (θ), results in the best performance amongst the tested
21 models with $CC=0.971$ and $RMSE=0.054$ in the prediction of training dataset used in this
22 paper.
23
24
25
26
27
28
29
30
31

32 At times the accuracy of models for testing data set is higher than that of training data set.
33 This issue relates to the model’s configuration such as the number of neurons in hidden layer,
34 and also to the modeling approach, here ANN.
35
36
37

38 Similar to ANNs, GP produced its best results for the S3 model structure input parameters
39 (Table 6), this model is referred as the best GP model (hereafter S3-GP), and all of the
40 contributing parameters in Eq. (7) play an important role in the prediction of the maximum scour
41 depth and cannot be neglected in the modeling process. Fig. 7 illustrates the scattered diagram of
42 measured and predicted the relative maximum scour depth (S_{max}/H_{nb}) for the case of S3-
43 ANN(5) and S3-GP. As indicated in this figure, the high accuracy of the S3-ANN(5) and S3-GP
44 models is reflected in the data points’ short distance from the fit line and their low scatter index.
45
46
47
48
49
50

51 As indicated in Table 3, both developed models predict the maximum scour depth for
52 various datasets with high precision in comparison to the existing empirical formulas. However it
53 appears that S3-GP performs marginally better than S3-ANN(5) in predicting of S_{max}/H_{nb} for all
54 cases, the main advantage of the GP on other soft computing methods, such as ANNs, being its
55 ability to produce explicit and accurate equations for the prediction of the phenomena at hand.
56
57
58
59
60
61
62
63
64
65

Using this capability, the following formula has been developed, by means of S3-GP, for predicting the non-breaking wave-induced maximum scour depth:

$$\frac{S_{\max}}{H_{nb}} = \frac{\sinh Cr}{(\theta - 1.223)^2 + 60.84(h_{toe}/L_{nb})^2} + \tan\left[\left(h_{toe}/L_{nb}\right) \cdot \tan\left(\frac{0.097Cr}{h_{toe}/L_{nb}}\right)\right] + \frac{H_{nb}}{L_{nb}} - \tan\left(\frac{h_{toe}}{L_{nb}}\right) \sqrt[3]{6.55\left(Cr + \frac{h_{toe}}{L_{nb}}\right)} \quad (15)$$

Unlike the empirical formulas of Xie (1981), Sumer and Fredsøe (2000), and Lee and Mizutani (2008), the S3-GP developed equation (Eq. (15)) includes the effects of the structural configuration, the bed sediments' properties and the wave characteristics, and it can be applied in predicting of the maximum scour depth of the vertical, sloped, and submerged structures under the action of fully or partially standing waves. The values of the reflection coefficient for Xie (1981), and Sumer and Fredsøe (2000) datasets, not measured during their experiments, were calculated using Eq. (6). On the opposite, the values of Cr for the Lee and Mizutani (2008) dataset were measured during their experiments, hence no use is here made of (6) for such Cr . The reflection coefficient measured by Lee and Mizutani (2008), and appearing in Eq. (15), includes the effects of the breakwater slope (α), the breakwater submergence (h_c), the breakwater width (B), the size of the armor stones (D_{50}) and the permeability of the breakwater (P). Although it may appear that Eq. (15) cannot give proper account of the effects of bottom slope, this is actually modeled through the role of Cr and h_{toe}/L_{nb} , see also the analysis of Hajivalie et al. (2008) on the role of the bottom slope in altering the steady streaming during wave shoaling.

As already mentioned the findings here proposed are characterized by limitations related with both scale phenomena and geometry. In particular, the 2D geometry and the wave forcing in use suggest the proposed results be taken as reliable to describe the maximum scour depth induced by non-breaking waves at the vertical trunk section of breakwaters.

The mathematical functions used in Eq. (15) like \sinh , \tan , $\sqrt[3]{\quad}$, $\sqrt{\quad}$ are the functions that have been used in the above-mentioned available formulas (Section 3). The main advantages of the proposed GP model are that it contains all of the most important parameters affecting the

1
2
3
4 maximum scour depth at breakwaters, has wide applicability, and it is very accurate in predicting
5 the maximum scour depth in comparison to the mentioned existing formulas. However, it is
6 fairly complex in its use and interpretation, such deficiencies being easily resolved by numerical
7 computing.
8
9

10
11 The proposed equation (15) is in good agreement with both the engineering sense and the
12 existing knowledge of the scour process. Fig. 8 shows the dependence of the maximum scour
13 depth predicted by Eq. (15) on fundamental input parameters, the others being kept constant. As
14 discussed by Sumer and Fredsøe (2000) and Lee and Mizutani (2008), among others, the
15 maximum scour depth increases with an increase in the reflection coefficient, the Shields
16 parameter and the non-breaking wave steepness, and decreases when the relative water depth at
17 the toe of the breakwater increases. As expected, Fig. 8 displays that the maximum scour depth is
18 directly proportional to the wave reflection, the non-breaking wave steepness, and the Shields
19 parameter. However, it decreases when the relative water depth at the toe of the breakwater
20 increases. Furthermore, Fig. 8 illustrates that the relationship between S_{\max}/H_{nb} and θ is almost
21 linear, which shows that the Shields parameter is not as important as Cr or h_{toe}/L_{nb} (see the
22 nonlinear dependence). Similarly, the trend between S_{\max}/H_{nb} and H_{nb}/L_{nb} is almost linear. This
23 is in line with the existing understanding of the relative importance of the parameters governing
24 the maximum scour depth at breakwaters due to non-breaking waves (Section 6).
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

40 **6. Uncertainty and reliability assessment**

41
42 After developing the final models with ANNs and GP, their predictions for other test data
43 sets may be biased and it is possible that the models cannot perform adequately well for the
44 available data spread over the entire domain of data set. To have more trustworthy models, a
45 resampling technique, the K -fold cross validation, has been applied to the data set as a whole. In
46 this technique, the whole data set is randomly partitioned into K equal-sized folds; $K-1$ folds are
47 used for training and the remaining one is used for testing. This process is repeated K times, with
48 each of the K folds used exactly once as the testing data set. The advantage of the K -fold cross
49 validation is that all the examples in the data set are used for both training and testing (Mahjoobi
50 and Adeli Mosabbebi, 2009). The error of the models is estimated through the average error rate:
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9

$$E = \frac{1}{K} \sum_{i=1}^K E_i \quad (16)$$

10 where E_i is the error for a single estimation. In this work, a 4-fold cross validation is used.
11 Hence, the whole data set (96 data points) is randomly partitioned into 4 subsets (folds), so that
12 each fold contains 24 data points ($96/4$). For each time, one fold is selected as the test set and
13 the other 3 folds are used as the training set. The results of the cross validation of both ANNs and
14 GP are given in Table 7. The results show that the uncertainty of the developed models is very
15 low and so these models can be implemented for various data sets.
16
17
18
19

20
21
22
23
24
25
26
27
28
29
30
31

Uncertainty assessment ensures that the final GP solution, Eq. (14), is not sensitive to the data selection and, thus, can be implemented for different data sets, ranging from small-scale to field-scale data. As indicated in Table 7, using various data sets as training and testing data, the performance of Eq. (14) in predicting the maximum scour depth does not change significantly. Furthermore, uncertainty assessment shows that the final model has not experienced overlearning problem.

