

Article

Estimation of Scour Propagation Rates around Pipelines While Considering Simultaneous Effects of Waves and Currents Conditions

Mohammad Najafzadeh ^{1,*}, Giuseppe Oliveto ²  and Farshad Saberi-Movahed ³

¹ Department of Water Engineering, Faculty of Civil and Surveying Engineering, Graduate University of Advanced Technology, Kerman P.O. Box 76315-116, Iran

² School of Engineering, University of Basilicata, 85100 Potenza, PZ, Italy; giuseppe.oliveto@unibas.it

³ College of Engineering, North Carolina State University, Raleigh, NC 22606, USA; fsaberi@ncsu.edu

* Correspondence: m.najafzadeh@kgut.ac.ir

Abstract: Seabed offshore pipelines are widely applied to carry fluid over long distances of the seafloor. The design of offshore pipelines is conducted to bear quite a few environmental loading circumstances in order to provide a well-guarded and reliable fluid transition. Fluid leakage and pipeline vibration due to a failure of the pipeline are the prime causes of accidental catastrophes. Scour phenomena occur around offshore pipelines due to currents and/or wave conditions, consequently causing the susceptibility to pipeline failure. Then, scouring propagation rates require to be studied in three dimensions, namely beneath and normal to the offshore pipeline and the longitudinal direction of itself. In this research, Artificial Intelligent (AI) models are used to derive new regression equations based on the laboratory data for the estimation of 3D scour propagation patterns while seafloor offshore pipelines are exposed to simultaneous impacts of currents and waves. In this way, chiefly based on the experimental investigations conducted by Cheng and colleagues, seven sets of dimensional parameters were given in terms of the Shields' parameter due to currents and waves, the Keulegan–Carpenter number, the ratio of embedment depth to pipeline diameter, the ratio of orbital velocity to current velocity, and the wave/current angle of attack. Dimensionless parameters were used to provide regression-based equations to evaluate scour propagation rates in three dimensions. The performance of AI models was evaluated by various statistical measures. The model based on our proposed equations performed better than the reported models in the literature. Even more importantly, we indicated that our model inherently has a reliable physical consistency for variations of dimensionless parameters against the scour propagation patterns.

Keywords: seabed pipeline; combined waves and currents; scour propagation; Artificial Intelligence models; regression analysis



Citation: Najafzadeh, M.; Oliveto, G.; Saberi-Movahed, F. Estimation of Scour Propagation Rates around Pipelines While Considering Simultaneous Effects of Waves and Currents Conditions. *Water* **2022**, *14*, 1589. <https://doi.org/10.3390/w14101589>

Academic Editors: Kun Shi and Kaishan Song

Received: 28 April 2022

Accepted: 13 May 2022

Published: 16 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Marine pipelines, as one of the most well-known lifeline structures, are used for various fundamental purposes: removal of industrial and urban wastewater into the sea for cooling water in nuclear power plants, and conveyance of oil, gas, and water along the seabed. It has long been recognized that the scouring process occurs whenever the hydrodynamic forces of moving water exceed the resistive forces holding sediment particles at rest. Scouring beneath undersea pipelines has become the cornerstone of many investigations during the past three decades. Indeed, the scouring process causes damage to the fluid-carrying pipelines in terms of stability, vibration, and leaks, thus, experts have paid great attention to this issue [1,2]. Generally, an initial phase of scouring development is generated at many places along the pipeline due to either the naturally established piping or irregular level of seabed after the installation operation of the pipeline [2]. As a scour hole is initially formed in the near vicinity of the pipeline, free spans are generated along

they presented regression-based equations for the estimation of the longitudinal rate of scour propagation for both flow circumstances.

According to the above-mentioned investigation, knowledge of scour rate prediction in three dimensions needs to be developed for deriving regression equations. Previously derived equations were validated for longitudinal scour rates and have no capability to predict the scour rate in other directions [1,2]. Previous investigations generally demonstrated that traditional equations come from various restrictions such as simultaneous variations in the range of laboratory variables for longitudinal scour propagation rates around pipelines in the various flow conditions: currents, waves, and combined waves/currents. In addition to this, regression-based equations were not provided for other directions of scour rate propagation through experimental investigations [1,2,11,12]. In fact, the existence of empirical equations for all directions of scour propagation around pipelines can help deeply understand the mechanism of 3D scour rates around pipelines. Therefore, there is an essential need to obtain regression-based equations for analyzing three-dimensional scour rates.

With the emergence of Artificial Intelligence (AI) models, there is a wide range of comprehensive solutions to the perception of sophisticated real-world problems. Hence, the application of AI models can be used for reducing the complexity of complicated systems in different fields of science. In recent years, a few investigations have been performed to predict 3D scour rates by various AI models such as the Group Method of Data Handling (GMDH) [13], Artificial Neural Networks (ANNs) [14], Gene-Expression Programming (GEP), Multivariate Adaptive Regression Splines (MARS), and Evolutionary Polynomial Regression (EPR) [15]. These AI techniques demonstrated exceptional performance in the estimation of 3D scour rates around pipelines in wave conditions. In this case, the performance of experimental and field monitoring of scour rate propagation in three directions is expensive and requires various types of high/low-frequency cutting-edge amenities such as acoustic wireless sensors, multi-hop communication networks, and other Autonomous Underwater Vehicles (AUVs).

A review of reliable literature demonstrated that the major capability of these AI models originates in non-linear regression equations, providing an accurate estimation of natural patterns of scouring propagation rates around underwater pipelines exposed to waves only [13–16]. In the case of combined waves/currents, AI models have not been employed yet for the prediction of scouring propagation rates. Some of the soft computing models such as the MARS, GEP, EPR, and MT have intrinsic advantages of presenting non-linear relationships for a limited number of features, detecting the number of effective factors, and making the behavior of the produced mathematical expressions understandable. To the best of the authors' knowledge, the present research follows these outlines, introducing contemporary explicit relationships to estimate 3D scour rates at underwater pipelines relying on crucial non-dimensional factors in the combined waves/currents conditions. These regression-based-equations given by AI models would be more consistently precise in performance than traditional techniques, besides controlling physical patterns of contributory factors on the 3D scour propagation rates.

2. Overview of Databases

2.1. Dimensional Analysis

Over the last decade, previous investigations have shown that the propagation scour rates depend on the pipeline geometry, flow conditions (i.e., currents, waves, and combined waves/currents), and physical properties of bed materials. Even though there are several investigations on the propagation scour rate around pipelines exposed to the waves and currents, most experimental studies are dedicated to predicting the scour depth below the pipeline at equilibrium status, e.g., [1,2,7,8,11,12,17]. In the case of combined waves/currents, Cheng et al. [1] presented the results of an experimental investigation on the three-dimensional local scour below a rigid pipeline subjected to the combined wave and current. Due to the simultaneous presence of waves and currents, the primary function

which can establish a relationship between the three-dimensional scour propagation and effective variables is expressed as [1];

$$F_1(V_H, V_L, V_R, U_*, e, D, U_W, U_C, d_{50}, T, \tan \Phi, \alpha, \rho, \rho_s, \mu, g) = 0 \tag{1}$$

where $V_H, V_L,$ and V_R are, respectively, the non-dimensional scour propagation rate along the pipeline, at left-hand shoulder of the pipeline, and at right-hand shoulder of the pipeline, U_* is the shear velocity of the seabed, e is the embedment depth, D is the pipeline diameter, U_W is the maximum undisturbed orbital rate at the sea bottom just above the boundary of the waves, U_C is the flow velocity due to current, d_{50} is the median sediment size, T is the wave period, Φ is the repose angle for bed sediments, α is the flow incident angle to the pipeline, ρ is the mass density of water, ρ_s is the mass density of bed material, μ is the dynamic viscosity of water, and g is the acceleration due to the gravity.

To increase the accuracy level of explicit regression equations, all investigations performed in this field demonstrated that once Artificial Intelligence approaches and traditional methods were fed by grouped non-dimensional parameters, introduced as input-output vectors, they could produce better predictions of scouring rates than those obtained by dimensional (or raw) variables [1,2,14,15]. Hence, this study uses π -theorem to provide a set of dimensionless parameters to perform AI models.

In this way, $\rho, D,$ and U_W are selected as repeating variables, and therefore dimensionless parameters ($\pi_1, \pi_2, \pi_3, \dots, \pi_{12}, \pi_{13}, \pi_{14}$) were generated by the Buckingham theorem as follows:

$$F_2\left(\frac{V_H}{U_W}, \frac{V_R}{U_W}, \frac{V_L}{U_W}, \frac{e}{D}, \frac{U_W T}{D}, \frac{d_{50}}{D}, \phi, \alpha, \frac{\rho_s}{\rho}, \frac{\rho U_W D}{\mu}, \frac{\rho U_C D}{\mu}, \frac{U_C}{U_W}, \frac{g D}{U_W^2}, \frac{g D}{U_C^2}\right) = 0 \tag{2}$$

where $\pi_5, \pi_{10},$ and π_{11} are the Keulegan–Carpenter (KC) number and the Reynolds number of the pipeline due to regular wave (Re_W) and current (Re_C), respectively. On the basis of Cheng et al.’s [1] research, many π -variables in Equation (2) need to be combined and categorized as follows:

$$\begin{aligned} \pi'_7 &= \tan \phi; \pi'_8 = \sin \alpha; \pi'_9 = \frac{\rho_s}{\rho} - 1; \pi'_{12} = \frac{U_C}{U_C + U_W} \\ \pi'_{13} &= \frac{1}{\pi_{13} \pi'_9 \pi_6} = \frac{U_W^2}{g[\rho_s/\rho - 1]d_{50}} \\ \pi'_{14} &= \frac{1}{\pi_{14} \pi'_9 \pi_6} = \frac{U_C^2}{g[\rho_s/\rho - 1]d_{50}} \\ \pi'_1 &= \pi'_7 \sqrt{\pi_{13} \left(\frac{\pi_1}{\pi_6}\right)^2} = \frac{V_H \cdot D \cdot \tan \phi}{\sqrt{g[\frac{\rho_s}{\rho} - 1]d_{50}^3}} \\ \pi'_2 &= \pi'_7 \sqrt{\pi_{13} \left(\frac{\pi_2}{\pi_6}\right)^2} = \frac{V_R \cdot D \cdot \tan \phi}{\sqrt{g[\frac{\rho_s}{\rho} - 1]d_{50}^3}} \\ \pi'_3 &= \pi'_7 \sqrt{\pi_{13} \left(\frac{\pi_3}{\pi_6}\right)^2} = \frac{V_L \cdot D \cdot \tan \phi}{\sqrt{g[\frac{\rho_s}{\rho} - 1]d_{50}^3}} \end{aligned}$$

where $\pi'_1, \pi'_2,$ and π'_3 denote $V_H^*, V_R^*,$ and V_L^* , respectively. Additionally, π'_{11} is the Shields’ parameter for both wave and current conditions, which have a key role to play in the scour propagation below pipelines. Therefore, the functional relationship (2) is reduced to

$$F_2(V_H^*, V_L^*, V_R^*, e/D, KC, \theta_C, \theta_W, Re_W, Re_C, m, \sin \alpha) = 0, \tag{3}$$

where $V_H^*, V_L^*,$ and V_R^* are, respectively, the non-dimensional scour propagation rate along the pipeline, at the left-hand shoulder of the pipeline, and at the right-hand shoulder

of the pipeline, KC is the Keulegan–Carpenter number, θ_C is the Shields' parameter due to steady current, θ_W is the wave Shields' parameter, Re_W is the wave-induced Reynolds number, and Re_C is the current-induced Reynolds number, which is calculated by the following function respectively; also, m is the velocity ratio $[U_C/(U_C + U_W)]$.

