https://doi.org/10.1016/j.engfailanal.2020.104505

PREDICTION OF MAXIMUM PITTING CORROSION DEPTH IN OIL AND GAS PIPELINES

Mohamed El Amine Ben Seghier [1,2,*], Behrooz Keshtegar[3], Kong Fah Tee[4], Tarek Zayed[5], Rouzbeh Abbassi[6], Nguyen Thoi Trung[1,2]

[1] Division of Computational Mathematics and Engineering, Institute for Computational Science, Ton Duc Thang University, Ho Chi Minh City, Vietnam
[2]Faculty of Civil Engineering, Ton Duc Thang University, Ho Chi Minh City, Vietnam
[3]Department of Civil Engineering, Faculty of Engineering, University of Zabol, P.B. 9861335856, Zabol, Iran
[3]School of Engineering, University of Greenwich, Kent, UK
[4]Building and Real Estate Dept., The Hong Kong Polytechnic University, Hong Kong.
[5] School of Engineering, Faculty of Science and Engineering, Macquarie University, Sydney, NSW, Australia

* Corresponding author: benseghier@tdtu.edu.vn

ABSTRACT

Avoiding failures of corroded steel structures are critical in offshore oil and gas operations. An accurate prediction of maximum depth of pitting corrosion in oil and gas pipelines has significance importance, not only to prevent potential accidents in future but also to reduce the economic charges to both industry and owners. In the present paper, efficient hybrid intelligent model based on the feasibility of Support Vector Regression (SVR) has been developed to predict the maximum depth of pitting corrosion in oil and gas pipelines, whereas the performance of well-known meta-heuristic optimization techniques, such as Genetic Algorithm (GA), Particle Swarm Optimization (PSO) and Firefly Algorithm (FFA), are considered to select optimal SVR hyper-parameters. These nature-inspired algorithms are capable of presenting precise optimal predictions and therefore, hybrid models are developed to integrate SVR with GA, PSO, and FFA techniques. The performances of the proposed models are compared with the traditional SVR model where its hyper-parameters are attained through trial and error process on the one hand and empirical models on the other. The developed models have been applied to a large database of maximum pitting corrosion depth. Computational results indicate that hybrid SVR models are efficient tools, which are capable of conducting a more precise prediction of maximum pitting corrosion depth. Moreover, the results revealed that the SVR-FFA model outperformed all other models considered in this study. The developed SVR-FFA model could be adopted to support pipeline operators in the maintenance decision-making process of oil and gas facilities.

Keywords: Hybrid intelligent models; Support vector regression; Firefly algorithm; Pitting corrosion; Oil and gas pipelines.

1. INTRODUCTION

For decades, pipelines have been used all over the world as the fastest and safest way of oil and gas transportation. Any failure in pipeline transmission systems has a direct impact on the economics of the oil and gas industry. Various failure modes may affect transmission systems [1]. For several years, researchers have studied these failure modes [2,3]. Most of the investigations on the evaluating of the different failure modes for oil and gas pipelines reveal that the corrosion is one of the most common causes of failures in transmission systems.

Analysis of existing reports on pipeline failure incidents reveals that corrosion, as the highest mode of failure, has the most negative impact in the degradation of oil and gas pipelines [4,5]. Lam and Zhou [6] analyzed the statistical failure incidents of onshore gas pipelines using the well-known PHASMA database from 2002 to 2013. The results confirm that 23.7% of pipeline failure incidents may occur from external corrosion while 8.7% from internal corrosion. Valor et al. [7,8] explained that 60% of failures in Mexican oil and gas transportation systems are caused by pitting corrosion on the external walls of pipelines. This fact of high incidents occur due to corrosion relates to the complexity of the environment that surrounds pipelines, including a large variety of the soil properties, water and transported products using the pipeline (Oil or Gas). All these features provoke the development and the growth of numerous types of corrosion on the metallic surfaces of a pipeline. Among them, uniform and localized corrosion are the most common types of pipelines failures, known as pitting corrosion [9–11]. Therefore, to predict the failure of a pipeline, and schedule the maintenance accordingly, it is important to estimate the maximum depth

of the pitting corrosion precisely. Moreover, a precise prediction of pits depth allows operators to better analyze and manage the risk in the transmission pipeline system [5,12].

Several researchers have proposed different models to predict the maximum depth of pitting corrosion in oil and gas pipelines by attending to its importance in the safety assessment and the structural reliability analysis. Various theoretical basis methods have been adopted analytically or by using real field-databases integrated with statistical approaches, which result in the development of different empirical to fully stochastic models.