32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

In order to achieve the models reliability, the box plot of the various models' discrepancy is very useful (Etemad-Shahidi and Ghaemi, 2011). The discrepancy is the distance between lower extreme and upper extreme in box plots and shows the models uncertainty. As indicated in Fig. 9, box plots are larger for the existing equations that mean they are more conservative and need higher values of safety factor. However, the discrepancy for the evolved models has small values and so they are more reliable. Furthermore, it was found that the empirical equations of Xie (1981), Sumer and Fredsøe (2000), and Lee and Mizutani (2008) are more conservative in comparison with the S3-GP and S3-ANN(5) evolved model. Also, Fig. 9 shows that the lower and upper quartiles of the data average in the empirical equations have significantly different numbers that indicates their uncertainty. This gap is not as large as empirical equations in the developed models (the S3-ANN(5) and S3-GP models). Having larger box height, the empirical formulas need larger safety factors to cover all the range of the predicted scour depths. However, the S3-GP and S3-ANN(5) models are more accurate and reliable in comparison with the empirical formulas.

7. Sensitivity Analysis

Sensitivity analysis is a conventional method for determining the relative significance of input parameters in the modeling process. Employing irrelevant or insignificant input parameters can lead to complex models, which are very difficult to evaluate and interpret. One of the major capabilities of genetic programming is its inherent power in the determination of the variables importance in the evolved model, in a way that the unimportant variables are gradually omitted in the final evolved model.

To determine the importance of the input parameters contributing to S3-GP model, the Liong et al. (2002) approach was implemented, where only one input parameter varies while the others are constant and a variation of $\pm 15\%$, $\pm 10\%$, $\pm 5\%$ for each input parameter was considered at each stage. The influence of the modification procedures on the proposed formula for prediction of S_{\max}/H_{nb} is measured in terms of the average percentage change (*APC*) as:

$$APC = \frac{100}{N} \sum_{i=1}^N \left\{ \frac{\left(\frac{S_{\max}}{H_{nb}} \right)_o - \left(\frac{S_{\max}}{H_{nb}} \right)_m}{\left(\frac{S_{\max}}{H_{nb}} \right)_o} \right\}_i \quad (17)$$

where $\left(\frac{S_{\max}}{H_{nb}} \right)_o$ is the predicted relative scour depth obtained by means of S3-GP using the original values of the input variables, and $\left(\frac{S_{\max}}{H_{nb}} \right)_m$ is the modified predicted relative scour depth due to variation of a particular variable and N is the number of data points. The procedure is repeated for all of the input variables. The significance of the input parameters resulting from the sensitivity analysis is summarized in Table 8. As seen, S_{\max}/H_{nb} is mostly affected by Cr and followed by h_{toe}/L_{nb} , respectively. This result is in line with the experimental findings of Sumer and Fredsøe (2000) and Lee and Mizutani (2008), which have pinpointed the reflection coefficient as the most important parameter for the scour depth at breakwaters.

The ANN weight matrix can be used to assess the relative importance of the various input parameters on the output parameter (Kasiri et al., 2008). The following equation was proposed based on partitioning of connection weights by Garson (1991):

$$I_j = \frac{\sum_{m=1}^{m=N_h} \left[\left(|W_{mn}^{ho}| \times \left(|W_{jm}^{ih}| / \sum_{l=1}^{N_i} |W_{kl}^{ih}| \right) \right) \right]}{\sum_{k=1}^{k=N_i} \left\{ \sum_{m=1}^{m=N_h} \left(|W_{km}^{ih}| / \sum_{k=1}^{N_i} |W_{kl}^{ih}| \right) \times |W_{mn}^{ho}| \right\}_i} \quad (18)$$

where I_j is the relative importance of the j th input variable on output variable, N_i and N_h are the number of input and hidden neurons, respectively. W_s are connection weights, the superscripts i, h and o refer to input, hidden, and output layers, respectively; and subscripts k, m and n refer to input, hidden, and output neurons, respectively. Table 9 shows the results of the ANN sensitivity analysis based on Eq. (18). As indicated, the wave reflection is the most impressive parameter on the scour depth and, again, the relative importance of the various input parameters calculated by Eq. (18) is in line with the S3-GP model results (Table 8). This shows that the S3-ANN(5) model is very sensitive to the wave reflection and the relative water depth at the toe, and the Shields parameter and the wave steepness are as important as the mentioned input parameters.

8. Summary and Conclusion

In this study, the non-breaking wave-induced scour depth at the trunk section of breakwaters has been studied by Genetic Programming (GP) and Artificial Neural Networks (ANNs) methodologies. Experimental data sets collected from the available literature have been used for developing the models. The developed models predict the relative scour depth (S_{max}/L_{nb}) as function of the reflection coefficient (Cr), the non-breaking wave steepness (H_{nb}/L_{nb}), the Shields parameter (θ), and the relative water depth at the toe of the breakwater (h_{toe}/L_{nb}).

In order to achieve the best possible accurate models, various ANNs and GP models have been developed using different combinations of input parameters. The results of the developed models indicate that the S3 model structure (S3) yields the best results and that all of the mentioned parameters play an important role in predicting the maximum scour depth. The results of the developed models have been compared with those of the equations by Xie (1981), Sumer and Fredsøe (2000), and Lee and Mizutani (2008) in terms of error measures. The results

1
2
3
4 indicate that the developed models (S3-ANN(5) and S3-GP) predict the S_{max}/L_{nb} better than those
5
6 of the existing empirical equations in terms of accuracy. Also, unlike the existing formulas,
7
8 which do not include the effects of many parameters of importance of the scouring, the
9
10 developed models include the effects of most of the parameters that have an important influence
11
12 on the maximum scour depth at breakwaters.

13
14 To verify that the developed models are consistent with the existing findings and reliable,
15
16 the scour depth predicted by S3-GP has been analysed as function of four main input parameters,
17
18 each varied while keeping the others fixed. It is shown that the GP best model (S3-GP) behavior
19
20 is in complete agreement with the known findings. In order to examine the reliability of the
21
22 developed models, a probabilistic approach has been applied using box plots. According to the
23
24 box plot, the discrepancy between observed scour depths and those predicted by means of the
25
26 ANNs and GP models is smaller than that achieved when predicting the scour through empirical
27
28 equations. A sensitivity analysis based on APC formulation has been conducted to explore the
29
30 effects of the various input parameters contribution in the ANNs and GP developed models. The
31
32 results of the sensitivity analysis proved that the reflection coefficient is the most important
33
34 parameter in the scour process at breakwaters. The latter result is clearly in line with the
35
36 experimental result of Sumer and Fredsøe (2000) and Lee and Mizutani (2008).