$$V_H^* = \frac{V_H}{\left[\sqrt{g \left(\left(\frac{\rho_s}{\rho} \right) - 1 \right) d_{50}^3 / D \tan \Phi} \right]} \quad (4)$$

$$V_L^* = \frac{V_L}{\left[\sqrt{g \left(\left(\frac{\rho_s}{\rho} \right) - 1 \right) d_{50}^3 / D \tan \Phi} \right]} \quad (5)$$

$$V_R^* = \frac{V_R}{\left[\sqrt{g \left(\left(\frac{\rho_s}{\rho} \right) - 1 \right) d_{50}^3 / D \tan \Phi} \right]} \quad (6)$$

$$KC = \frac{U_W \cdot T}{D} \quad (7)$$

$$\theta_C = \frac{u_*^2}{g(S-1)d_{50}} \quad (8)$$

$$\theta_W = \frac{\tau_w}{\rho g(S-1)d_{50}} \quad (9)$$

in which S is the specific gravity of sediment grain $[\rho_s/\rho]$ and τ_w is the wave-induced shear stress on the seabed, which is expressed as

$$U_W^{-2} \tau_w = \frac{1}{2} \rho f_w, \quad (10)$$

where f_w is the wave-induced shear stress on the seabed and was proposed by Soulsby [18]. In particular,

$$f_w = 0.237r^{-0.52} \quad (11)$$

$$r = U_W T / 4\pi d_{50} \quad (12)$$

$$Re_W = \frac{\rho U_W D}{\mu} \quad (13)$$

$$Re_C = \frac{\rho U_C D}{\mu} \quad (14)$$

2.2. Experimental Case

In the case of experimental datasets conditions, Cheng et al. [1] performed three-dimensional scour experiments in a wave flume with 50 m in length, 4 m in width, and 2.5 m in depth. A concrete sandpit of 4 m long, 4 m wide, and 0.25 m deep was built in the test section. The transitions from the original flume bed to the test section and from the test section back to the original flume bed were achieved through two 1:20 slopes on either end of the sandpit. The upstream end of the sandpit was 13.5 m from the flow inlet, and the downstream end was 6.5 m from the flow outlet. A clear pipeline with a smooth surface, diameter of 50 mm, and wall thickness of 8 mm was utilized in their study. All tests were conducted under combined wave and current conditions, with the KC number ranging from 2.0 to 12.0, velocity ratio m ranging from 0.181 to 0.813, and initial pipeline embedment depth varying from $0.1D$ to $0.5D$. Also, median sediment size and relative density of sediments grain $[S = \rho_s/\rho]$ were 0.37 mm and 2.7, respectively. The tests were conducted under live-bed conditions. Additionally, the non-dimensional parameters are taken into account as constant parameters such as: $\Phi = 32^\circ$, $\alpha = 0^\circ$, $\rho = 1000 \text{ Kg/m}^3$, $\rho_s = 2650 \text{ Kg/m}^3$, and $\mu = 0.001 \text{ Pa.s}$. Cheng et al. [1] reported that the maximum Shields' parameter for the

motion of the bed sediments was evaluated based on the Soulsby [18] approach. The only formula suggested by Cheng et al. [1] was employed to estimate the three-dimensional equilibrium scour propagation rate induced by combined waves/currents is expressed as

$$\frac{V_H}{\sqrt{g\left(\frac{\rho_s}{\rho} - 1\right)d_{50}^3/D \tan \Phi}} = K_{WC} \left(1 - \frac{e}{D}(1 + \sin \alpha)\right) ((1 - m)\theta_w + m\theta_c)^{5/3} F \tag{15}$$

where K_{WC} is a constant and F , that was proposed by Sumer and Fredsøe (2002), is given by the following empirical equations:

$$F = \begin{cases} \frac{5}{3}(KC)^a \exp(2.3b) & 0 < m \leq 0.7 \\ 1 & 0.7 < m < 1 \end{cases} \tag{16}$$

$$a = \begin{cases} 0.557 - 0.912(m - 0.25)^2 & 0 \leq m \leq 0.4 \\ -2.14m + 1.46 & 0.4 < m \leq 0.7 \end{cases} \tag{17}$$

$$b = \begin{cases} -1.14 + 2.24(m - 0.25)^2 & 0 \leq m \leq 0.4 \\ 3.3m - 2.5 & 0.4 < m \leq 0.7 \end{cases} \tag{18}$$

3. Strategy on Selection of Effective Parameters

Effects of non-dimensional parameters on scouring propagation rate, which have been recognized by dimensional analysis technique, are considered to develop the AI models for predicting the three-dimensional scour rate. To reach a reasonable consistency between the results of AI models and experimental datasets, optimal selection of contributory factors draws great attention. Generally, applying the number of typical input parameters and mathematical structure associated with AI model-based relationships are well-matched and consistent with those obtained by empirical equations. Dimensionless parameters, by which Equations (15)–(18) were developed, can be used to estimate propagation scour rates (V_H^* , V_R^* , V_L^*). Hence, six non-dimensional parameters (KC , θ_C , θ_W , e/D , $\sin \alpha$, and m) were initially considered as input parameters to develop the AI models. As an alternative, according to Equation (15) given by Cheng et al. [1], regression terms of F and parameters of a and b are merged into the list of input variables. In addition to this, two typical parameters of Reynolds number (Re_W and Re_C) were excluded from the list of effective parameters because values of Reynolds number stood at the turbulent states, demonstrating insignificant influence on the scour rate propagation during the Cheng et al. [1] experiments. Hence, Equation (3) is re-written as,

$$F_3(V_H^*, V_L^*, V_R^*e/D, KC, \theta_C, \theta_W, m) = 0 \tag{19}$$

In Table 1, various combinations of input variables are presented.

Table 1. Optimal selection of input variables.

Set No.	List of Input Variables
Set 1	$e/D, KC, m, \theta_W, \theta_C$
Set 2	$1 - e/D, KC, (1 - m)\theta_W + m\theta_C$
Set 3	$1 - e/D, F, \theta_W, \theta_C$
Set 4	$1 - e/D, F, (1 - m)\theta_W + m\theta_C$
Set 5	$e/D, F, m, \theta_W, \theta_C$
Set 6	$1 - e/D, KC, (1 - m)\theta_W, m\theta_C$
Set 7	$1 - e/D, KC, [(1 - m)\theta_W + m\theta_C]^{5/3}$

In the case of combined waves/currents conditions, Cheng et al. [1] performed 65 experiments in which scour propagation was not observed in the 9 experiments. Additionally, 14 experiments were given in the state of onset of the scour multiple places along the pipeline for small values of embedment depths (for example, $e/D = 0.1$). Generally, 23 experimental observations were excluded from feeding AI models due to the fact that rates of the scour propagation need to be measured in three dimensions as what was considered in the previously-related works for the wave-only conditions [13–15]. The datasets used to predict the scour propagation rate was collected from Cheng et al. [1] (42 datasets). Table 2 presents the ranges of datasets for modeling the scour rates. Out of the datasets, about 75% (31 datasets) and 25% (11 datasets) were selected randomly to perform training and testing stages for AI models, respectively.

Table 2. Statistical descriptions of variables.

Parameters	Data Range	Average	Standard Deviation	Coefficient of Variation
T (s)	1–2	1.536	0.419	0.273
U_W (m/s)	0.11–0.34	0.261	0.074	0.284
E (cm)	0.5–2.5	1.57	0.586	0.3717
u^* (m/s)	0.0036–0.0282	0.0171	0.00769	0.449
U_C (m/s)	0.06–0.47	0.291	0.130	0.445
V_H (mm/s)	0.33–4.55	1.669	0.893	0.535
m	0.15–0.79	0.505	0.187	0.371
θ_W	0.0278–0.1315	0.08057	0.03094	0.361
θ_C	0.00048–0.0297	0.01356	0.00908	0.669
KC	2.2–13.6	8.314	3.723	0.448
e/D	0.2–0.5	0.319	0.102	0.319
V_H^*	0.26–3.61	1.326	0.709	0.535
V_R^*	0.609–7.9023	2.411	1.371	1.634
V_L^*	0.233–4.59	1.952	1.036	0.6023

To obtain a deeper understanding of the effective parameter variations, histograms for non-dimensional parameters [see Equation (19)] are depicted in Figure 2 from (a) to (h). Overall, the histograms can take advantage of the laboratory observations in which the frequency distributions for the five parameters affect the pipeline scouring while considering the simultaneous influence of currents and waves. Moreover, the statistical analysis of histograms gives investigators new insights into considering unseen ranges of effective factors. Figure 2a illustrates the frequency of the e/D ratio with a rather symmetric distribution, showing 13 values of $e/D = 0.3$ as the maximum frequency during experiments. In Figure 2b, the frequency of KC parameter has a fragmented distribution with the maximum proportion for $KC = 7.9$. Figure 2c shows that the distribution of m is completely partitioned without symmetrical patterns. The minimum and maximum frequencies were 3 and 12, respectively, which were related to $m = 0.363$ and 0.683 . As seen in Figure 2d, the Shields' parameter due to waves, θ_W , was thoroughly distributed, starting from $\theta_W = 0.05$ with four frequencies to ending $\theta_W = 0.245$ with seven frequencies. In Figure 2e, the Shields' parameter was not symmetrically distributed due to the currents (θ_C). For instance, three levels of θ_C (i.e., 0.0233, 0.0656, 0.1079) have the same value of frequency (10). Figure 2 from (f) to (h) indicates that the patterns of dimensionless scour rate distributions are not symmetrical in a way that with an increase in V_H^* (or V_R^* and V_L^*) and the frequency generally had a downward trend. Consequently, it can be inferred

from the fragmentation of histograms that five dimensionless parameters (as mentioned in Equation (17)) played a significant role in the scouring experiments.