Most of the existing empirical models predicting pitting corrosion depth in metals are developed based on the power-law model. The model was developed based on the studies of Rossum in 1969 [13], where it can be expressed using Equation 1 in which parameters *k* and *n* are constants, d represents the maximum depth while T is the time in years.

$$d(T) = kT^n \qquad (1)$$

In the past few decades, several attempts were made to enhance and improve the power-law model by developing more precise approaches for formulating *k* and *n*. Among the widely used models based on Equation 1 to describe pitting depths in pipelines, the model proposed by Velázquez et al. (2009) [14] is one of the most well-known. This empirical model is a modified version of Equation 1 in which a new factor is considered in the model to represent pit initiation time (T_0). In addition, to developing this model, a large database of real-field corroded pipelines buried in various soil types was considered. The collected database includes several important parameters from the maximum depths of pitting corrosion to different soil properties in each location. Statistical analysis was then conducted to fit the database into the proposed model using a non-linear regression technique. The two contacts in the model (*K* and *n)* were derived based on all possible combinations of factors influencing pitting corrosion. Although the new model has

improved the basic power-law model proposed by Rossum, the discussion of the model's performance evaluation was limited to the Coefficient of determination (R^2), where the highest results were equal to 0.9. Moreover, the applied technique based on non-linear regression suffers from the chaotic pattern of the data, and mostly results in the inaccurate fitting. Similarly, Alamilla and Sosa (2009) [15] proposed a new model that describes the damage velocity of pitting corrosion in operating buried steel pipelines. Unlike the Velázquez model, this model introduces two types of pit rates which are initial and long-term rates, and it is only the second one that is assumed to be dependent on soil properties. Several in-line inspection results were used to support model development. Moreover, the proposed model was developed mathematically using different known operations such as Tylor series and convolution equation, yet the performance evaluation and the ability of the model were not investigated.

In the stochastic models, Markovian approaches are widely applied to describe the growth of pitting defects in oil and gas pipelines [16]. Several researchers have been conducted using Markov chain methodologies for predicting the pit development. Transition probabilities between pit ‘‘states’’ are modeled as functions of environmental factors following the original model of pit growth rate developed by Velázquez et al. (2009) [17–19]. Previous studies on predicting the maximum depth of pitting corrosion have mostly focused on using simple and analytical methods to achieve their formulation. No previous study has given sufficient consideration to evaluate the performance of the model that predicts the maximum depth of pitting corrosion or employs hybrid artificial intelligence (hybrid AI) approaches to solve such a problem.

Recently, Artificial Intelligence (AI) has attracted much attention due to the success of its techniques in solving engineering problems [20,21] . Support Vector Regression (SVR) is among these techniques that have been widely used as a strong tool to formulate the relationship between

inputs and outputs variables [22,23]. The SVR has been applied in various fields including regression estimation problems, time-series predictions, pattern recognition and dynamic systems inverse solution of [24–27]. Despite the advantages of SVR techniques in problem-solving, SVR suffers from the un-optimum hyper-parameters selection. This returns to the fact that SVR traditionally chooses its hyper-parameters using trial and error approach. This method is time-consuming and leads to un-optimum hyper-parameters, which decrease the performance of the SVR technique. In order to overcome such limitations, integrating an optimization method with the SVR technique has been proposed as a solution to achieve optimum selections of the hyper-parameters. Wen et al [28] developed an SVR model in order to predict the corrosion rate within 3C steel. Five different seawater environments were studied. Their research reveals that the SVR results show more precious prediction than the back-propagation neural network (BPNN) results. However, sample numbers of the reported database is very limited, which was insufficient to train the predictive models in terms of performance and accuracy.

Thus, considering the aforementioned shortcomings, the present research attempts to develop a model for predicting the maximum depth of pitting corrosion in oil and gas pipelines using hybrid artificial intelligence approaches. Therefore, different hybrid meta-heuristic optimization algorithms i.e., Practical Swarm Optimization (PSO), Genetic Algorithms (GA) and Firefly Algorithms (FFA), are integrated with SVR to select its optimal hyper-parameters. The hybrid models i.e., SVR-GA, SVR-PSO, and SVR-FFA are then used to predict the maximum pitting corrosion in pipelines. To achieve this aim, a large experimental database containing the maximum depths of pitting corrosion with different related soil properties is used. Therefore, the objectives of this study are to (1) develop the framework of the proposed hybrid models-based SVR. (2)

Implement both the proposed models i.e. SVR, SVR-GA, SVR-PSO, SVR-FFA and excited well-known empirical correlations on a large database of maximum pitting corrosion depth. (3) Evaluate and compare the prediction results in terms of performance, efficiency, and accuracy using different statistical and graphical criteria. This study outlines a new framework for more accurately predicting the maximum depth of pitting corrosion in oil and gas pipelines. The application of such a model enhances operational safety and reliability in the oil and gas industry.