37 **Acknowledgments**

38
39 The authors are grateful to Professor William Allsop and Professor Marcel R.A. van Gent
40
41 for their constructive comments. Also, the authors express their sincere gratitude to Professors
42
43 M. Sumer, J. Fredsøe, K.H. Lee, N. Mizutani, S.L. Xie for providing their valuable experimental
44
45 data for this paper.
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

References

- Akib, S., Mohammadhassani, M., Jahangirzadeh, M.A., 2014. Application of ANFIS and LR in prediction of scour depth in bridges. *Computers & Fluids* 91, 77–86.
- Azamathullah, H.M., 2012. Gene expression programming for prediction of scour depth downstream of sills. *Journal of Hyrology* 460, 156-159.
- Azamathullah, H.M., Guven, A., Kagan Demir, Y., 2011. Linear genetic programming to scour below submerged pipeline. *Ocean Engineering* 38, 995-1000.
- Balouchi, B., Nikoo, M.R., Adamowski, J., 2015. Development of expert systems for the prediction of scour depth under live-bed conditions at river confluences: Application of different types of ANNs and the M5P model tree, *Applied Soft Computing* 34, 51-59.
- Baquerizo, A., Losada, M.A., 1998. Longitudinal current induced by oblique waves along coastal structures, *Coastal Engineering* 35, 211-230.
- Carter, T.G., Liu, L.F.P., Mei, C.C., 1973. Mass transport by waves and offshore sand bedforms. *ASCE J. Waterways, Harbors Coastal Engineering*, ASCE 99 (WW2), 165–184.
- De Best, A., Bijker, E.W., Wichers, J.E.W., 1971. Scouring of sand in front of a vertical breakwater, *Conference on Port and Ocean Engineering under Artic Conditions*, the Norwegian institute of technology, Trondheim, Norway, pp. 1077-1086.
- Etemad-Shahidi, A., Ghaemi, N., 2011. Model tree approach for prediction of pile groups scour due to waves. *Ocean Engineering* 38 (13), 1522-1527.
- Garson, G.D., 1991. *Interpreting Neural-Network Connection Weights*, AI Expert; Miller FRreeman, Inc.: San Francisco,p 46.
- Hajivalie, F., Yeganeh-Bakhtiary, A., Gotoh, H., 2008. A comparison between standing wave pattern in front of vertical breakwater with horizontal and slope bed. *The 8th international conference on coasts, ports and marine structures (ICOPMASS 2008)*, Tehran, Iran.
- Hughes, S.A., Fowler, J.E., 1991. Wave-induced scour prediction at vertical walls. *ASCE Proc. Conf. Coastal Sediments* 91, 1886–1899.
- Jain, P., Deo, M.C., 2007. Real-time wave forecasts off the western Indian coast. *Applied Ocean Research* 29,72-79.

- 1
2
3
4 Kasiri, M.B., Aleboyeh, H., Aleboyeh, A., 2008. Modeling and Optimization of Heterogenous
5 Photo-Fenton Process with Response Surface Methodology and Artificial Neural Networks.
6 Environmental Science and Technology 42, 7970-7975.
7
8
9
10 Kızılöz, B., Çvikb, E., Burak A., 2015. Apor Estimation of scour around submarine pipelines
11 with ArtificialNeural Network. Applied Ocean Research 51, 241–251.
12
13
14 Koza, J. 1992 Genetic programming, On the Programming of Computers by Natural Selection.
15 MIT Press, Cambridge, MA.
16
17 Lee, K.-H., Mizutani, N., 2008. Experimental study on scour occurring at a vertical impermeable
18 submerged breakwater. Applied Ocean Research 30 (2), 92-99.
19
20
21 Lillycrop, W.J., Hughes, S.A., 1993. Scour hole problems experienced by the corps of engineers,
22 data presentation and summary. Miscellaneous papers, CERC-93-2. U.S. Army Engineers
23 Waterways Experiment Station, Coastal Engineering Research Center, Vicksburg, MS.
24
25
26 Liong, S.Y., Gautam, T.R., Khu, S.T., Babaovic, V., Muttill, N., 2002. Genetic programming, A
27 new paradigm in rainfall-runoff modeling. Journal of American Water Resources
28 Association. 38 (3), 219-231.
29
30
31
32 Losada, M.A., Gimenez-Curto, L.A., 1981. Flow characteristics on rough, permeable slopes
33 under wave action. Coastal Eng. 4(3), 187–206.
34
35
36 Mahjoobi, J., Adeli Mosabbeb, E., 2009. Prediction of significant wave height using regressive
37 support vector machines. Ocean Engineering 36, 339-347.
38
39 **Mattioli, M., Alsina, J.M., Mancinelli, A., Miozzi, M., Brocchini, M., 2012. Experimental**
40 **investigation of the nearbed dynamics around a submarine pipeline laying on different types**
41 **of seabed: the interaction between turbulent structures and particles. Adv. Water Res. 48, 31-**
42 **46.**
43
44
45
46 **Mattioli, M., Mancinelli, A, Brocchini, M., 2013. Experimental investigation of the wave-**
47 **induced flow around a surface-touching cylinder. J. Fluids Struct. 37, 62-87.**
48
49
50 Oumeraci, H., 1994. Review and analysis of vertical breakwater failure-lessons learned. Coastal
51 Engineering 22, 3-29.
52
53
54 Poli, T., McPhee, N.F., 2008. Covariant parsimony pressure in genetic programming. Technical
55 Report, CES-480. Department of Computing and Electronic Systems, University of Essex.
56
57
58
59
60
61
62
63
64
65

- 1
2
3
4 Pourzangbar, A. 2012. Determination of the most effective parameters on scour depth at seawalls
5 using genetic programming (GP). The 10th international conference on coasts, ports and
6 marine structures (ICOPMASS 2012), Tehran, Iran.
7
8
9
10 Rogers, L.L., Dowla, F.U., 1994. Optimization of groundwater remediation using artificial neural
11 networks with parallel solute transport modeling. *Water resources research* 30 (2), 457-481.
12
13 Samadi, M., Jabbarib, E., Azamathulla, H.M., Mohammad Mojallald, M., 2015. Estimation of
14 scour depth below free overfall spillways using multivariate adaptive regression splines and
15 artificial neural networks. *Engineering Applications of Computational Fluid Mechanics* 9(1),
16 291-300.
17
18
19
20 Sawaragi, T., 1966. Scouring due to wave action at the toe of permeable coastal structure. In:
21 *Proc. 10th Conference on Coastal Engineering, Tokyo, Japan Vol. 2*, pp. 1036–1047.
22
23
24 Sumer, B.M., Fredsøe, J., 2000. Experimental study of 2D scour and its protection at a rubble-
25 mound breakwater. *Coastal Engineering* 40 (1), 59-87.
26
27
28 Van Ghent, M.R.A., Van den Boogard, H.F.P., Pozueta, B., Medina, J.R., 2007. Neural network
29 modelling of wave overtopping at coastal structures. *Coastal Engineering* 54, 586–593.
30
31
32 Whitehouse, R.J.S., 1998: *Scour at marine structures*. Thomas Telford, ISBN 07277-26552,
33 London.
34
35
36 Xie, S.L., 1981. Scouring patterns in front of vertical breakwaters and their influence on the
37 stability of the foundations of the breakwaters. Department of Civil Engineering, Delft
38 University of Technology, the Netherlands, 61 pp.
39
40
41 Yeganeh-Bakhtiary. A., Ghorbani. M.A., Pourzangbar. A., 2012. Determination of the Most
42 Important Parameters on Scour at Coastal Structures. *Journal of Civil Engineering and*
43 *Urbanism*. 2(2), 68-71.
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Figure captions

Fig. 1. Sketch illustrating the role of variables and the waves' characteristics that contribute to Eqs. (6) and (15).