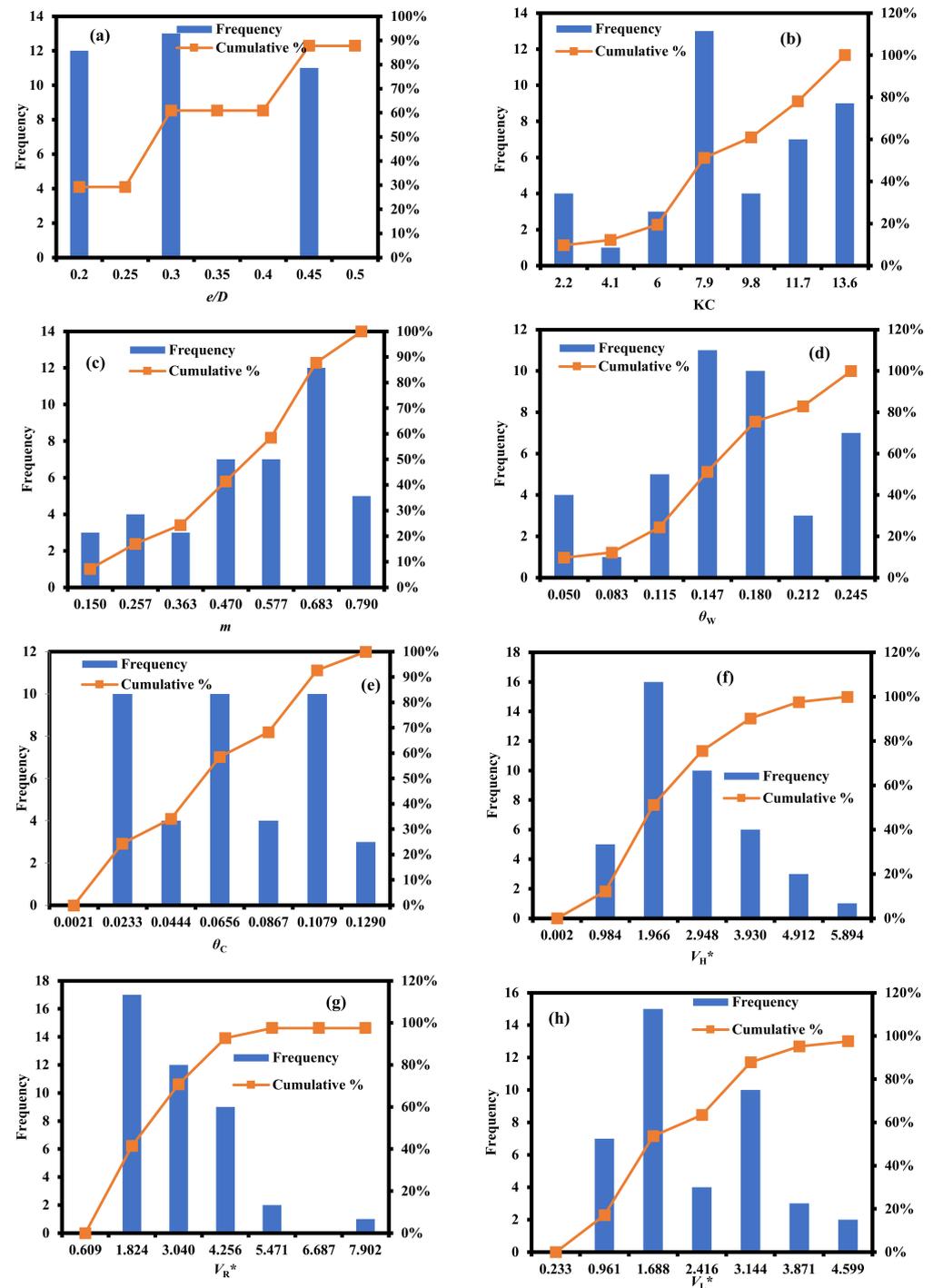


Figure 2. Distribution patterns for the effective parameters regarded in this research: (a) ratio of the pipeline embedded depth to pipeline diameter, e/D ; (b) Keulegan–Carpenter number, KC ; (c) ratio of the maximum undisturbed orbital rate at the sea bottom just above the boundary of the waves to the flow velocity due to currents, m ; (d) Shields’ parameter due to waves, θ_w ; (e) Shields’ parameter due to currents, θ_C ; (f) dimensionless scouring propagation rate around offshore pipeline at the longitudinal direction, V_H^* ; (g) dimensionless scouring propagation rate around offshore pipeline at the right-hand shoulder of the pipeline, V_R^* ; (h) dimensionless scouring propagation rate around offshore pipeline at the left-hand shoulder of the pipeline, V_L^* .

4. Implementation of Soft Computing Models

4.1. Gene-Expression Programming

GEP, classified into Evolutionary Computing (EC) models, is inextricably bound up with genetic algorithms (GA) and genetic programming (GP). The GEP can potentially provide a relationship to interpret variations of dataset patterns based on evolutionary computing principles [19,20]. In the case of computer programming, GEP generally consists of complicated tree structures, providing predictive models. These models can be efficiently trained by varying their sizes, shapes, and composition, like a chromosome. In the GEP model, simple linear chromosomes of fixed length are considered living organisms in the encoded computer programs. Therefore, GEP is grouped into a genotype–phenotype system in which a simple genome and a complicated phenotype can efficiently incorporate into (i) saving and transmitting the genetic data and (ii) finding variations of dataset patterns for various real-world problems. In the structure of the GEP model, evolutionary computing employs populations of individuals, selects individuals according to goodness, and determines genetic variation by one or more genetic operators. In the GEP model, the linear chromosomes of fixed length were inherited from GA and consequently, the expressive parse trees of various sizes and shapes were rooted in GP. The existing chromosomes of the GEP model usually consisted of more than one gene of equal length. The structure of each gene is like a sub-expression tree (sub-ET), interacting with each other in various ways. In the final step of the GEP model, the sub-ETs are linked by addition or other functions such as division, multiplication, and subtraction.

In this study, GeneXproTools software was employed to implement the GEP model. The performance of the GEP was assessed by fitness function values. Calibration and validation stages were evaluated by mean square error (MSE). As shown in Table 2, the GEP model was performed seven times for each scour propagation rate. In Table 3, correlation values of the GEP performances for the seven combinations of non-dimensionless parameters were evaluated. According to Table 3, the correlation coefficient yielded in the calibration and validation stages indicated that performances of the GEP model for Set 5 (0.7796 and 0.7029), Set 6 (0.6645 and 0.7425), and Set 7 (0.7552 and 0.7833) had the best results in the prediction of V_H^* , V_R^* , and V_L^* , respectively. This study presented the results of the best GEP performances for three-dimensional scour rates. Table 4 gives characterizations of the GEP performance for the 3D scour rate propagation.

Table 3. Correlations of GEP performance in the calibration and validation stage for 3D scour rates.

Set No.	V_H^*		V_R^*		V_L^*	
	Calibration	Validation	Calibration	Validation	Calibration	Validation
1	0.5921	0.748	0.6482	0.8345	0.6115	0.4135
2	0.6739	0.5056	0.5175	0.5768	0.5692	0.6290
3	0.7481	0.6118	0.6160	0.7913	0.5090	0.6619
4	0.6151	0.8933	0.59359	0.8438	0.6587	0.7778
5	0.7796	0.7029	0.6755	0.6251	0.7823	0.7071
6	0.7185	0.8477	0.6645	0.7425	0.73617	0.6746
7	0.4607	0.4867	0.46055	0.6911	0.7552	0.7833

small to continue. Regarding this investigation, the MARS model predicted 3D scour rates propagation by Equation (23),

$$V_H^*, V_L^*, V_R^* = \chi_0 + \sum_{j=1}^{NBF} \chi_j \cdot BF_j(e/D, \theta_C, \theta_W, m, KC) \tag{23}$$

in which χ_0 , χ_j , BF, and NBF are the bias, the constant coefficients related to the basis functions, the basis functions, and the number of basis functions, respectively. The forward pass occasionally creates an overfitted model. After that, the backward pass is carried out to prune the primarily-calibrated MARS model. This methodology eliminates terms one by one, having insignificant effective terms at each stage until it yields the most promising sub-model. The performance of model subsets is assessed by the generalized cross-validation (GCV) criterion that has frequently been applied in the different fields of machine learning techniques [22,23].

All the steps of the MARS model development were implemented by a computer programming code in MATLAB software. The best performance of the MARS model for V_H^* prediction was related to Set 2 with a GCV value of 1.498, whereas the input combination of Set 6 resulted in the most promising predictions for both V_L^* (GVC = 1.139) and V_R^* (GVC = 2.297). To further extend, Table 5 demonstrates the best performance of MARS models for seven combinations of dimensionless parameters in the calibration and validation for V_H^* (0.7705 and 0.7513), V_R^* (0.7325 and 0.5906), and V_L^* (0.8439 and 0.8949). The total effective number of parameters for all the best MARS models was 11. Consequently, the following relationships are the best MARS models for the prediction of the 3D scour propagation rates below pipelines:

$$V_H^* = 4.51 + 27.377 \times BF_1 - 50.568 \times BF_2 - 24.35 \times BF_3 - 20.731 \times BF_4 \tag{24}$$

$$V_L^* = -0.2622 + 141.27 \times BF_1 - 265.05 \times BF_2 + 33.301 \times BF_3 + 0.51102 \times BF_4 - 199.59 \times BF_5 + 93.833 \times BF_6 \tag{25}$$

$$V_R^* = 3.1098 + 51.833 \times BF_1 - 9.2312 \times BF_2 - 56.263 \times BF_3 - 2925 \times BF_4 \tag{26}$$

in which BF relationships related to Equations (24)–(26) are given in Table 6. As seen in Table 6, all the three dimensionless parameters $(1 - e/D, F, (1 - m)\theta_W + m\theta_C)$ and four dimensionless parameters $(1 - e/D, (1 - m)\theta_W, m\theta_C, KC)$ were incorporated into the prediction of V_H^* and V_L^* , respectively; whereas KC parameter has no contribution to the estimation of V_R^* . All the coefficients in Equations (24)–(26) were obtained using the Particle Swarm Optimization (PSO) algorithm, providing MSE = 0.636, 0.336, and 1.028 as the most promising results.

Table 5. Correlations of MARS performance in the calibration and validation stages for 3D scour rates.

Set No.	V_H^*		V_R^*		V_L^*	
	Calibration	Validation	Calibration	Validation	Calibration	Validation
1	0.837	0.558	0.8113	0.4777	0.6361	0.44
2	0.7705	0.7513	Nan	Nan	Nan	Nan
3	0.8370	0.5586	0.3945	0.5965	0.6144	0.4405
4	0.6142	0.5622	0.3945	0.5965	0.7744	0.4408
5	0.7943	0.4148	0.7392	0.2263	0.6835	0.5542
6	0.7578	0.6955	0.7325	0.5903	0.8439	0.8949
7	0.4396	0.6438	0.8475	0.3638	0.6317	0.7374

Table 6. Coefficients and basis functions for MARS models in scour depth prediction.

Basis Functions Used in the Prediction of V_H^*	
BF ₁	$\max(0, 0.61751 - F)$
BF ₂	$\max(0, 0.11276 - (1 - m)\theta_W - m\theta_C)$
BF ₃	$\max(0, 0.73254 - (1 - e/D))$
BF ₄	$\max(0, 0.7 - (1 - e/D)) \times \max(0, F - 0.30369)$
Basis functions used in the prediction of V_R^*	
BF ₁	$\max(0, m\theta_C - 0.03862)$
BF ₂	$\max(0, m\theta_C - 0.064886)$
BF ₃	$\max(0, 0.064886 - m\theta_C)$
BF ₄	$\max(0, 6.3 - KC)$
BF ₅	$\max(0, 0.7 - (1 - e/D)) \times \max(0, 0.078754 - ((1 - m)\theta_W))$
BF ₆	$\max(0, (1 - e/D) - 0.6) \times \max(0, (1 - m)\theta_W - 0.068215)$
Basis functions used in the prediction of V_L^*	
BF ₁	$\max(0, m\theta_C - 0.03044)$
BF ₂	$\max(0, 0.7 - (1 - e/D))$
BF ₃	$\max(0, 0.0478720 - (1 - m)\theta_W)$
BF ₄	$\max(0, (1 - m)\theta_W - 0.0478720) \times \max(0, m\theta_C - 0.017761)$

4.3. M5 Model Tree

M5 model tree [24] is a generalization of regression decision trees, where multilinear regression models are attached to the leaf (terminal) nodes to perform regression tasks. One problem with regular regression decision trees is their high computational cost for datasets with high dimensionality [25,26]. The M5 model tree takes a divide-and-conquer approach to solve this issue by splitting the original high dimensional parameter space into parameter sub-spaces for each of which a separate regression model is trained. The final solution is then obtained by combining these sub-problems. The M5 algorithm implements this solution in three steps [27–29]. It first grows the tree, followed by pruning the tree to avoid overfitting. Finally, a smoothing process is used to account for the discontinuities that occur between linear models, which thereby results in better predictions. The structure of the tree built by the M5 algorithm looks like a binary decision tree with a splitting criterion based on the standard deviation reduction (SDR) of the target values at each node [30]. For each sub-space, when the data reaches a node, it gets divided into two groups.

Thus, the M5 algorithm uses the standard deviation of the target values at a node as an error measure at that node for each data attribute. Among all possible splits, the M5 model chooses the attribute as the splitting parameter that maximizes SDR. As the branches of the tree grow, the SD values of data in the child nodes get smaller than those in the parent nodes. The splitting process continues until the SD of the target values in a node change slightly or a few examples below a certain threshold remain. At this point, the tree has reached its leaf node. This process could lead to a large tree that needs to be pruned to prevent over-fitting. For this purpose, the sub-trees in the overgrown model tree are replaced by a linear regression model. In the end, the smoothing process is applied in such a way that the linear models at the leaf nodes have nearly closed to the predicted values.

In the present research, Weka3.9 software has been employed to implement M5MT for the evaluation of 3D scour propagation rates around offshore pipelines exposed to simultaneous impacts of current and wave. M5MT is implemented for all combinations of

dimensionless parameters to provide multilinear regression models. For instance, M5MT presents the following relationship to predict V_H^* for Set 1:

$$V_H^* = a_0 + a_1 \times \theta_C + a_2 \times \theta_W + a_3 \times \left(1 - \frac{e}{D}\right) + a_4 \times m + a_5 \times KC \quad (27)$$

in which a_0 is the bias term and a_1 to a_5 are weighting coefficients that are computed by the least square technique.