2. MODELLING METHODOLOGY

2.1. Support Vector Regression (SVR)

In the 1960s, Vapnik introduced a new tool to solve classification problems based on the concepts of Structural Risk Minimization (SRM) and Support Vector Machine (SVM) [29,30]. Subsequently, at the end of the 20th century, the general applicability of the method had been improved by adopting the principle of the ε-insensitive loss function [31]. This amelioration has permitted SVM to solve nonlinear regression estimation problems. Inspiration from SVM, by including the new loss function to solve non-linear problems, has promoted Support Vector Regression. SVR as a non-linear kernel method used for regression problems is a powerful machine learning technique. The goal of the method is to define the best hyper-plan for regression, aiming to minimize as much as possible the risk for high dimensional feature space [32,33].

SVR concept is to split the training data $\{x_i, y_i\}$ in order to determine the hyper-planes with the maximum margin. A hyper-plane is defined by using Equation 5 [34,35]:

$$f(x) = \langle w, x \rangle + b \tag{5}$$

where w represents the regression coefficient vector, while b represents the bias. Therefore, minimizing the regularized risk problem represented in Equations 6 and 7 is a solution for the

regression. Thus, the estimation of w and b can be conducted based on minimizing the next optimization formulation:

$$R(f) = \frac{1}{n}\sum_{i=1}^{n} L(f(x_i) - y_i) + \frac{1}{2}\|w\|^2 \tag{6}$$

$$L(f(x) - y) = \begin{cases} \|f(x) - y\| - \varepsilon & for\ |f(x) - y| \geq \varepsilon \\ 0 & otherewise \end{cases} \tag{7}$$

Equation 6 represents the so-called ε -sensitive loss function, whereas, ε defines the precision parameter that represents the tube radius of this function. This tube surrounds the regression function $f(x)$. This zone is called the ε-sensitive zone that represents the internal tube region, where the predicted value loss in this zone is equal to zero. Outside this region, the predicted value of the loss is equal to the difference magnitude between the predicted value and the radius ε.

The optimum parameters are obtained using Equations 8 and 9, which formulate a constrained optimization problem.

$$Minimize\ \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi + \xi_i^*) \tag{8}$$

Subjected to constraints:

$$Subjected\ to \begin{cases} y_i - \langle w, x_i\rangle - b \leq \varepsilon + \xi_i \\ \langle w, x_i\rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \tag{9}$$

C represents a modifying constant-coefficient superior to zero, where the trade-off between model complexity and training error is determined using this constant. In Equation 8, the term $\frac{1}{2}\|w\|^2$, improves the generalization ability by maximization of the regression function smoothness due to the concept, distance increasing of two separated training data [34]. The other term uses ε -sensitive loss function to penalize training errors of $f(x)$ [36]. Therefore, optimum parameters

values can be found using the Lagrange equation based on the width of the tube (ε), the constant *C* and the kernel function *K* [37]. Hence, SVR can be represented using Equations10 and 11.

$$maximize -\frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i-\alpha_i^*)(\alpha_j-\alpha_j^*)\left(x_i,x_j\right)-\varepsilon\sum_{i=1}^{n}(\alpha_i-\alpha_i^*)+\sum_{i=1}^{n}y_i(\alpha_i-\alpha_i^*) \tag{10}$$

with the following constraints:

$$S.t\begin{cases}\sum_{i=1}^{n}(\alpha_i-\alpha_i^*)=0\\ 0\le\alpha_i\le C\\ 0\le\alpha_i^*\le C\\ i,j=1,2,\dots,N\end{cases} \tag{11}$$

α_i, α_i^* are Lagrange multipliers and the solution for the dual problem, they have non-zero values. Using the abovementioned maximization function, SVR for the fitting function is illustrated by adopting Equation 12.

$$f(x,w)=(\alpha_i-\alpha_i^*)\langle x_i,x\rangle+b \tag{12}$$

Finally, in the dual space, SVR function is expressed by using Equation 13.

$$f(x,w)=(\alpha_i-\alpha_i^*)K(x_i,x)+b \tag{13}$$

For obtaining the kernel function *K*, several equations are available (i.e. linear kernel, radial basis kernel, polynomial kernel, sigmoid kernel, … etc.). In the current study, the radial basis function is used as it commonly gives the best outcome and is well adaption with datasets. Therefore, the radial basis function is expressed by Equation 14 [38].

$$\mathrm{K}\left(x_i,x_j\right)=\exp\left(\frac{1}{2}\frac{\left\|x_i-x_j\right\|^2}{\sigma^2}\right) \tag{14}$$