Fig. 2. An example of GP final solution ($x^2 + (5 + \sqrt{y})$) in the form of Parse Tree.

Fig. 3. Typical genetic programming flowchart (Koza, 1992).

Fig. 4. (a) Crossover between ($x^2 + (5 + \sqrt{y})$) and ($y + 4$), (b) Mutation operation applied on a parent ($x^2 + (5 + \sqrt{y})$).

Fig. 5. The ANN structure used in this study.

Fig. 6. Comparison between the measured (dataset) and predicted (formula) maximum scour depth from the specific works of (a) Xie (1981), (b) Sumer and Fredsøe (2000) and (c) Lee and Mizutani (2008). Panel (d) gives the maximum scour depth predicted by the formulas of Xie (1981) and Sumer and Fredsøe (2000) on the dataset made of all the data coming from the mentioned works.

Fig. 7. Comparison between the measured and predicted maximum scour depth by the developed models (S3-ANN(5) model, and S3-GP model (Eq. 15)) for (a) Xie (1981) dataset, (b) Sumer and Fredsøe (2000) dataset, (c) Lee and Mizutani (2008) dataset, (d) the full dataset.

Fig. 8. Scour depth predicted by Eq. (15) as function of: (A) the relative water depth (h_{toe}/L_{nb}); (B) the reflection coefficient (Cr); (C) the Shields parameter (θ) and (D) the wave steepness (H_{nb}/L_{nb})

Fig. 9. Box plots of the developed models and the existing empirical equations in predicting of the scour depth for all datasets used in this paper.

Table captions

Table 1
The characteristics of employed Genetic Programming models.

Table 2
Ranges of the parameters employed to train and test the GP model.

Table 3
Performance indices of the developed models and the existing empirical equations in the prediction of various datasets.

Table 4
Model structures for the non-breaking scour depth prediction.

Table 5
Comparison of various structures of ANN in the prediction of training and testing data set.

Table 6
Performance indices of various GP models in the prediction of the relative scour depth.

Table 7
Error statistics of predicted maximum scour depth by S3-ANN(5) and S3-GP (4-fold cross validation).

Table 8
Average Percentage Change (APC) in the relative scour depth predicted by Eq. (15) due to changes in specific variables.

Table 9
The relative importance of input parameters in predicting scour depth based on S3-ANN(5).

Table 1

The characteristics of employed Genetic Programming models.

| | |
|---------------------------------------|--|
| Function set | $+$, $-$, \times , \div , Exp, x^2 , $x^{1/2}$, $x^{1/3}$, $\sinh(x)$ |
| Number of chromosomes | 30 |
| Head size | 8 |
| Number of genes | 3 |
| Linking function | Addition |
| Fitness function | RMSE |
| Mutation rate | 0.044 |
| One-point and two-point recombination | 0.3 |
| Gene transposition | 0.1 |
| Constants per gene | 2 |
| Range of constants | -10 to 10 |

Table 2

Ranges of the parameters employed to train and test the GP model.

| Parameters | Test data range | Full data range | Minimum | Average | Maximum |
|------------------|-----------------|-----------------|---------|---------|---------|
| Cr | 0.273-1.00 | 0.225-1.00 | 0.225 | 0.719 | 1.00 |
| H_{nb}/L_{nb} | 0.011-0.043 | 0.008-0.058 | 0.008 | 0.026 | 0.058 |
| h_{toe}/L_{nb} | 0.060-0.199 | 0.045-0.199 | 0.045 | 0.114 | 0.199 |
| θ | 0.020-0.187 | 0.023-0.187 | 0.020 | 0.081 | 0.187 |
| S_{max}/H_{nb} | 0.011-0.938 | 0.011-0.938 | 0.011 | 0.368 | 0.938 |

Table 3
Performance indices of the developed models and the existing empirical equations in the prediction of various datasets.

| Equation/Model | Dataset | <i>CC</i> | <i>RMSE</i> | <i>SI (%)</i> | <i>BIAS</i> |
|---|----------------------------------|-----------|-------------|---------------|-------------|
| Xie (1981) formula (Eq. 2) | Xie (1981) dataset | 0.806 | 0.162 | 31.71 | 0.015 |
| | All dataset | 0.528 | 0.286 | 77.90 | 0.154 |
| Sumer and Fredsøe (2000) formula (Eq. 3) | Sumer and Fredsøe (2000) dataset | 0.820 | 0.158 | 68.37 | 0.074 |
| | All data set | 0.763 | 0.163 | 44.24 | 0.050 |
| Lee and Mizutani (2008) formula (Eq. 4) | Lee and Mizutani (2008) dataset | 0.952 | 0.069 | 23.67 | -0.054 |
| S3-ANN(5) | Xie (1981) | 0.943 | 0.081 | 15.7 | -0.0120 |
| | Sumer and Fredsøe (2000) | 0.951 | 0.065 | 28.3 | +0.0120 |
| | Lee and Mizutani (2008) | 0.931 | 0.051 | 17.4 | +0.0050 |
| | All data set | 0.959 | 0.068 | 18.40 | -0.0003 |
| S3-GP (Eq. 15) | Xie (1981) | 0.946 | 0.076 | 14.75 | 0.0102 |
| | Sumer and Fredsøe (2000) | 0.960 | 0.062 | 26.99 | 0.017 |
| | Lee and Mizutani (2008) | 0.956 | 0.041 | 14.21 | -0.006 |
| | All data set | 0.964 | 0.062 | 16.82 | 0.006 |

Table 4
Model structures for the non-breaking scour depth prediction.

| Model structure | Input parameters combination | Output parameter |
|-----------------|---|------------------|
| S1 | $Cr, h_{10e}/L_{nb}, H_{nb}/L_{nb}$ | S_{max}/H_{nb} |
| S2 | $Cr, h_{10e}/L_{nb}, \theta$ | S_{max}/H_{nb} |
| S3 | $Cr, h_{10e}/L_{nb}, H_{nb}/L_{nb}, \theta$ | S_{max}/H_{nb} |

Table 5
Comparison of various structures of ANN in the prediction of training and testing data set.