In the development phases, tree structures of M5MT were pruned and smoothed. Results of Table 7 indicated that the correlation coefficient yielded in calibration and validation phases provided the most efficient performances of the M5MT model for Set 5 (0.7650 and 0.7810), Set 5 (0.7450 and 0.7956), and Set 6 (0.6916 and 0.7760) in the estimation of V_H^* , V_R^* , and V_L^* , respectively. The most promising results of M5MT performances demonstrated that using Set 5 (e/D , F , m , θ_W , θ_C) provided V_H^* and V_R^* predictions with the most accurate results, whereas the most precise predictions of V_L^* values were simulated by Set 6. As seen in Table 8, three linear equations along with three rules were used to provide V_H^* predictions for Set 5. The performance of M5MT indicates that the F parameter has no role to play in the prediction of scouring rates in the longitudinal direction. Similarly, Table 9 indicated that the F parameter was not incorporated into the estimation of V_R^* and consequently, two rules were used to provide a linear regression equation along with the splitting parameter of θ_C (0.087). In Table 10, equations extracted by M5MT demonstrated that when m is combined by two types of Shields' parameters, the KC parameter has no role in the estimation of V_L^* .

Table 7. Correlations of M5MT performance in the calibration and validation stage for 3D scour rates.

Set No.	V_H^*		V_R^*		V_L^*	
	Calibration	Validation	Calibration	Validation	Calibration	Validation
1	0.65	0.75	0.361	0.681	0.4866	0.6394
2	0.434	0.5962	0.361	0.680	0.4866	0.6394
3	0.756	0.775	0.361	0.680	0.486	0.639
4	0.4409	0.5960	0.7887	0.670	0.4866	0.6394
5	0.765	0.781	0.745	0.7956	0.6794	0.7740
6	0.753	0.774	0.723	0.820	0.6916	0.7760
7	0.4339	0.596	0.361	0.681	0.4944	0.6440

Table 8. Lists of M5MT details in V_H^* estimates.

Rules of M5MT#1 with Focusing on Pruning and Smoothing Phases
Rule: 1 IF $e/D \leq 0.35$ $\theta_C \leq 0.087$ THEN $V_H^* = -5.6911 \times \theta_W + 27.9909 \times \theta_C - 6.8389 \times m - 2.8633 \times e/D + 6.0401$
Rule: 2 IF $e/D > 0.35$ THEN $V_H^* = -6.015 \times e/D + 4.3233$
Rule: 3 $V_H^* = +4.5309$

Table 9. Lists of M5MT details in V_R^* estimates.

Rules of M5MT#1 with Focusing on Pruning and Smoothing Phases	
Rule: 1	
IF	$\theta_C \leq 0.087$
THEN	$V_R^* = -4.3517 \times \theta_W + 17.5692 \times \theta_C - 6.868 \times m - 6.0857 \times e/D + 7.0983$
THEN	$V_R^* = -9.5347 \times e/D + 6.8673$

Table 10. Lists of M5MT details in V_L^* estimates.

Rules of M5MT#1 with Focusing on Pruning and Smoothing Phases	
Rule: 1	
IF	$1 - e/D > 0.65$ $m\theta_C \leq 0.052$
THEN	$V_L^* = 2.4945 \times (1 - e/D) + 8.6366 \times (1 - m)\theta_W + 8.5379 \times m\theta_C - 0.6196$
Rule: 2	
	$V_L^* = -10.8666 \times (1 - e/D) - 4.5984$

4.4. Evolutionary Polynomial Regression

The Evolutionary Polynomial Regression (EPR) is a model selection technique that leverages an evolutionary algorithm to find an optimum pseudo-polynomial-based model for an underlying physical system [31]. The implementation of the EPR Multi-Objective Genetic Algorithm (MOGA) consists of a two-stage process [32–34]. First, an evolutionary algorithm is used to search for model structures. Second, a linear regression algorithm is applied to find the optimum model parameters using the least-squares technique. This multi-objective approach leads to the search for optimum models while maintaining a balance between prediction accuracy and model complexity. On the one hand, higher prediction accuracy is achieved by minimizing the sum of squared error (SSE) between real data and predicted values. On the other hand, the model complexity is measured by the number of pseudo-polynomial terms and/or the number of input variables in the model. Therefore, the EPR MOGA approach allows one to evaluate what level of model complexity results in a higher accuracy [35].

In the present study, the following mathematical expression was applied to develop equations in order to estimate 3D scouring propagation rates:

$$V_H^*, V_R^*, V_L^* = \text{bias} + \sum_{j=1}^{\text{NOT}} \left[L_j \times (\theta_C)^{\text{ES}(j,1)} \times \left(\frac{e}{D}\right)^{\text{ES}(j,2)} \times (\theta_W)^{\text{ES}(j,3)} \times (\text{KC})^{\text{ES}(j,4)} \times (m)^{\text{ES}(j,5)} \right. \\ \left. \times Q\left((\theta_C)^{\text{ES}(j,6)} \times \left(\frac{e}{D}\right)^{\text{ES}(j,7)} \times (\theta_W)^{\text{ES}(j,8)} \times (\text{KC})^{\text{ES}(j,9)} \times (m)^{\text{ES}(j,10)}\right) \right] \quad (28)$$

Equation (28) is an example of the general expression of EPR when fed by Set 1. *NOT* is the maximum mathematical terms, L_j is a set of coefficients, Q is a user-defined function, and *ES* function is a range of exponents explored by the EPR model. According to the most recent research works on scouring around pipelines, applying natural logarithmic as an inner function indicated the best performance in the evaluation of the scouring propagation rates due to waves and currents [15,16]. During the development of the EPR model, the number of generations was obtained in three levels: 6480 (Set 2 and Set 7), 8640 (Set 4, Set 3, and Set 6), and 10,800 (Set 2 and Set 7) for various combinations of dimensionless parameters.

Generally, 21 EPR expressions were yielded for seven combinations of dimensionless inputs and three scour rates. Results of the EPR model’s performance are presented in Table 11. As indicated in Table 11, the correlation coefficient values which were obtained in the calibration and validation stages provided the most promising performances of EPR expressions for Set 6 (0.8794 and 0.8842), Set 5 (0.7458 and 0.6651), and Set 5 (0.8725 and 0.7853) in the evaluation of V_H^* , V_R^* , and V_L^* , respectively. Moreover, Table 11 applying

the fifth combination of inputs ($e/D, F, m, \theta_W, \theta_C$) yielded the best results of EPR models for V_R^* and V_L^* predictions whereas the most accurate predictions of V_H^* were obtained by the sixth combination ($1 - e/D, KC, (1 - m)\theta_W, m\theta_C$). The optimum mathematical expressions given by EPR runs were given as:

$$\begin{aligned}
 V_H^* = & 43.1683 + 239316 \left(\frac{m\theta_C}{1-\frac{e}{D}}\right)^{0.5} \text{Ln} \left[\left(\frac{(1-m)\theta_W}{1-\frac{e}{D}}\right)^{1.5} \right] + 1837.1028 \frac{(m\theta_C)^{1.5}}{KC} \text{Ln} \left[(KC)^2 \right] \\
 & + 16.7884((1 - m)\theta_W)^{0.5} \text{Ln} \left[\left(\frac{1-\frac{e}{D}}{KC}\right)^2 \right] + 0.56651((1 - m)\theta_W)^{1.5} (KC)^2 \text{Ln} \left[\frac{(1-\frac{e}{D})^{0.5}}{((1-m)\theta_W)^{1.5}} \right] \\
 & + 3.2992(1 - \frac{e}{D})^{0.5} \text{Ln} \left[\frac{((1-m)\theta_W)(m\theta_C)^{0.5}}{1-\frac{e}{D}} \right] + 4831.8475(1 - \frac{e}{D})^2((1 - m)\theta_W)(m\theta_C)^2
 \end{aligned} \tag{29}$$

$$\begin{aligned}
 V_R^* = & 0.076855 + 0.57687 \text{Ln} \left[\frac{1}{(\theta_C)^{1.5}} \right] + 0.28563 \text{Ln} \left[(\theta_W)^{1.5} \right] \\
 & + 3.4109 \left(\frac{e}{D}\right)^2 \text{Ln} \left[(\theta_W)^{1.5} \right] + 439.956(\theta_C)^2 F
 \end{aligned} \tag{30}$$

$$\begin{aligned}
 V_L^* = & 0.70796 + 40.3526 \text{Ln} \left[F^{0.5} \right] + 2.8658 \text{Ln}[\theta_W] + 2.9314 \times F \times \text{Ln} \left[\theta_C \left(\frac{e}{D}\right)^{0.5} \right] \\
 & + 826.7353(\theta_C)^2 (F)^{0.5} \text{Ln} \left[\frac{F}{\left(\frac{e}{D}\right)^{0.5}} \right] + 68.351(\theta_W)^{0.5} \text{Ln} \left[\frac{m}{(\theta_C)^{0.5}} \right] + 1.591 \frac{\theta_W}{(\theta_C)^{0.5}} \text{Ln} \left[\frac{1}{\left(\frac{e}{D}\right)^{0.5}} \right]
 \end{aligned} \tag{31}$$

Table 11. Correlations of EPR performance in the calibration and validation stages for 3D scour rates.

Set No.	V_H^*		V_R^*		V_L^*	
	Calibration	Validation	Calibration	Validation	Calibration	Validation
1	0.7780	0.4137	0.7413	0.6800	0.79173	0.2661
2	0.6512	0.8881	0.5898	0.8194	0.6771	0.7373
3	0.6887	0.6018	0.7219	0.6467	0.8788	0.3685
4	0.6382	0.7297	0.5323	0.8396	0.6921	0.7264
5	0.6887	0.6031	0.7458	0.6651	0.8725	0.7853
6	0.8794	0.8842	0.7432	0.6556	0.7378	0.6403
7	0.6258	0.5769	0.5800	0.8285	0.6956	0.6735

According to the Equations (29)–(31), all the input parameters in Set 5 and Set 6 were incorporated into the prediction of V_H^* and V_L^* , respectively, whereas velocity ratio [$m = U_C/(U_C + U_W)$] has no role to play in the estimation of V_R^* parameter (as seen in Equation (30)).

5. Results and Discussion

5.1. Definition of Statistical Indices

To investigate the evaluation of soft computing models’ efficiency for both calibration and validation phases, coefficient of correlation (CC), root mean square error (RMSE), mean absolute percentage error (MAPE), scatter index (SI), and discrepancy ratio (DR) have been utilized. These statistical measures are defined as:

$$\text{CC} = \frac{\sum_{i=1}^N \left[\left(V_{H,R,L}^{*(i)}(\text{Est}) - \bar{V}_{H,R,L}^{*(i)}(\text{Est}) \right) \left(V_{H,R,L}^{*(i)}(\text{Exp}) - \bar{V}_{H,R,L}^{*(i)}(\text{Exp}) \right) \right]}{\sqrt{\left[\sum_{i=1}^N \left(V_{H,R,L}^{*(i)}(\text{Est}) - \bar{V}_{H,R,L}^{*(i)}(\text{Est}) \right)^2 \times \sum_{i=1}^N \left(V_{H,R,L}^{*(i)}(\text{Exp}) - \bar{V}_{H,R,L}^{*(i)}(\text{Exp}) \right)^2 \right]}} \tag{32}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N \left(V_{H,R,L}^{*(i)}(\text{Est}) - V_{H,R,L}^{*(i)}(\text{Exp}) \right)^2}{N}} \tag{33}$$

$$\text{MAPE} = 100 \times \frac{\sum_{i=1}^N \left| \frac{V_{H,R,L}^{*(i)}(\text{Est}) - V_{H,R,L}^{*(i)}(\text{Exp})}{V_{H,R,L}^{*(i)}(\text{Exp})} \right|}{N} \quad (34)$$

$$\text{SI} = \frac{\sqrt{\left(\frac{1}{N}\right) \sum_{i=1}^N \left[\left(V_{H,R,L}^{*(i)}(\text{Est}) - \bar{V}_{H,R,L}^{*(i)}(\text{Est}) \right) - \left(V_{H,R,L}^{*(i)}(\text{Exp}) - \bar{V}_{H,R,L}^{*(i)}(\text{Exp}) \right) \right]^2}}{\bar{V}_{H,R,L}^{*(i)}(\text{Exp})} \quad (35)$$

$$\text{DR} = \frac{\sum_{i=1}^N \frac{V_{H,R,L}^{*(i)}(\text{Est})}{V_{H,R,L}^{*(i)}(\text{Exp})}}{N} \quad (36)$$

where $V_{H,R,L}^{*(i)}(\text{Exp})$, $V_{H,R,L}^{*(i)}(\text{Est})$, $\bar{V}_{H,R,L}^{*(i)}$, and N are the experimental values of non-dimensional scour propagation rate, the estimated values of non-dimensional scour propagation rate, the average values of non-dimensional scour propagation rate, and the number of scouring tests, respectively. In the case of correlation values, when CC values vary between +1 and -1, this means that the soft computing model indicates the best performance. In contrast, if CC is equal to 0, the worst performance is met. RMSE, MAPE, and SI values are known as error functions, varying from 0 to $+\infty$. If the DR value is equal to 1, the soft computing technique indicates the most efficient performance. Where DR is greater than one, it shows over predictions, otherwise, if $\text{DR} < 1$, underprediction is met.