The addressed objective function for the optimization approach through this study is the average absolute relative deviation (AARD) of SVR technique outcomes. Thus, the AAPD of SVR outcomes is formulated as explained by Equation 15.

$$AAPD = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{d_{max}^{i} - \hat{d}_{max}^{i}}{d_{max}^{i}}\right| \quad (15)$$

In which d_{Max}^{i}, $\hat{d}_{Max}^{i}$ are the i[th] field measured and predicted value of maximum depth of pitting corrosion, correspondingly, and *n* represents the initial population samples. In other words, the optimization problem can be expressed as illustrated by Equation 16.

$$Min\ F(C, \varepsilon, \gamma) = Min(AAPRD) \quad (16)$$

2.2. SVR hyper-parameters optimization using Meta-heuristic algorithms

2.2.1. Particle Swarm Optimization (PSO)

Particle swarm optimization (PSO) is a powerful meta-heuristic algorithm developed from the inspiration of swarms dynamic movement and social behaviour such as birds, fishes and insects [39]. PSO stands for a stochastic approach based on the concept of population-based which aims to solve space problems by a continuous search [40]. In this study, PSO is considered to determine SVR optimum hyper-parameters, which are referred to as SVR-PSO. Hybrid SVR-PSO steps begin with random initializations of the locations and velocities population.

Each particle fitness is evaluated using the statistical function in Equation 15 based on train database. The next step is the stopping criteria test. If the best particle fitness rate satisfies the condition, the parameters are supposed as being optimum. Otherwise, an enhanced rate must be achieved using particle velocity and position updating.

Therefore, in the case where the fitness result of the global best is inferior to the fitness result of the particle, the parameters of the best global should be updated. In the other case, where best particle fitness is less than best global fitness, updating goes for the best particle. The new velocity and position can be formulated by adopting Equations 17 and 18 respectively.

$$V_{i+1} = \omega.V_i + C_1.rand(\quad).(P_{i\,best} - X_i) + C_2.rand(\quad).(g_{i\,best} - X_i) \quad (17)$$

$$X_{i+1} = X_i + V_{i+1} \quad (18)$$

Where $V_{i+1}, V_i, X_i, X_{i+1}$ are the velocity and position at ith and i+1th moment respectively, rand () represent random value in [0,1] range; C_1 and C_2 are called learning factors, and ω is a weighting factor used to accelerate the convergence speed. The mathematical formulation of ω is provided by Equation 19.

$$\omega = \omega_{min} + (iter_{max} - iter).(\omega_{max} - \omega_{min})/iter_{max} \quad (19)$$

Terms $\omega_{min}, \omega_{max}$ represent the smallest and highest weighting factors, where "Iter" refers to the number of iterations. Therefore, the next step is to re-evaluate the following particles.

2.2.2. Genetic Algorithm (GA)

Genetic algorithm (GA) represents a well-known randomized searching algorithm for handling optimization problems. Its principals are inspired by the theory of genetics and natural selection [41]. GA is selected to search the optimum hyper-parameters of SVR due to the algorithm efficiency in solving various engineering optimization problems [41]. The involved steps of the hybrid SVR-GA start by initializing a random population. Next, using the statistical fitness function, each individual is evaluated. Subsequently, an examination of the agreement of each individual, based on the stopping criterion, is made. If results are found to be acceptable, the algorithm must stop and the global best satisfactory individual has been generated. Therefore, the optimum parameters are achieved. If results are not acceptable, fitness evaluation is the next step to be considered. It includes three essential operations starting with the selection of the fittest individuals, then, the crossover is followed by the mutation in order to overcome the stopping criterion [42].

2.2.3. FireFly Algorithm (FFA)

Firefly algorithm (FFA) is a new swarm intelligence optimization technique developed by Yang [43], inspired by fireflies dynamic movements. The algorithm has obtained considerable attention in the past few years due to its efficiency to determine the optimal solutions for numerous optimization problems [44–46]. FFA is deployed in this study to obtain the optimal hyper-parameters of SVR. Hybrid SVR-FFA steps are commenced by randomly generating the population and FFA parameters i.e., original light intensity (β_0) adsorption coefficient (γ) and attractiveness. By using the statistical fitness function, fireflies brightness is evaluated. If the brightness of the firefly fulfills the stopping criteria, then the algorithm must cut and the supposed optimal parameters have been generated. Otherwise, another two-steps must be considered, first fireflies must move to the brightest firefly, and then the fireflies brightness should be updated. The movement of an "*i*" firefly towards another brighter "*j*" firefly is expressed by adopting Equation 20.