| ANN Referred Model | Model structure | Number of hidden layer neurons | Training data set | | | | Testing data set | | | |
|--------------------|-----------------|--------------------------------|-------------------|--------------|--------------|----------------|------------------|--------------|-------------|---------------|
| | | | CC | RMSE | SI (%) | Bias | CC | RMSE | SI (%) | Bias |
| S1-ANN(3) | S1 | 3* | 0.957 | 0.066 | 17.8 | 0.0001 | 0.938 | 0.085 | 23.7 | -0.009 |
| S1-ANN(4) | | 4 | 0.951 | 0.084 | 19.8 | 0.0002 | 0.928 | 0.081 | 25.7 | -0.011 |
| S1-ANN(5) | | 5 | 0.954 | 0.059 | 16.27 | 0.0001 | 0.905 | 0.113 | 31.3 | -0.025 |
| S1-ANN(6) | | 6 | 0.922 | 0.091 | 24.07 | 0.0001 | 0.708 | 0.221 | 61.5 | 0.021 |
| S2-ANN(3) | S2 | 3 | 0.915 | 0.091 | 24.75 | 0.001 | 0.933 | 0.087 | 24.3 | 0.008 |
| S2-ANN(4) | | 4 | 0.953 | 0.069 | 18.65 | 0.0012 | 0.929 | 0.089 | 24.8 | 0.0005 |
| S2-ANN(5) | | 5* | 0.960 | 0.063 | 16.94 | -0.0002 | 0.933 | 0.092 | 25.6 | -0.003 |
| S2-ANN(6) | | 6 | 0.940 | 0.082 | 21.68 | 0.0004 | 0.804 | 0.150 | 41.8 | -0.015 |
| S3-ANN(3) | S3 | 3 | 0.939 | 0.077 | 20.76 | 0.0006 | 0.910 | 0.104 | 29 | 0.016 |
| S3-ANN(5) | | 5* | 0.971 | 0.054 | 14.19 | -0.0016 | 0.929 | 0.104 | 29 | 0.003 |
| S3-ANN(6) | | 6 | 0.966 | 0.058 | 15.71 | 0.0011 | 0.925 | 0.103 | 28.7 | -0.001 |
| S3-ANN(7) | | 7 | 0.971 | 0.053 | 14.39 | 0.0004 | 0.925 | 0.105 | 29.2 | 0.013 |
| S3-ANN(8) | | 8 | 0.965 | 0.060 | 16.02 | 0.0002 | 0.894 | 0.119 | 33.2 | 0.017 |

*The results in bold show the best number of hidden layers for each specified model.

Table 6
Performance indices of various GP models in the prediction of the relative scour depth.

| Model (Equation) | Used dataset | <i>CC</i> | <i>RMSE</i> | <i>SI(%)</i> | <i>BIAS</i> |
|-------------------------|-------------------|--------------|--------------|--------------|--------------|
| S1-GP | Training data set | 0.900 | 0.099 | 26.65 | 0.003 |
| | Testing data set | 0.894 | 0.108 | 30.01 | -0.009 |
| S2-GP | Training data set | 0.915 | 0.092 | 25.51 | -0.001 |
| | Testing data set | 0.912 | 0.098 | 25.53 | -0.008 |
| S3-GP | Training data set | 0.981 | 0.050 | 13.23 | 0.005 |
| | Testing data set | 0.922 | 0.093 | 25.82 | 0.007 |

Table 7

Error statistics of predicted maximum scour depth by S3-ANN(5) and S3-GP (4-fold cross validation).

| Model | Used dataset | <i>CC</i> | <i>RMSE</i> | <i>SI(%)</i> | <i>BIAS</i> | <i>Mean ± std</i> |
|--|--------------|-----------|-------------|--------------|-------------|-------------------|
| S3-ANN(5) (4-fold cross validation) | All data set | 0.919 | 0.089 | 23.5 | -0.003 | 0.337 ± 0.252 |
| S3-ANN(5) (non- cross validation) | All data set | 0.959 | 0.068 | 18.4 | -0.0003 | 0.368 ± 0.241 |
| S3-GP (4-fold cross validation) | All data set | 0.934 | 0.078 | 19.8 | 0.011 | 0.348 ± 0.247 |
| S3-GP (non- cross validation) | All data set | 0.964 | 0.062 | 16.8 | 0.006 | 0.372 ± 0.223 |

Table 8
Average Percentage Change (APC) in the relative scour depth predicted by Eq. (15) due to changes in specific variables.

| Considered variable | Average Percentage Change (APC) | | | | | | Significance order |
|---------------------|---------------------------------|--------|--------|--------|--------|--------|--------------------|
| | -15% | -10% | -5% | +5% | +10% | +15% | |
| Cr | 20.68 | 11.76 | 3.58 | -14.39 | -25.43 | -58.75 | 1 |
| h_{ro}/L_{nb} | -38.53 | -26.43 | -14.79 | 4.07 | 12.81 | 21.66 | 2 |
| θ | -3.11 | -3.74 | -4.38 | -5.66 | -6.31 | -6.97 | 3 |
| H_{nb}/L_{nb} | -2.70 | -3.47 | -4.24 | -5.79 | -6.56 | -7.34 | 4 |

Table 9
The relative importance of input parameters in predicting scour depth based on S3-ANN(5).

| Variable under analysis | Relative importance (%) | Significance order |
|-------------------------|-------------------------|--------------------|
| Cr | 54.40 | 1 |
| h_{toe}/L_{nb} | 24.16 | 2 |
| θ | 9.14 | 3 |
| H_{nb}/L_{nb} | 3.30 | 4 |

Figure. 1

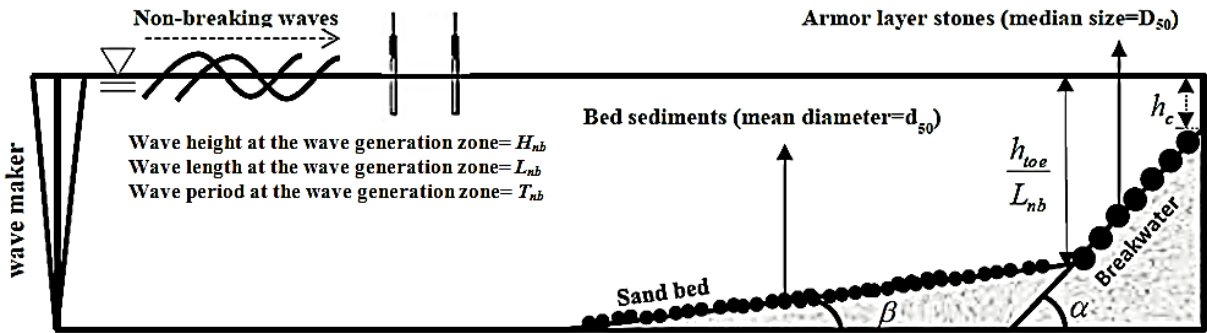


Fig. 1. Sketch illustrating the role of variables and the waves' characteristics that contribute to Eqs. (6) and (15).

Figure. 2

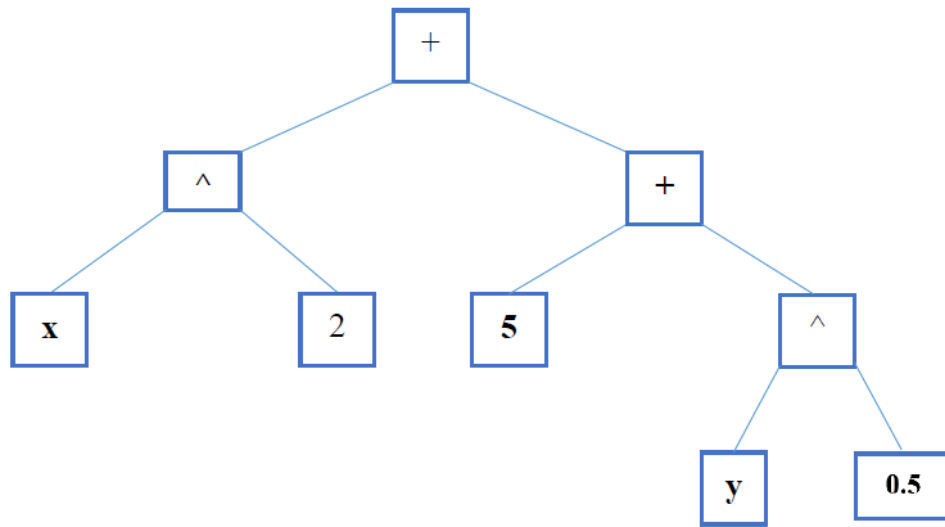


Fig. 2. An example of GP final solution $(x^2 + (5 + \sqrt{y}))$ in the form of Parse Tree