5.2. Statistical Performance of Soft Computing Techniques

Table 12 presents quantitative results of the best performances of soft computing models for both calibration and validation in the prediction of V_H^* . It can be inferred from statistical measures that the EPR model [Equation (29)] indicated the highest accuracy ($\text{CC} = 0.879$, $\text{RMSE} = 0.595$, and $\text{MAPE} = 0.224$) in the calibration phase when compared to other soft computing techniques. According to Table 12, GEP model ($\text{CC} = 0.779$, $\text{RMSE} = 0.801$, and $\text{MAPE} = 0.348$) stood at the second rank in terms of accuracy, followed by MARS ($\text{CC} = 0.770$, $\text{MAPE} = 0.398$, and $\text{DR} = 1.167$) and MTM5 ($\text{CC} = 0.765$, $\text{MAPE} = 0.474$, and $\text{DR} = 1.286$). Over/under predictions of the soft computing models were quantified by DR criterion. In this way, Equation (29) obtained by the EPR model has the lowest level of overprediction ($\text{DR} = 1.046$) whereas predictions of V_H^* were indicative of the highest level with $\text{DR} = 1.286$. Results of quantitative comparisons from both validation and calibration phases indicated that the EPR model had no sufficient potential of estimating V_H^* . In Table 12, values of three statistical criteria (RMSE, MAPE, and SI) showed that MARS could be chosen as the best soft computing model in the validation phase ($\text{RMSE} = 0.929$, $\text{MAPE} = 0.498$, and $\text{SI} = 0.516$). GEP model provided relatively accurate predictions ($\text{RMSE} = 1.239$, $\text{MAPE} = 0.565$, and $\text{DR} = 1.403$) than those obtained by M5MT ($\text{RMSE} = 1.543$, $\text{MAPE} = 0.692$, and $\text{DR} = 1.661$). In addition, the statistical criteria CC, RMSE, and DR indicated that EPR ($\text{CC} = 0.884$, $\text{RMSE} = 1.509$, and $\text{DR} = 0.927$) had a better potential for predicting V_H^* in comparison with M5MT. Overall, EPR ($\text{DR} = 0.927$) and MARS ($\text{DR} = 0.897$) had underprediction whereas M5MT ($\text{DR} = 0.1661$) and GEP ($\text{DR} = 1.403$) techniques indicated overprediction.

The performance of soft computing models for the estimation of V_H^* in calibration and validation stages is depicted in Figure 3. In Figure 3a, all the soft computing techniques overpredicted the observed values of V_H^* between 0.5 and 1.5 for the calibration stage. For $V_H^* = 5.89$, the MARS model had significant underprediction in comparison with EPR and GEP models. In the validation stage, Figure 3b illustrated overprediction for observed values of V_H^* between 1 and 2.5. Additionally, for $V_H^* = 5.89$, the MARS model indicated remarkable underprediction when compared to other soft computing models.

Table 12. The best statistical performances of the soft computing models for prediction of V_H^* .

Soft Computing Models	Calibration Stage				
	CC	RMSE	MAPE	SI	DR
MARS (Developed by Set 2)	0.770	0.987	0.398	0.424	1.167
EPR (Developed by Set 6)	0.879	0.595	0.224	0.256	1.046
M5MT (Developed by Set 5)	0.765	0.817	0.474	0.348	1.286
GEP (Developed by Set 5)	0.779	0.801	0.348	0.344	1.183
Soft Computing Models	Validation Stage				
	CC	RMSE	MAPE	SI	DR
MARS (Developed by Set 2)	0.751	0.929	0.498	0.516	0.897
EPR (Developed by Set 6)	0.884	1.509	0.796	0.832	0.927
M5MT (Developed by Set 5)	0.781	1.543	0.692	0.529	1.661
GEP (Developed by Set 5)	0.703	1.239	0.565	0.538	1.403

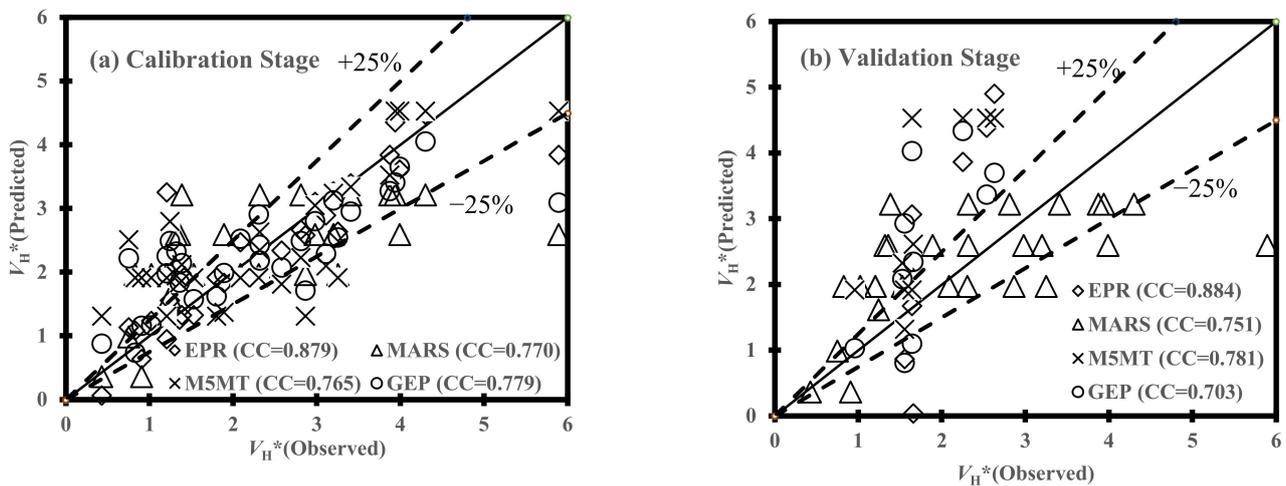


Figure 3. Performance of soft computing models in the prediction of V_H^* for (a) calibration and (b) validation stages.

In Table 13, the statistical evaluation of calibration stages indicated that EPR and M5MT provided relatively the same performance in the estimation of V_R^* . More specifically, the accuracy metrics of the EPR model [Equation (30)] were $CC = 0.746$, $RMSE = 1.003$, and $MAPE = 0.308$, which were relatively close to those of the M5MT model ($CC = 0.745$, $RMSE = 1.006$, and $MAPE = 0.312$). According to Table 13, SI and DR proved marginal improvement of EPR ($SI = 0.393$ and $DR = 1.120$) over M5MT ($SI = 0.394$ and $DR = 1.163$). Additionally, Table 13 indicated that the MARS model stood at the third stage on the account of accuracy with CC, RMSE, and SI values of 0.732, 1.041, and 0.408, respectively. Overall, the GEP model resulted in the lowest value of correlation (0.664) and the highest values of error functions ($RMSE = 1.142$ and $SI = 0.447$). In the validation stage, GEP model demonstrated slightly better performance ($CC = 0.742$, $MAPE = 0.315$, and $SI = 0.223$) in the estimation of V_R^* than MARS ($CC = 0.590$, $MAPE = 0.384$, and $SI = 0.446$), and EPR ($CC = 0.665$, $MAPE = 0.764$, and $SI = 0.878$) did. With reference to all statistical measures except CC value, M5MT provided V_R^* predictions with comparatively lower accuracy in terms of SSE (0.1322) and MAE (0.0944) than those reported by the GEP model. As seen in Figure 4a, all the predicted values of V_R^* in the calibration stage were almost placed in $\pm 25\%$ error lines. Figure 4a depicted significant overprediction for the observed V_R^* of 1.321. In Figure 4b, the qualitative performance of the validation stage indicated relatively successful efficiency because the majority of data points were placed inside $\pm 25\%$

error lines. More specifically, Figure 4b indicated that EPR, MARS, and M5MT provided overprediction in the state of observed V_R^* varying from 1.334 to 3.277.

Table 13. The best statistical performances of the soft computing models for prediction of V_R^* .

Soft Computing Models	Calibration Stage				
	CC	RMSE	MAPE	SI	DR
MARS (Developed by Set 6)	0.732	1.041	0.388	0.408	1.111
EPR (Developed by Set 5)	0.746	1.003	0.308	0.393	1.120
M5MT (Developed by Set 5)	0.745	1.006	0.312	0.394	1.163
GEP (Developed by Set 6)	0.664	1.142	0.373	0.447	1.174
Soft Computing Models	Validation Stage				
	CC	RMSE	MAPE	SI	DR
MARS (Developed by Set 6)	0.590	0.981	0.384	0.446	1.632
EPR (Developed by Set 5)	0.665	2.236	0.764	0.878	1.234
M5MT (Developed by Set 5)	0.795	1.388	0.581	0.527	1.415
GEP (Developed by Set 6)	0.742	1.142	0.315	0.223	1.259

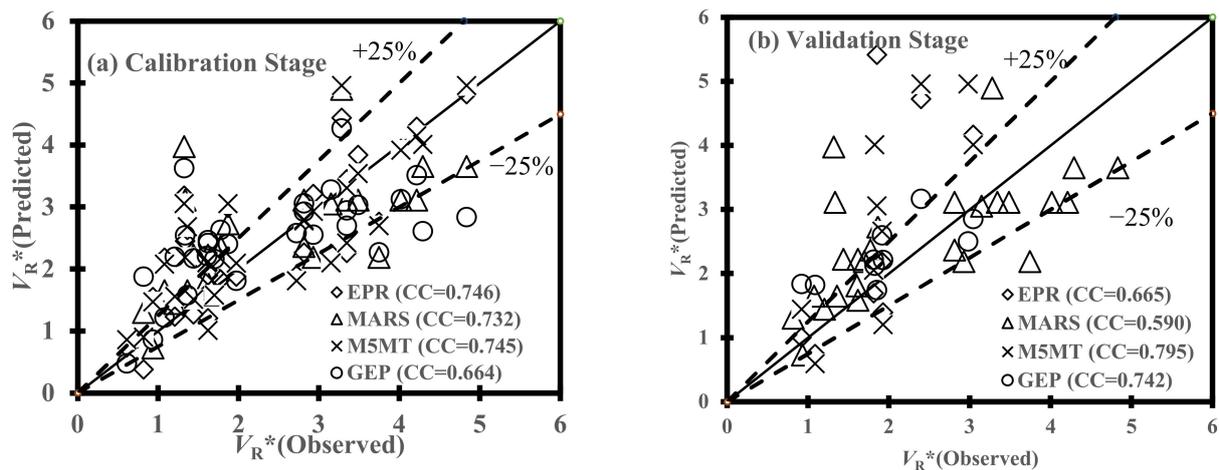


Figure 4. Performance of soft computing models in the prediction of V_R^* for (a) calibration and (b) validation stages.