$$X_i = X_i + \beta_0 e^{-\gamma r_{i,j}^2}\left(X_i - X_j\right) + \alpha(rand(\quad) - \frac{1}{2}) \tag{20}$$

Where the attractiveness is represented by the term $\beta_0 e^{-\gamma r_{i,j}^2}$; $(rand(\quad) - \frac{1}{2})$ is a term indicating a randomization in the range [-0.5,0.5], β_0 is the light intensity and γ refers to the adsorption coefficient. Thus, a distance between two fireflies *i* and *j* at X_i and X_j, respectively, is a Cartesian distance formulated by using Equation 21.

$$r_{i,j} = \sqrt{\sum_{k=1}^{D}\left(X_{i,k} - X_{j,k}\right)} \tag{21}$$

The second step is to rank the fireflies according to their fitness and determine the best option. The framework of hybrid SVR approaches i.e. SVR-PSO, SVR-GA, SVR-FFA are depicted in Figure1 where the initial random populations are divided into training and test datasets. The optimization

algorithms i.e. PSO, GA, FFA are applied on the training dataset then the performance and accuracy of the output results generalization in each algorithm are tested using the other dataset as shown in Figure 1.

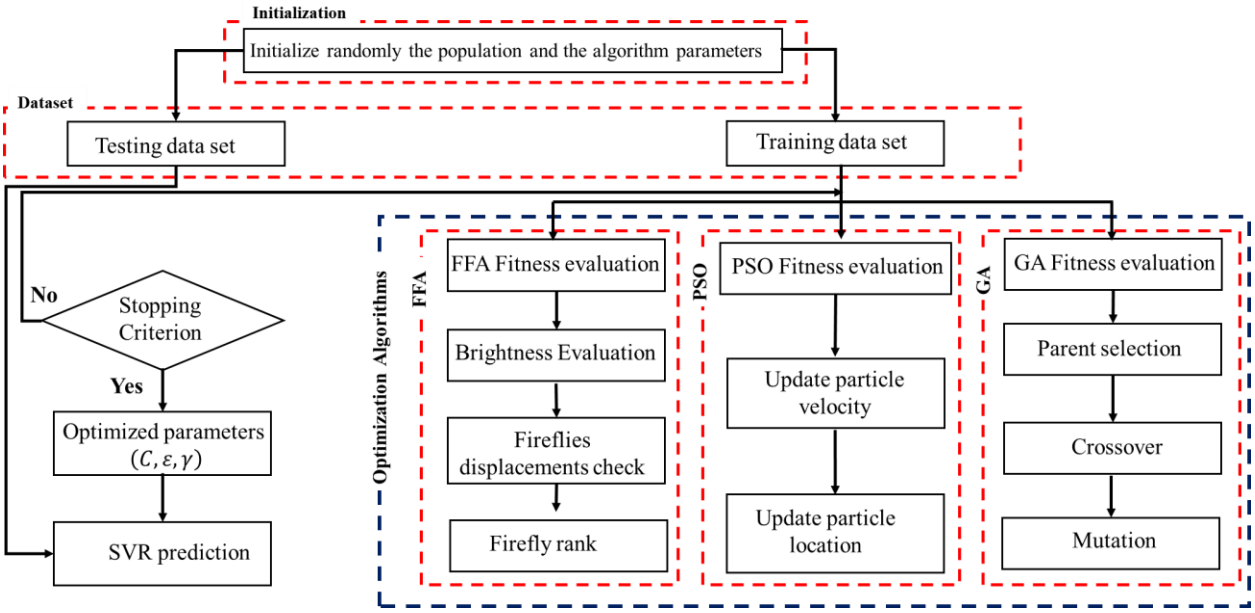


Fig. 1. Flowchart of the proposed hybrid models (i.e. SVR-GA, SVR-PSO and SVR-FFA)

3. Comparative criteria

The performance and accuracy of the proposed intelligent techniques i.e., SVR, SVR-GA, SVR-PSO, SVR-FFA and empirical correlations against the studied database is demonstrated through different comparative criteria. Thus, the differences between field measured data and predicted data using the above models are determined. Subsequently, a comparative study is considered according to several statistical indicators. These indicators include the root mean square error (*RMSE*), the mean absolute error (*MAE*), the Nash-Sutcliffe Efficiency (*NSE*), the Willmott index of agreement (*d*), and the coefficient of determination (R^2). Models pertaining to these criteria are explained in Equations 22-28 [47].