Figure. 3

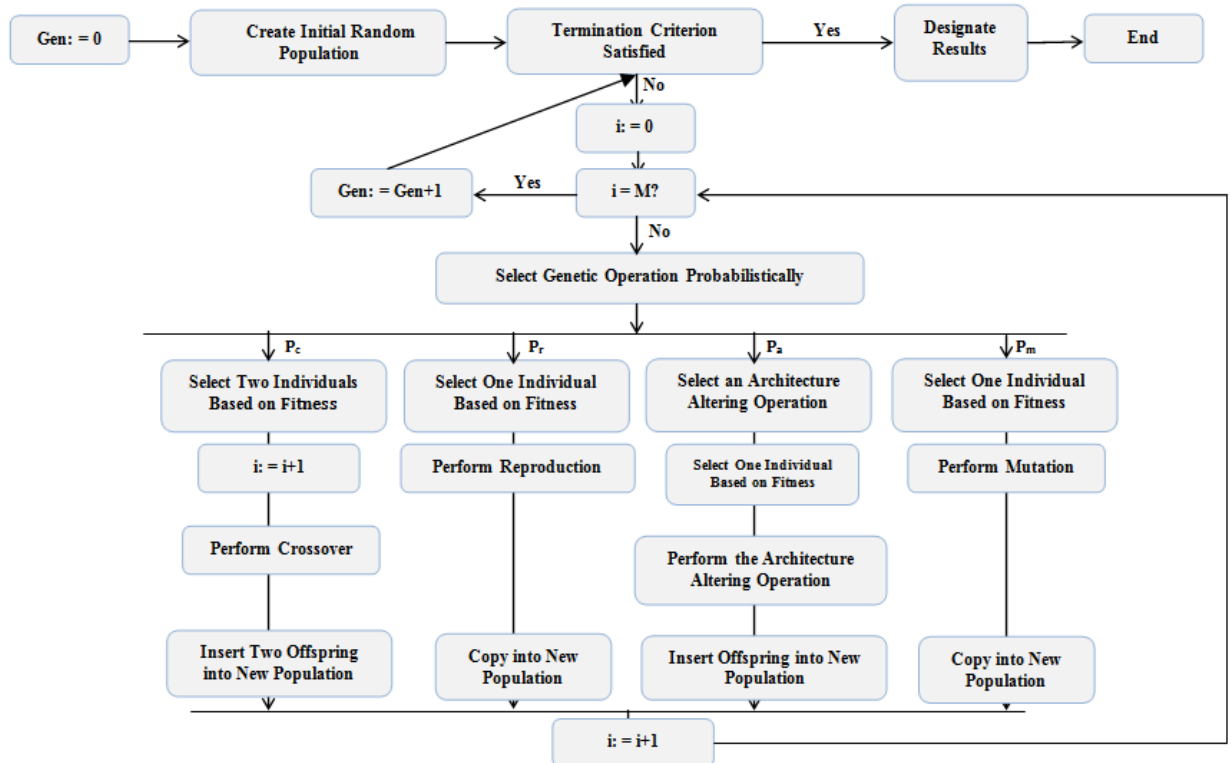


Fig. 3. Typical genetic programming flowchart (Koza, 1992).

Figure. 4

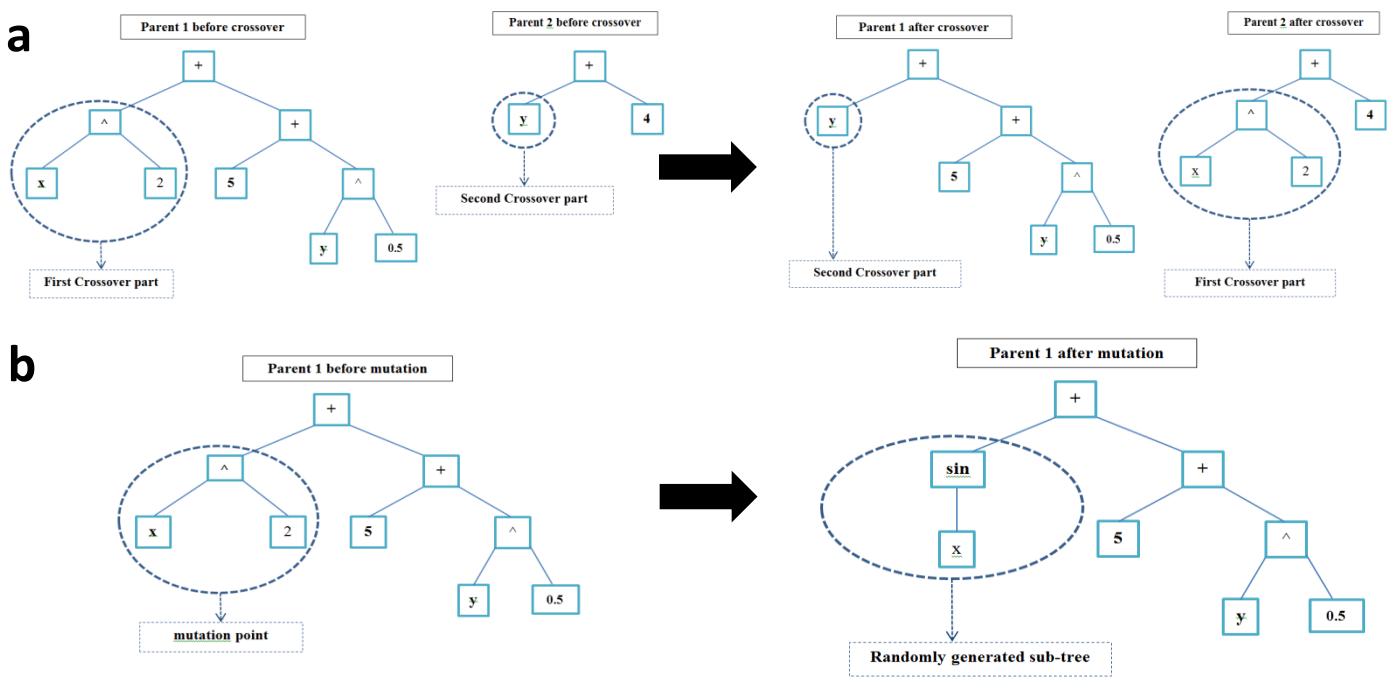


Fig. 4. (a) Crossover between $(x^2 + (5 + \sqrt{y}))$ and $(y + 4)$, (b) Mutation operation applied on a parent $(x^2 + (5 + \sqrt{y}))$