As seen in Table 14, the EPR model indicated the most promising performance (CC = 0.872, RMSE = 0.544, and MAPE = 0.290) in the calibration stage when compared with MARS (CC = 0.844, RMSE = 0.602, and MAPE = 0.314) and GEP (CC = 0.755, RMSE = 0.739, and MAPE = 0.447). In the calibration phase, M5MT stood at the lowest level of precision according to CC = 0.691, RMSE = 0.807, and MAPE = 0.598. Additionally, SI values presented in Table 14 demonstrated the superiority of the EPR model (SI = 0.259) over other soft computing models such as MARS (SI = 0.286), M5MT (SI = 0.383), and GEP (SI = 0.352). Although the EPR model had the best performance in the calibration phase, the MARS model [Equation (25)] provided the most efficient performance in the prediction of V_R^* for the validation stage. All the statistical measures obtained by the performance of MARS (CC = 0.895, RMSE = 0.440, and MAPE = 0.248) in the validation phases indicated its superiority over other soft computing techniques. In addition to this, SI and DR criteria proved the most efficient performance of the MARS model. In contrast, all the statistical measures except CC value showed that EPR had the lowest performance (RMSE = 2.203, MAPE = 0.991, and MAPE = 0.991). Moreover, the results of the validation stage demonstrated that the GEP model had a better assessment in the estimation of V_L^* compared with

M5MT (RMSE = 1.197, MAPE = 0.679, and SI = 0.991). From DR values, it can be said that the MARS model had a lower overprediction of V_L^* (DR = 1.139) and followed by GEP (DR = 1.231), M5MT (DR = 1.666), and EPR (1.991). To make deeply qualitative comparisons in the calibration stage, Figure 5a indicated that all the soft computing techniques had remarkable underprediction for the observed value $V_L^* = 0.427\text{--}1.321$. As seen in Figure 5a, M5MT had no required capability of prediction in the $V_L^* = 2.565$ in comparison with other soft computing models. Additionally, for $V_L^* = 0.984\text{--}1.515$, the GEP model had overpredictions in the less level (see Figure 5b) than other soft computing models.

Table 14. The best statistical performances of the soft computing models for prediction of V_L^* .

Soft Computing Models	Calibration Stage				
	CC	RMSE	MAPE	SI	DR
MARS (Developed by Set 6)	0.844	0.602	0.314	0.286	1.108
EPR (Developed by Set 5)	0.872	0.544	0.290	0.259	1.129
M5MT (Developed by Set 6)	0.691	0.807	0.598	0.383	1.374
GEP (Developed by Set 7)	0.755	0.739	0.447	0.352	1.144
Soft Computing Models	Validation Stage				
	CC	RMSE	MAPE	SI	DR
MARS (Developed by Set 6)	0.895	0.440	0.248	0.253	1.139
EPR (Developed by Set 5)	0.785	2.203	0.991	0.991	1.991
M5MT (Developed by Set 6)	0.776	1.197	0.679	0.439	1.666
GEP (Developed by Set 7)	0.783	0.674	0.358	0.369	1.231

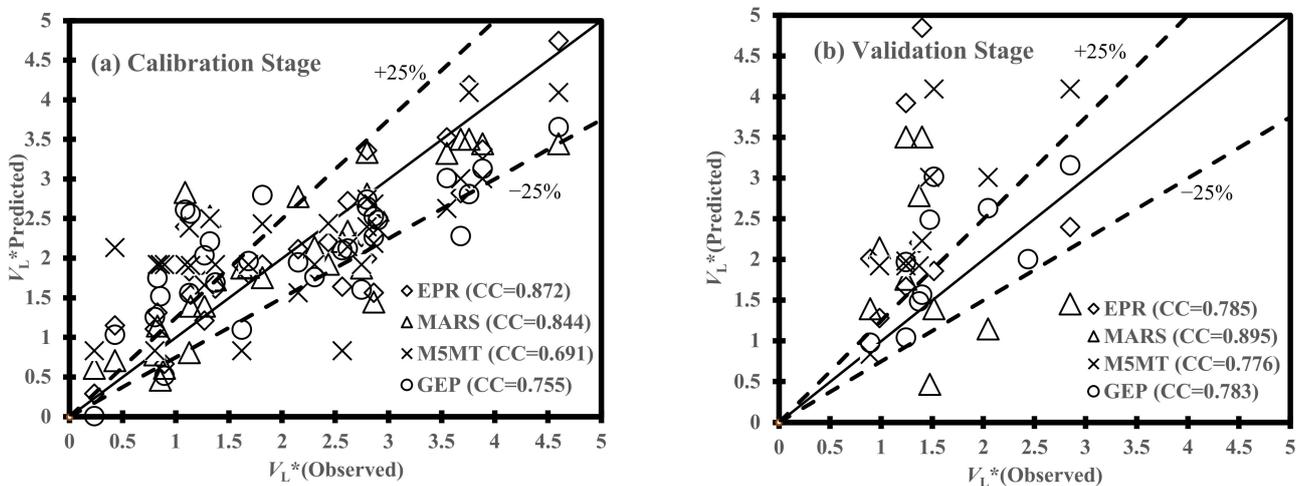


Figure 5. Performance of soft computing models in the prediction of V_L^* for (a) calibration and (b) validation stages.

5.3. Complexity of Soft Computing Techniques Performance

In this part of the research, reasonable judgment between the complexity of soft computing models and their performance in the assessment of 3D scour rates were studied. As mentioned in the descriptions of experimental data, the dataset related to the combined wave and current has not yet been applied for the estimation of 3D scour rates. This lack of experimental data imposed some limitations on our model developments, which thereby had an unavoidable impact on the accuracy level of the resulting soft computing models.

In the prediction of V_H^* , the EPR model was efficiently calibrated using Set 6 and consequently, this model provided complex expressions compared to M5MT and MARS [Equation (24)]. In fact, the sixth combination of dimensionless inputs (see Table 1) contains input parameters (KC, m, θ_W), which are representative of wave impacts on the scour propagation rates. On the contrary, the best linear regression by M5MT (see Table 8) does not consist of the KC parameter, and additionally, linear equations by M5MT are simpler than expressions by EPR and GEP models. Equation (24) obtained by the MARS model had the most accurate assessment in comparison with using other performances by input combination. This means that the number of input parameters related to Set 4 ($1 - e/D, F(1 - m)\theta_W + m\theta_C$) is small and three parameters (θ_W, m, θ_C) were combined and converted into one dimensionless parameter.

For V_R^* prediction, the EPR model provided a non-linear expression [Equation (30)] excluding velocity ratio (m), which generally had simpler mathematical structures than that developed by the GEP model [Equation (21)]. In the GEP models, all the sixth input parameters (see Table 1) were employed to obtain an expression with three genes; the GEP model demonstrated the lowest performance in the calibration stage than other soft computing models. In contrast, the MARS model [Equation (26)] with four inputs and more complex expressions (second polynomial as seen in Table 6) stood at a higher level of accuracy in the calibration stage than M5MT.

Compared to the GEP model [Equation (22)] for the V_L^* prediction, MARS [Equation (25)] and EPR [Equation (31)] models have algebraic mathematics with lower complexity as well as superiority over the GEP performance in the validation stage. In the present research, the MARS model excluded the KC parameter even though experimental investigations carried out by Cheng et al. [1] proved that KC played a key role in the formation of the scour propagation around pipelines exposed to waves. As seen in Table 6, the MARS model created three linear splines (for instance, $BF_1 = \max(0, m\theta_C - 0.03044)$) and one polynomial expression consisting of two splines.

6. Effects of Velocity Ratio on the Scour Propagation Rates

Figure 6 illustrates the variation of V_H^* values versus the velocity ratio m at four levels of e/D values varying from 0.2 to 0.5. In Figure 6a, with the exception of M5MT, all soft computing models provided the same pattern for $e/D = 0.1$. For instance, the EPR model indicated a downward trend, plummeting from $V_H^* = 3.844$ in $m = 0.15$ to $V_H^* = 1.331$ in $m = 0.62$, then, increasing to $V_H^* = 4.531$ in $m = 0.79$. Similar to what is inferred from Figure 6a, this trend has been seen in Figure 6b–d. Generally, for all levels of e/D , M5MT demonstrated that variation of V_H^* against velocity ratio remained slightly steady. From Figs. 6a–d, it can be said that scour propagation rate along the pipeline increased for all values of e/D , thereby, m ratio increased over 0.5. Additionally, for $m = 0.5$ – 0.6 , all the soft computing models except M5MT hit minimum values of V_H^* for all values of e/D .

Table 15 shows the error values of the soft computing models' performance for detecting patterns of scour rate propagation versus velocity ratios. According to Table 15, M5MT had the lowest level of accuracy at three levels of e/D (RMSE = 1.105, 1.229, and 0.835 for $e/D = 0.2, 0.3$, and 0.5 , respectively) in comparison with EPR, GEP, and MARS models.

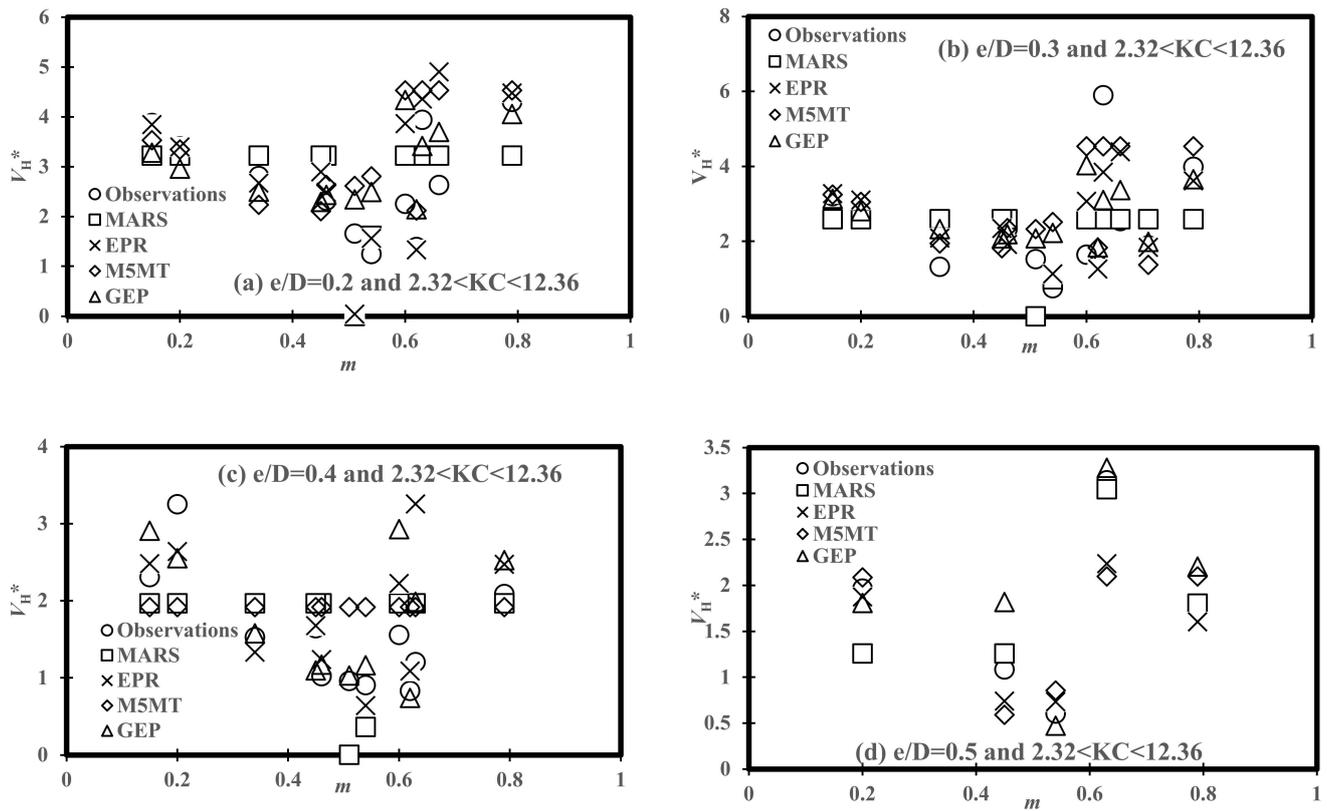


Figure 6. Comparisons of the soft computing models’ performance for variation of V_H^* versus m for various levels of e/D : (a) $e/D = 0.2$, (b) $e/D = 0.3$, (c) $e/D = 0.4$, and (d) $e/D = 0.5$.