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(d_{max}^{i} - \hat{d}_{max}^{i}\right)^2} \quad (22)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|d_{max}^{i} - \hat{d}_{max}^{i}\right| \quad (23)$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(d_{max}^{i} - \hat{d}_{max}^{i}\right)^2 \quad (24)$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(d_{max}^{i} - \hat{d}_{max}^{i}\right)^2}{\sum_{i=1}^{n}\left(d_{max}^{i} - \bar{d}_{max}^{i}\right)^2} \qquad -\infty \le NSE \le 1 \quad (25)$$

$$d = 1 - \frac{\sum_{i=1}^{n}\left(d_{max}^{i} - \hat{d}_{max}^{i}\right)^2}{\sum_{i=1}^{n}\left(\left|\hat{d}_{max}^{i} - \bar{d}_{max}^{i}\right| + \left|d_{max}^{i} - \bar{d}_{max}^{i}\right|\right)^2} \qquad 0 \le d \le 1 \quad (26)$$

$$CI = d \times NSE \quad (27)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(d_{max}^{i} - \hat{d}_{max}^{i}\right)^2}{\sum_{i=1}^{n}\left(\bar{d}_{max}^{i} - \hat{d}_{max}^{i}\right)^2} \quad (28)$$

In Equations, 22-28, *n* refers to the number of database samples, and d_{max}, $\hat{d}_{max}$, $\bar{d}_{max}$ are the measured field data, predicted and average values of maximum pitting corrosion depth respectively. The models with the smallest values of RMSE and MAE are considered to have more accurate performance, while models with higher R^2 and CI values (i.e. near to unit) show more robust and efficient performance.

4. DEVELOPED MODELS IMPLEMENTATION

4.1. Data Collection

It is of fundamental importance that for reliable and accurate AI models development, a reliable and accurate database should be used. Therefore, in order to implement the above modeling approaches and to prove their efficiency and accuracy against a large database, maximum depth data of pitting corrosion in oil and gas pipelines are collected from previous literature [48]. The

data contains 259 maximum pit depths gathered from coated steel pipes buried in different types of soil for an average period of 23 years. The pipes had been buried in various soil types including clay, sandy clay loam, clay loam, silty clay loam, silty clay and silt loam. The details of the database were published in a separate technical note by Velázquez et al, which included maximum depth of each pit, time that the pipe had been exposed to soil (buried) type of coating in each pipe and pipe to soil potential. Moreover, eight important factors that represent soil properties are reported being: redox potential, pH, resistivity, water content, bulk density, chloride, bicarbonate and sulfate content. The 259 datasets can be separated based on soil categories into 110 datasets from clay soil, 79 datasets from clay loam soil, 61 datasets from Sandy clay loam soil and the remainder of the samples are from silty clay and silty clay loam. Table 1 summarizes the range and the statistical properties of the 259 datasets, where all factors values are represented using minimum (X_{min}), maximum (X_{max}) , mean (X_{mean}) and standard deviation (X_{StD}) values.

Table. 1: Statistical attributes of Velázquez et al. datasets (2010) [48]

Variable, (Units)	X_{min}	X_{max}	X_{mean}	X_{StD}
Max. pit depth, (mm)	0,41	13,44	2,02	2,05
Exposure time (years)	5	50	22,99	9,12
Resistivity, (X–m)	1,9	399,5	50,15	55,92
Water content, (%)	8,8	66	23,9	6,66
Sulphate, (ppm)	0,99	1370,2	152,97	168,18
Bicarbonate, (ppm)	0,99	195,2	19,67	25,33
Chloride, (ppm)	0,99	672,7	47,73	75,16
pH	4,14	9,88	6,13	0,93
Pipe/soil potential, (V)	-1,97	-0,42	-0,88	0,24
Bulk density, (g/ml)	1,1	1,56	1,3	0,09
Redox potential, (mV)	2,1	348	167,04	85,48
Coating type	0,3	1	0,77	0,13

4.2. Empirical Models

4.2.1. Velazquez et al. model [14]

According to Velazquez et al (2009) [14],the Rossum model was modified and the maximum pit depth can be expressed as explained in Equation 29.

$$d(T) = K(T - T_0)^n \tag{29}$$

where *d* represents the maximum pit depth, T_0 is the pit initiation time, and *k* and *n* are regression parameters depending on the soil properties. In their study, Velazquez et al. (2009) [14]used the aforementioned database with nonlinear regression to extract the formulations of *k* and *n*. Therefore, the following equations express *k* and *n* formulas:

$$\begin{cases} K = k_0 + \sum_{i=1}^{n} k_i X_i \\ n = n_0 + \sum_{i=1}^{m} n_i X_i \end{cases} \tag{30}$$

where K_i and n_i are regression coefficients related to the predicted possible factor X_i. Table 2 represents the outcome of the regression results.