Figure. 5

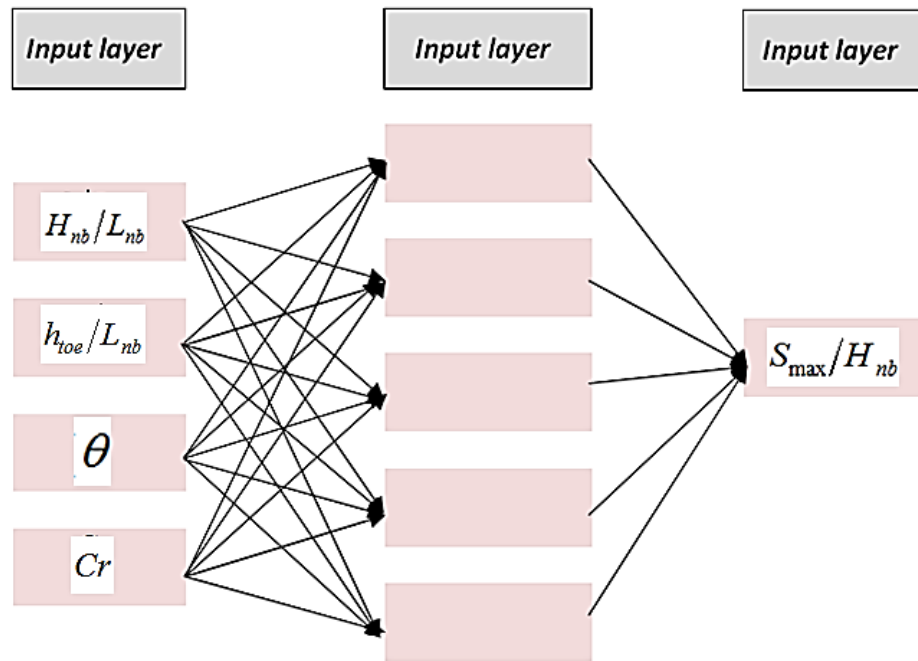


Fig. 5. The ANN structure used in this study

Figure. 6

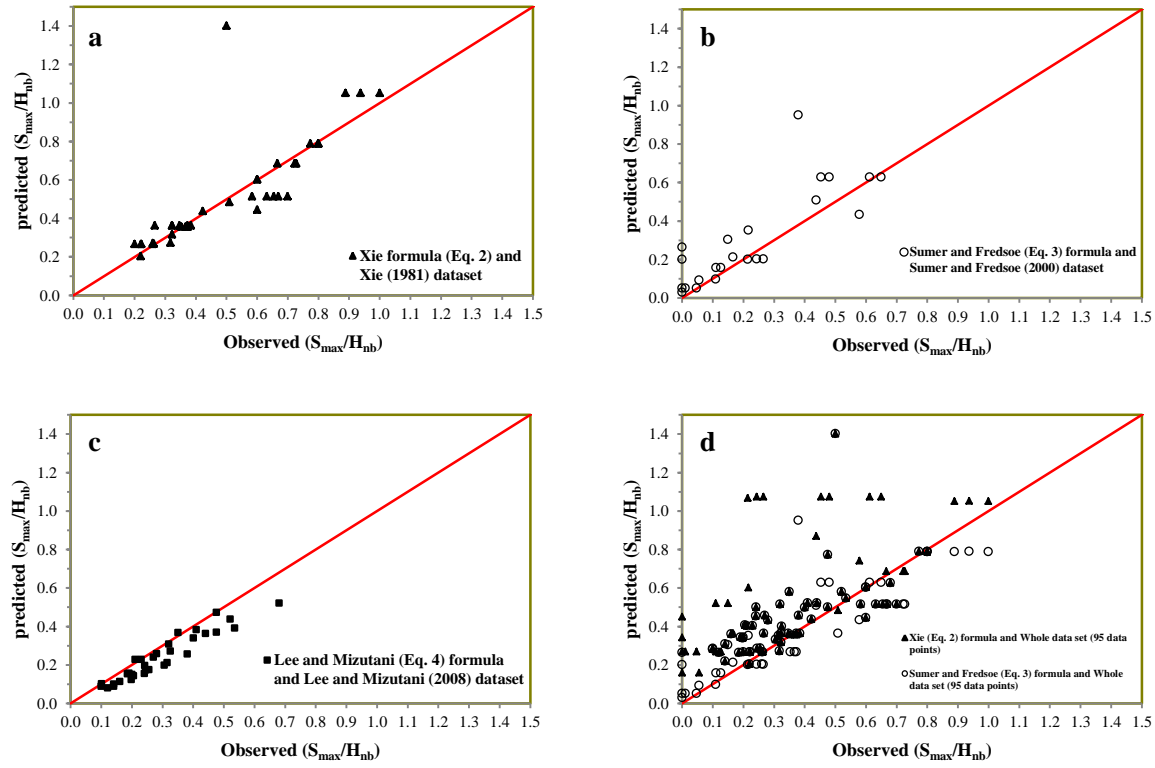


Fig. 6. Comparison between the measured (dataset) and predicted (formula) maximum scour depth from the specific works of (a) Xie (1981), (b) Sumer and Fredsøe (2000) and (c) Lee and Mizutani (2008). Panel (d) gives the maximum scour depth predicted by the formulas of Xie (1981) and Sumer and Fredsøe (2000) on the dataset made of all the data coming from the mentioned works.