Table 15. RMSE values for performance of the soft computing models for variation of velocity ratio.

Soft Computing Models	Variation of V_H^* vs. m for $2.32 < KC < 12.36$			
	$e/D = 0.2$	$e/D = 0.3$	$e/D = 0.4$	$e/D = 0.5$
MARS	0.943	1.251	0.696	0.563
EPR	0.947	1.038	0.812	0.390
M5MT	1.105	1.229	0.741	0.835
GEP	0.902	1.187	0.426	0.803
Soft Computing Models	Variation of V_R^* vs. m for $2.32 < KC < 12.36$			
	$e/D = 0.2$	$e/D = 0.3$	$e/D = 0.4$	$e/D = 0.5$
MARS	1.029	1.141	1.061	0.495
EPR	1.366	1.721	1.326	0.439
M5MT	1.182	1.361	0.862	0.577
GEP	0.842	1.323	1.008	0.439
Soft Computing Models	Variation of V_L^* vs. m for $2.32 < KC < 12.36$			
	$e/D = 0.2$	$e/D = 0.3$	$e/D = 0.4$	$e/D = 0.5$
MARS	0.437	0.671	0.642	0.318
EPR	1.685	1.157	0.623	0.626
M5MT	1.035	0.910	0.792	0.891
GEP	0.805	0.716	0.756	0.386

Figure 7 generally showed a downward trend for the variation of V_R^* against m . As seen in Figure 7a, M5MT and EPR models significantly overpredicted for $e/D = 0.2$ and $m = 0.45$. In Figure 7b, the EPR model was under-predicted for two values of m (0.6 and 0.63) and $e/D = 0.3$, whereas the performance of soft computing models indicated well agreement with experimental observation for $m = 0.15$ and 0.20. Figure 7c illustrates all the soft computing models provided V_R^* values with significant overprediction for $m = 0.6$, whereas moderate overprediction can be seen for $m = 0.79$ and $e/D = 0.4$. Moreover, Figure 7c indicated a permissible harmony between soft computing models and experimental observation at $m = 0.34$ and 0.46. As depicted in Figure 7d, EPR and M5MT had remarkable underprediction in $m = 0.63$ and additionally, MARS indicated significant underprediction in $m = 0.2$. In $m = 0.54$, all the soft computing techniques had permissible performance in the prediction of V_R^* . Table 15 presented RMSE values of predictive models' efficiency at various levels of e/D . EPR indicated the worst performance at three levels of e/D (RMSE = 1.366, 1.721, and 1.326 for $e/D = 0.2, 0.3,$ and $0.4,$ respectively) than MARS, GEP, and M5MT models. Furthermore, M5MT resulted in the lowest accuracy level (RMSE = 0.577) in $e/D = 0.5$ than EPR (RMSE = 0.439), GEP (RMSE = 0.439), and MARS (RMSE = 0.495).

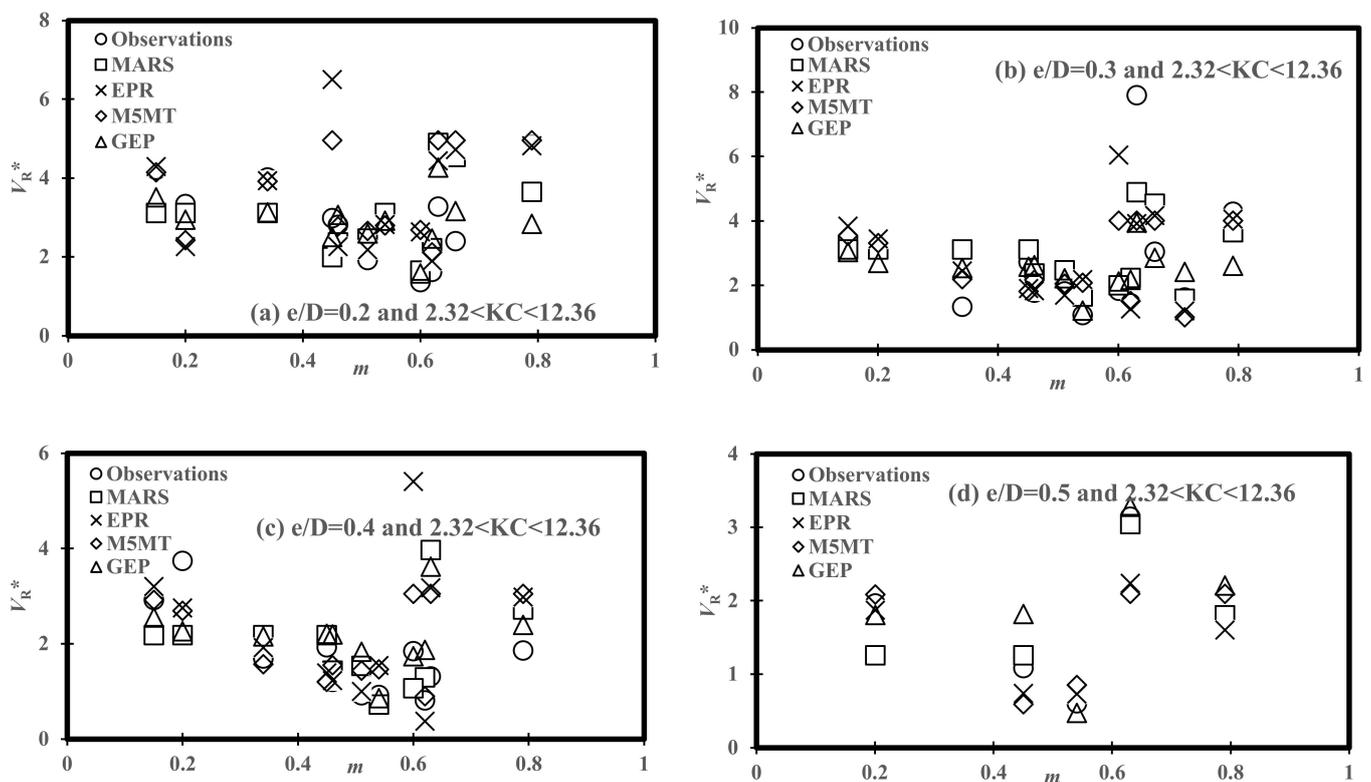


Figure 7. Comparisons of the soft computing models' performance for variation of V_R^* versus m in for various levels of e/D : (a) $e/D = 0.2$, (b) $e/D = 0.3$, (c) $e/D = 0.4$, and (d) $e/D = 0.5$.

Figure 8 illustrates the general pattern of V_L^* values against m for all levels of e/D . For instance, in $e/D = 0.2$, Figure 8a indicated that V_L^* values first have a downward trend, decreasing from 2.863 in $m = 0.15$ to 0.810 in $m = 0.6$, then, increasing to 3.158 in $m = 0.79$. In Figure 8a, the EPR model significantly overpredicted for two values of m (0.63 and 0.79). M5MT had overprediction of V_L^* in $m = 0.79$, whereas the MARS and GEP models performed well in agreement with experimental observations. Additionally, Figure 8a illustrated all the soft computing models had a promising performance for $m = 0.15$. Figure 8b demonstrated that EPR had significant overpredictions of V_L^* in $m = 0.66$ whereas all the soft computing models had moderate overestimation in $m = 0.6$. Figure 8c depicted that M5MT had remarkable overprediction in $m = 0.51$ and 0.54 whereas, for $m = 0.2$, all the pre-

dictive models underpredicted values of V_L^* . More specifically, experimental observations simulated by M5MT have no certain patterns against m variations. In Figure 8c, MARS, GEP, and EPR models significantly overpredicted in $m = 0.6$ and 0.63 . Figure 8d indicated that all the soft computing techniques had underprediction in $m = 0.63$. In $m = 0.45$, EPR and MARS models overpredicted V_L^* values, whereas M5MT and GEP models performed well and were in good agreement with experimental observations. According to the statistical results of Table 15, EPR model showed the best performance at e/D levels of 0.2 (RMSE = 0.437), 0.3 (RMSE = 0.671), 0.4 (RMSE = 0.623), and 0.5 (RMSE = 0.318).

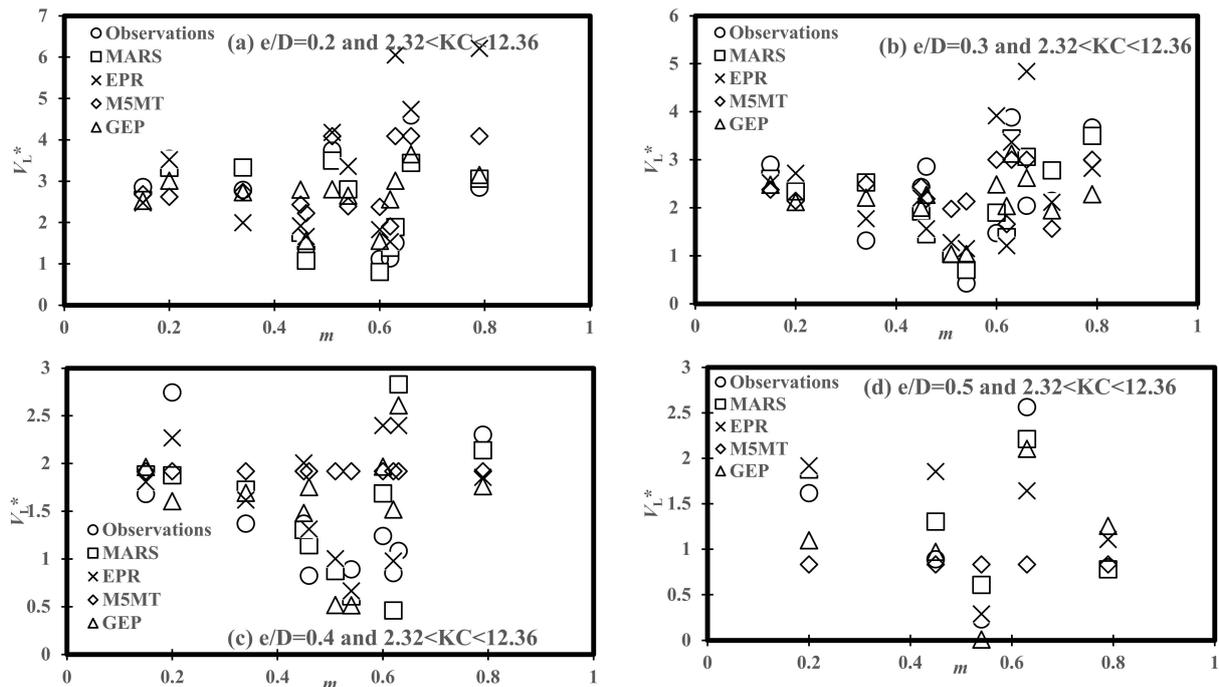


Figure 8. Comparisons of the soft computing models’ performance for variation of V_L^* versus m in for various levels of e/D : (a) $e/D = 0.2$, (b) $e/D = 0.3$, (c) $e/D = 0.4$, and (d) $e/D = 0.5$.