Table. 2: Regression results derived from Velazquez et al. (2009) [14] study

Coefficient	Factor	Value
T_0	Initial time	2,88
K_0	Constant	0,608
n_0	Constant	0,896
k_1	Redox potential,(*mv*)	-0,00018
k_2	Ph	-0,0654
k_3	Resistivity, (*X–m*)	-0,00026
k_4	Chloride, (*ppm*)	0,00087399
k_5	Bicarbonate, (*ppm*)	-0,000639
k_6	Sulphate, (*ppm*)	-0,000122
n_1	Pipe/soil potential, (*V*)	0,519
n_2	Water content, (*%*)	0,00465
n_3	Bulk density, (*g/ml*)	-0,099
n_4	Coating type	0,431

4.2.2. Alamilla et al. model [15]

According to Alamilla et al. (2009) [15] the mathematical formulation that describes the propagation of localized corrosion damage can be expressed by using Equation 31.

$$d(T) = v_p T + \frac{(v_0 - v_p)}{q_0}[1 - \exp(-q_0 T)] \quad (31)$$

Whereas previously *d* represented the maximum pit depth, v_0, v_p are respectively the initial and long- term corrosion velocities, and q_0 is regression constant. In their model, only $v_p$is considered as a function of surrounding environmental parameters. Therefore, the formula of long-term velocity was given by Alamilla et al (2009) as follow:

$$v_p = C_0 \exp[-(q_1 pH + q_2 \rho + q_3 E_{Redox} + q_4 E_{s-d})] \quad (32)$$

In Equation 32, C_0, v_0 and q_i are regression coefficients relating to the predicted possible environment factor as represented in the regression result of Table 3.

Table. 3: Regression results extracted from Alamilla et al. (2009) [15]study

Coefficient	Factor	Value
q_1	Hydrogen potential (*pH*)	0,6623
q_2	Resistivity *(q)/X m*	0,0069
q_3	Potential redox *(E_{Redox})/mV/SHE*	0,0027
q_4	Soil–structure potential*(E_{s-p})/V/C_u/C_{uSO4}*	0,981
v_0	Initial corrosion rate *mm/year*	0,6743
q_0	Constant *1 /year*	1,7326
C_0	Constant 2 *mm/ year*	12,2652

4.3. Implementation of the Developed Models

As indicated in the above framework described in Section 2, the results obtained by using the SVR technique are highly sensitive to accurate estimation of the SVR hyper-parameters. In order to conduct optimum performance, two different approaches are considered to obtain SVR hyper-parameters. First, the traditional trial and error approach of SVR is applied although; it is time-consuming and does not guarantee the selection of optimal solutions due to the manual search. Secondly, the meta-heuristic algorithms are applied where the performance of three inspired

optimization algorithms are investigated (i.e., SVR-GA, SVR-PSO and SVR-FFA). Control parameters of the developed hybrid models i.e. SVR-GA, SVR-PSO and SVR-FFA are listed in Table 4. As explained in Table 4, the initial random population used in the three algorithms are the same and equal to 30.

Table. 4: Default settings for model parameters

Algorithm	Parameters	Setting values
GA	Population size	30
	Crossover's probability	90%
	Mutation's probability	70%
	Type of replacement	Elitism (10% of the population)
	Type of selection	Linear ranking
	Max number of generation	30
PSO	Number of particles	30
	Maximum number of iteration	30
	C_1, $C2$	2.05
	ω_{Max}	1.2
	ω_{Min}	0.1
FFA	Number of fireflies	30
	Maximum number of iteration	30
	alpha	0.5
	beta	4
	gamma	1

5. Comparative predicted results of the models

To implement the proposed hybrid models, the database of section 4.1 with 259 samples is separated randomly into training (80% of the data) and testing (20% of the data) datasets. Initially, the SVR models are applied to the first dataset (Train), and then the outcome of the model is validated by using the second dataset (Test), where different trials are considered to obtain SVR model hyper-parameters(C, ε, γ). Statistical performances of the developed models and their prediction accuracy are illustrated in Table 5. Considering Table 5, the four SVR-based models

show the best performance, based on all statistical indexes, compared to empirical-based correlations developed by Velazquez et al. (2009) and Alamilla et al. (2009). Table 5 results also indicate clearly that the SVR-FFA hybrid model provides the most accurate results for maximum depth prediction of pitting corrosion in pipelines. Specifically, SVR-FFA yields the lowest *MAE, RMSE, MSE* values that equal to 0.0935; 0.0949 and 0.0090, respectively, using the training dataset while the same values using the test dataset are equal to 0.2359; 0.2909 and 0.5588, respectively. Moreover, the highest *CI* results in both phases of train and test are provided by SVR-FFA model being 0.9977 and 0. 8935, respectively.