Figure. 7

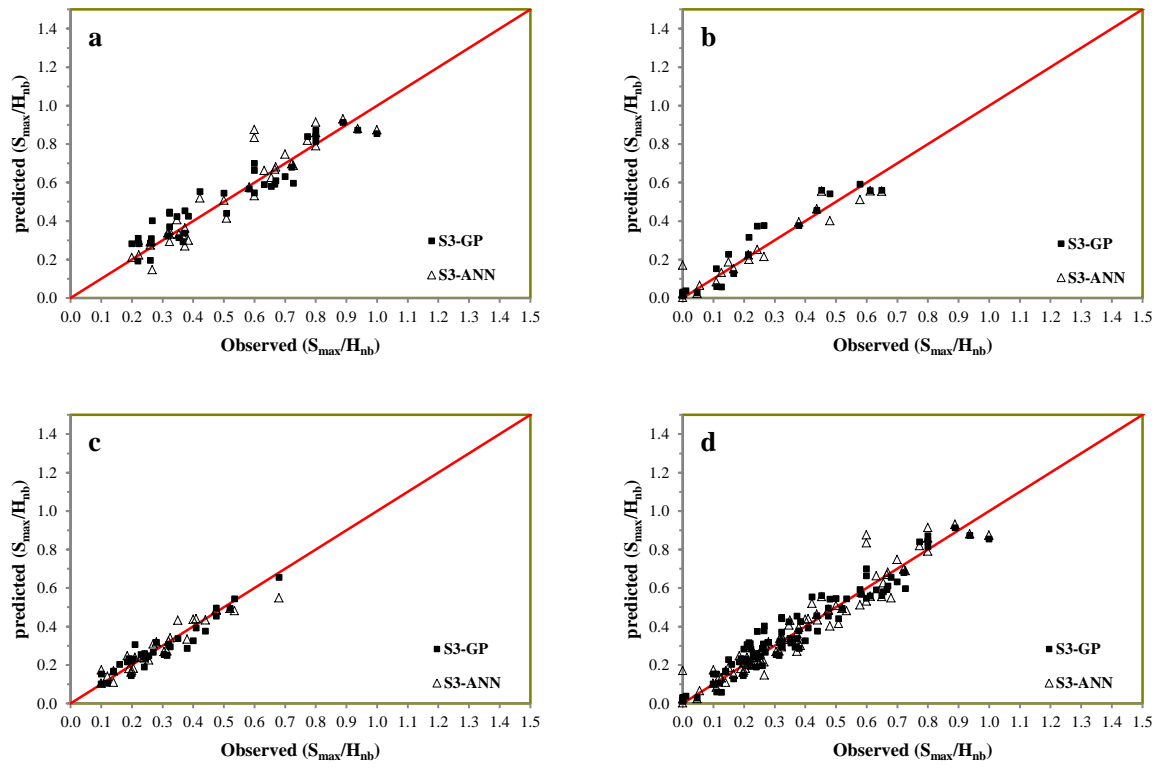


Fig. 7. Comparison between the measured and predicted maximum scour depth by the developed models (S3-ANN(5) model, and S3-GP model (Eq. 15)) for (a) Xie (1981) dataset, (b) Sumer and Fredsøe (2000) dataset, (c) Lee and Mizutani (2008) dataset, (d) the full dataset.

Figure. 8

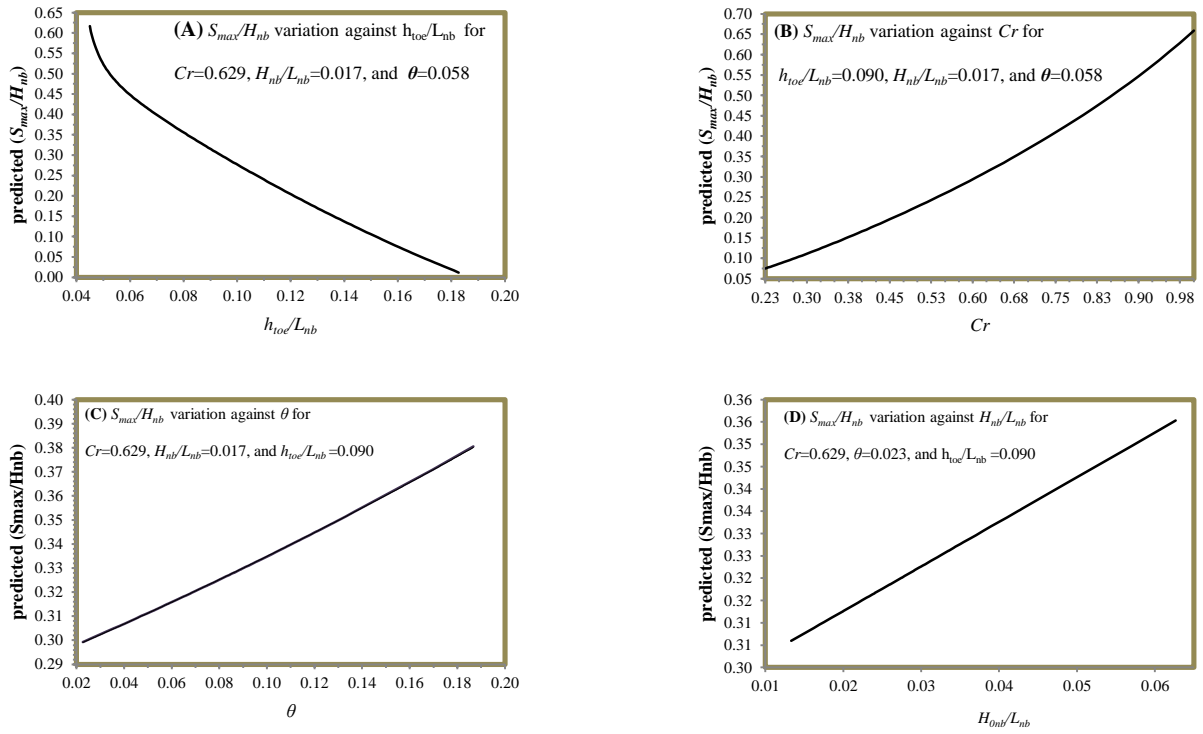


Fig. 8. Scour depth predicted by Eq. (15) as function of: (A) the relative water depth (h_{toe}/L_{nb}); (B) the reflection coefficient (Cr); (C) the Shields parameter (θ) and (D) the wave steepness (H_{nb}/L_{nb})

Figure. 9

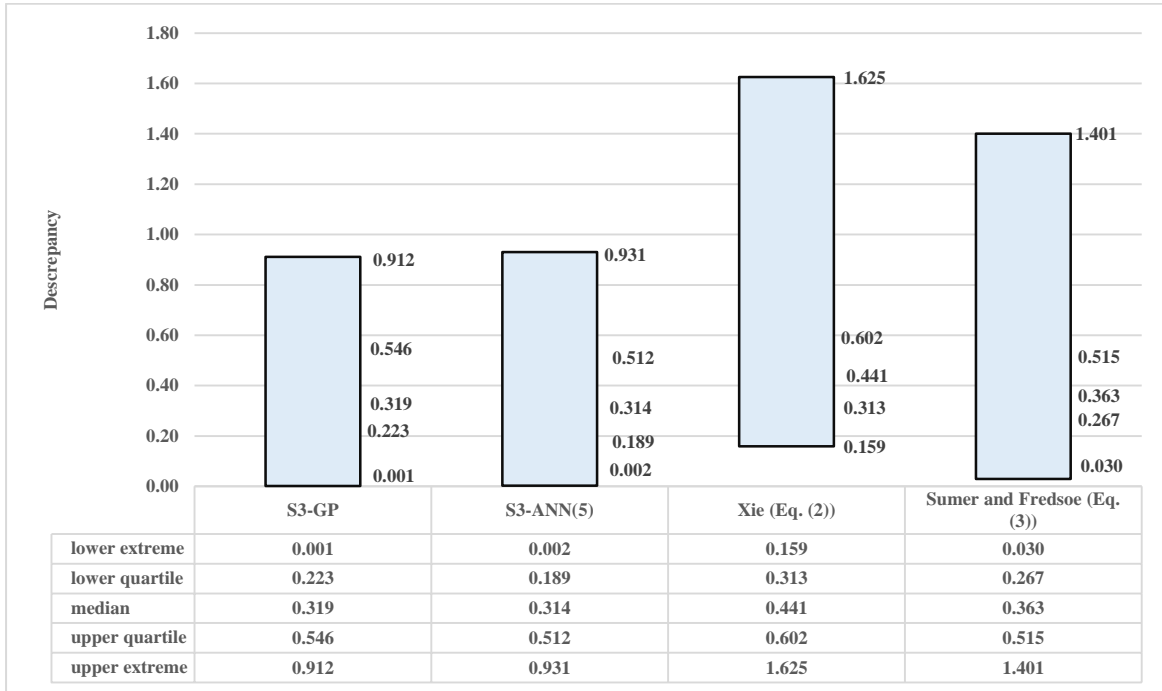


Fig. 9. Box plots of the developed models and the existing empirical equations in predicting of the scour depth for all datasets used in this paper.