7. Conclusions

The present research evaluated 3D scour propagation rates around seafloor offshore pipelines under simultaneous effects of current and wave conditions. Four AI models were used to assess the scour propagation rates. In this way, the 42 experimental works by Cheng et al. [1] were collected from the live-bed scouring tests in the currents and waves circumstances. To feed the AI models, the Buckingham theorem was applied to provide the key dimensionless parameters, namely θ_C , θ_W , e/D , KC , and m . From these parameters, seven sets of non-dimensional parameters were provided to have a reasonable assessment of the scour propagation rates for the currents and waves at the same time. Overall, the following findings can be expressed as:

- The developments of new equations by seven combinations of dimensionless parameters showed that the present predictive techniques raised three chief merits: (i) providing mathematical expressions by EPR and GEP with complicated terms (as naturally found in the scour propagation around offshore pipelines) when fed by a limited number of scouring tests, (ii) on the contrary, equations by M5MT and MARS models reduced the complexity of the evaluation of scouring process by providing simpler regression equations, and (iii) selecting optimal combination of the effective dimensionless parameters (from θ_C , θ_W , e/D , KC , and m) played a key role in the prediction of the scour propagation rates;
- From the calibration and validation phases, the performance of AI models indicated a reasonable degree of efficiency for the estimation of the scour propagation rates.

More specifically, statistical measures demonstrated that Equations (20) and (21) given by the GEP model had the best performance in the estimation of V_R^* and V_H^* , respectively, whereas the MARS model [Equation (25)] indicated the most accurate efficiency for the evaluation of V_L^* . On the other hand, the sixth combination of effective parameters (i.e., $m \theta_C, (1 - m)\theta_W, 1 - e/D, KC$) provided the best results for the scour propagation rates in the right and left hands of pipelines (V_R^* and V_L^*) when the GEP and MARS models fed, respectively. In the case of V_H^* , the MARS model fed by the fifth combination of dimensionless parameters (i.e., $F, \theta_C, \theta_W, m, e/D$) demonstrated the best performance. Generally, it was inferred that the performance of AI models stood at the highest level of accuracy when θ_C , θ_W , and m were not converted to one dimensionless parameter [$m\theta_C + (1 - m)\theta_W$].

- The physical consistency of the predictive models' results has been studied by analyzing scour propagation rates versus the ratio of pipeline embedment depth and pipeline diameter (e/D), the ratio of current velocity to orbital velocity (m), and Keulegan–Carpenter number (KC). From variations of the scour propagation patterns versus m , it was found that scour propagation rates followed a decreasing trend for $m = 0.2$ – 0.6 and all ranges of e/D and KC , then; increased up to $m = 0.8$.
- The effects of complexity level on the performance of AI models for 3D scour propagation rates was investigated. The EPR model was developed by Set 6 for the V_H^* prediction with complex expressions in comparison with M5MT and Equation (24) (MARS). Multivariate linear equations by M5MT were simpler than those obtained expressions by EPR and GEP models. Additionally, the EPR model provided V_R^* prediction with simpler mathematical structures than that developed by the GEP model [Equation (21)]. The sixth combination of input parameters has been applied to acquire a mathematical expression-based GEP model with three genes, indicating the lowest level of efficiency in the calibration stage in comparison with AI techniques. On the contrary, the MARS model with four dimensionless inputs (i.e., $m \theta_C, (1 - m)\theta_W, 1 - e/D, KC$) and the second-order polynomial had better performance in the calibration phase than linear equations by M5MT. Furthermore, explicit equations given by MARS and EPR models have lower complexity for the V_L^* estimation in comparison with the GEP model.

Author Contributions: Conceptualization, M.N., G.O. and F.S.-M.; methodology, M.N.; software, M.N.; validation, M.N. and G.O.; formal analysis, M.N. and G.O.; investigation, M.N. and G.O.; data curation, M.N. and G.O.; writing—original draft preparation, M.N., G.O. and F.S.-M.; writing—review and editing, M.N., G.O. and F.S.-M.; visualization, M.N., G.O. and F.S.-M.; supervision, M.N., G.O. and F.S.-M. All authors have read and agreed to the published version of the manuscript.

Funding: This research has been supported by Graduate University of Advanced Technology (Kerman-Iran) under grant number of 2594.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: This research has been supported by the Graduate University of Advanced Technology (Kerman-Iran) under grant number 2594.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cheng, L.; Yeow, K.; Zang, Z.; Li, F. 3D scour below pipelines under waves and combined waves and currents. *Coast. Eng.* **2014**, *83*, 137–149. [[CrossRef](#)]
2. Cheng, L.; Yeow, K.; Zhang, Z.; Teng, B. Three-dimensional scour below offshore pipelines in steady currents. *Coast. Eng.* **2009**, *56*, 577–590. [[CrossRef](#)]
3. Adegboye, M.A.; Fung, W.-K.; Karnik, A. Recent Advances in Pipeline Monitoring and Oil Leakage Detection Technologies: Principles and Approaches. *Sensors* **2019**, *19*, 2548. [[CrossRef](#)] [[PubMed](#)]

4. Sumer, B.M.; Fredsøe, J. *The Mechanics of Scour in the Marine Environment*; World Scientific: Singapore, 2002.
5. Sumer, B.M.; Fredsøe, J. Scour below pipelines in waves. *J. Waterway Port Coast. Ocean Eng.* **1990**, *116*, 307–323. [[CrossRef](#)]
6. Sumer, B.M.; Truelsen, C.; Sichmann, T.; Fredsøe, J. Onset of scour below pipelines and self-burial. *Coast. Eng.* **2001**, *42*, 313–335. [[CrossRef](#)]
7. Bernetti, R.; Bruschi, R.; Valentini, V.; Venturi, M. Pipelines placed on erodible seabeds. In Proceedings of the 9th International Conference on Offshore Mechanics and Arctic Engineering, ASME, Houston, TX, USA, 18–23 February 1990; Volume 5, pp. 155–164.
8. Lucassen, R.J. Scour Underneath Submarine Pipelines. In *MAT Rep. PL-42A, Marine Tech. Res.*; TU Delft: Delft, The Netherlands, 1984.
9. Sumer, B.M.; Fredsøe, J. Scour around pipelines in combined waves and current. In Proceedings of the 15th Conference on Offshore Mechanics Arctic Engineering, Florence, Italy, 16–20 June 1996; American Society of Mechanical Engineers: New York, NY, USA, 1996; Volume 5, pp. 595–602.
10. Fredsøe, J.; Sumer, B.M.; Arnskov, M.M. Time scale for wave/current scour below pipelines. *Int. J. Offshore Polar Eng.* **1992**, *2*, 13–17.
11. Wu, Y.; Chiew, Y.-M. Three-dimensional scour at submarine pipelines. *J. Hydraul. Eng.* **2012**, *138*, 788–795. [[CrossRef](#)]
12. Wu, Y.; Chiew, Y.-M. Mechanics of three-dimensional pipeline scour in unidirectional steady current. *J. Pipeline Syst. Eng.* **2013**, *4*, 3–10. [[CrossRef](#)]
13. Najafzadeh, M.; Saberi-Movahed, F. GMDH-GEP to predict free span expansion rates below pipelines under waves. *Mar. Georesources Geotechnol.* **2019**, *37*, 375–392. [[CrossRef](#)]
14. Ehteram, M.; Ahmed, A.N.; Ling, L.; Fai, C.M.; Latif, S.D.; Afan, H.A.; Banadkooki, F.B.; El Shafie, A. Pipeline scour rates prediction-based model utilizing a Multilayer Perceptron-Colliding Body algorithm. *Water* **2020**, *12*, 902. [[CrossRef](#)]
15. Najafzadeh, M.; Oliveto, G. Exploring 3D wave-induced scouring patterns around subsea pipelines with Artificial Intelligence techniques. *Appl. Sci.* **2021**, *11*, 3792. [[CrossRef](#)]
16. Najafzadeh, M.; Oliveto, G. Scour Propagation Rates around Offshore Pipelines Exposed to Currents by Applying Data-Driven Models. *Water* **2022**, *14*, 493. [[CrossRef](#)]
17. Hansen, E.A.; Staub, C.; Fredsøe, J.; Sumer, B.M. Time-development of scour induced free spans of pipelines. In Proceedings of the 10th Conference on Offshore Mechanics and Arctic Engineering (OMAE), Stavanger, Norway, 23–28 June 1991.
18. Soulsby, R. *Dynamics of Marine Sands: A Manual for Practical Applications*; Thomas Telford: London, UK, 1997.
19. Ferreira, C. Gene expression programming: A new adaptive algorithm for solving problems. *Complex Syst.* **2001**, *13*, 87–129.
20. Ferreira, C. *Gene Expression Programming*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2006.
21. Friedman, J.H. Multivariate adaptive regression splines. *Ann. Statist.* **1991**, *19*, 1–67. [[CrossRef](#)]
22. Mokhtia, M.; Eftekhari, M.; Saberi-Movahed, F. Feature selection based on regularization of sparsity based regression models by hesitant fuzzy correlation. *Appl. Soft Comput.* **2020**, *91*, 106255. [[CrossRef](#)]
23. Mokhtia, M.; Eftekhari, M.; Saberi-Movahed, F. Dual-manifold regularized regression models for feature selection based on hesitant fuzzy correlation. *Knowl.-Based Syst.* **2021**, *229*, 107308. [[CrossRef](#)]
24. Quinlan, J.R. Learning with Continuous Classes. In *Proceedings of the 5th Australian Joint Conference on Artificial Intelligence (AI '92)*; World Scientific: Singapore, 1992; pp. 343–348.
25. Granata, F.; de Marinis, G. Machine learning methods for wastewater hydraulics. *Flow Meas. Instrum.* **2017**, *57*, 1–9. [[CrossRef](#)]
26. Granata, F.; Saroli, M.; de Marinis, G.; Gargano, R. Machine learning models for spring discharge forecasting. *Geofluids* **2018**, *2018*, 8328167. [[CrossRef](#)]
27. Solomatine, D.P.; Xue, Y. M5 Model Trees and Neural Networks: Application to Flood Forecasting in the Upper Reach of the Huai River in China. *J. Hydrol. Eng.* **2004**, *9*, 491–501. [[CrossRef](#)]
28. Di Nunno, F.; Alves Pereira, F.; de Marinis, G.; Di Felice, F.; Gargano, R.; Miozzi, M.; Granata, F. Deformation of Air Bubbles Near a Plunging Jet Using a Machine Learning Approach. *Appl. Sci.* **2020**, *10*, 3879. [[CrossRef](#)]
29. Granata, F.; Di Nunno, F. Air Entrainment in Drop Shafts: A Novel Approach Based on Machine Learning Algorithms and Hybrid Models. *Fluids* **2022**, *7*, 20. [[CrossRef](#)]
30. Etemad-Shahidi, A.; Yasa, R.; Kazeminezhad, M.H. Prediction of wave-induced scour depth under submarine pipelines using machine learning approach. *Appl. Ocean Res.* **2011**, *33*, 54–59. [[CrossRef](#)]
31. Giustolisi, O.; Doglioni, A.; Savic, D.A.; Webb, B.W. A multi-model approach to analysis of environmental analysis. *Environ. Modell. Softw.* **2007**, *22*, 674–682. [[CrossRef](#)]
32. Giustolisi, O.; Savic, D.A. Advances in data-driven analyses and modelling using EPR-MOGA. *J. Hydroinf.* **2009**, *11*, 225–236. [[CrossRef](#)]
33. Laucelli, D.; Giustolisi, O. Scour depth modeling by a multi-objective evolutionary paradigm. *Environ. Model. Softw.* **2011**, *26*, 498–509. [[CrossRef](#)]
34. Savic, D.A.; Giustolisi, O.; Berardi, L.; Shepherd, W.; Djordjevic, S.; Saul, A. Modelling sewer failure by evolutionary computing. *Water Manag.* **2006**, *159*, 111–118. [[CrossRef](#)]
35. Berardi, L.; Giustolisi, O.; Kapelan, Z.; Savic, D.A. Development of pipe deterioration models for water distribution systems using EPR. *J. Hydroinf.* **2008**, *10*, 113–126. [[CrossRef](#)]