Further analysis of Table 5 results reveals that SVR-PSO shows more efficient and robust than SVR-GA using the statistical indicators Moreover, the trial and error approach for SVR hyper-parameters selection shows the satisfactory results in the training phase, though the test results have low accuracy compared to the others SVR-hybrid models. Furthermore, results in Table 5 reveal that the abilities and performances of the two correlation models developed by Velazquez et al. (2009) and Alamilla et al. (2009) using the listed indexes of statistical errors have a low accuracy compared to AI models. However, Velazquez et al. (2009) model shows better results than Alamilla et al. (2009) model.

Table. 5: The comprative statistical indicators for different models in training and testing phases

	Models	MAE	RMSE	MSE	d	CI	NSE
Train datasets	SVR	0,0972	0,1107	0,0122	0,9993	0,9969	0,9975
	SVR-GA	0,0955	0,0988	0,0097	0,9994	0,9975	0,998
	SVR-PSO	0,0986	0,1057	0,0111	0,9994	0,9971	0,9977
	SVR-FFA	**0,0935**	**0,0949**	**0,009**	**0,9995**	**0,9977**	**0,9981**
	Velazquez et al.	0,7011	1,1957	2,7445	0,6896	0,4152	0,6021

	Alamilla et al.	4,3805	5,6998	34,181	0,9393	-7,5530	-8,0410
Test datasets	SVR	0,8925	1,1633	1,3533	0,8616	0,3441	0,3994
	SVR-GA	0,5848	0,8389	0,7038	0,917	0,6305	0,6876
	SVR-PSO	0,5719	0,8002	0,6403	0,9185	0,6575	0,7158
	SVR-FFA	**0,2359**	**0,2909**	**0,5588**	**0,9285**	**0,8935**	**0,9624**
	Velazquez et al.	1,2165	2,3153	1,1311	0,5454	0,0620	0,1138
	Alamilla et al.	4,8449	5,8525	37,2137	0,8988	-4,1907	-4,6621

* Bold numbers represent the best-obtained results amongst the others

To illustrate the applicability of the proposed models, the ratio of *d/RMSE* based on Equations. (22) and (26) are determined and the results are plotted in Figure 2. As illustrated in Figure 2, the *d/RMSE* ratio have high values for all SVR models for the training phase. This means that the AI models are accurate and can train this problem with a nonlinear format while the robustness of predictions results depend on the tuning SVR parameters. It is clear that SVR-FFA results have the most accurate results when compared to all other SVR-based models and empirical correlations in both phases. Moreover, SVR-GA and SVR-PSO results are slightly similar, in which the *d/RMSE* values of SVR-GA provide less accuracy than SVR-PSO when considering the testing phase.

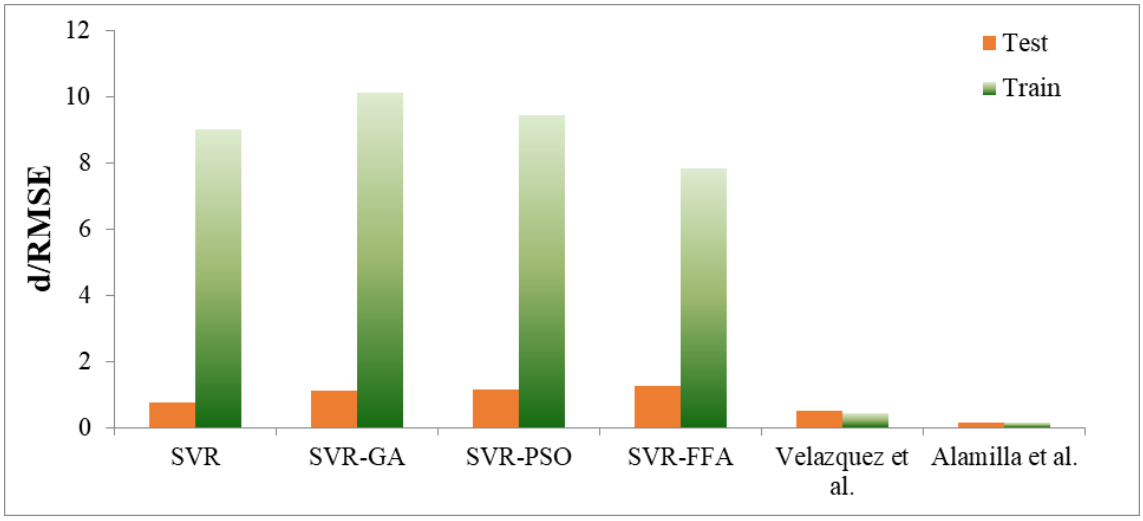
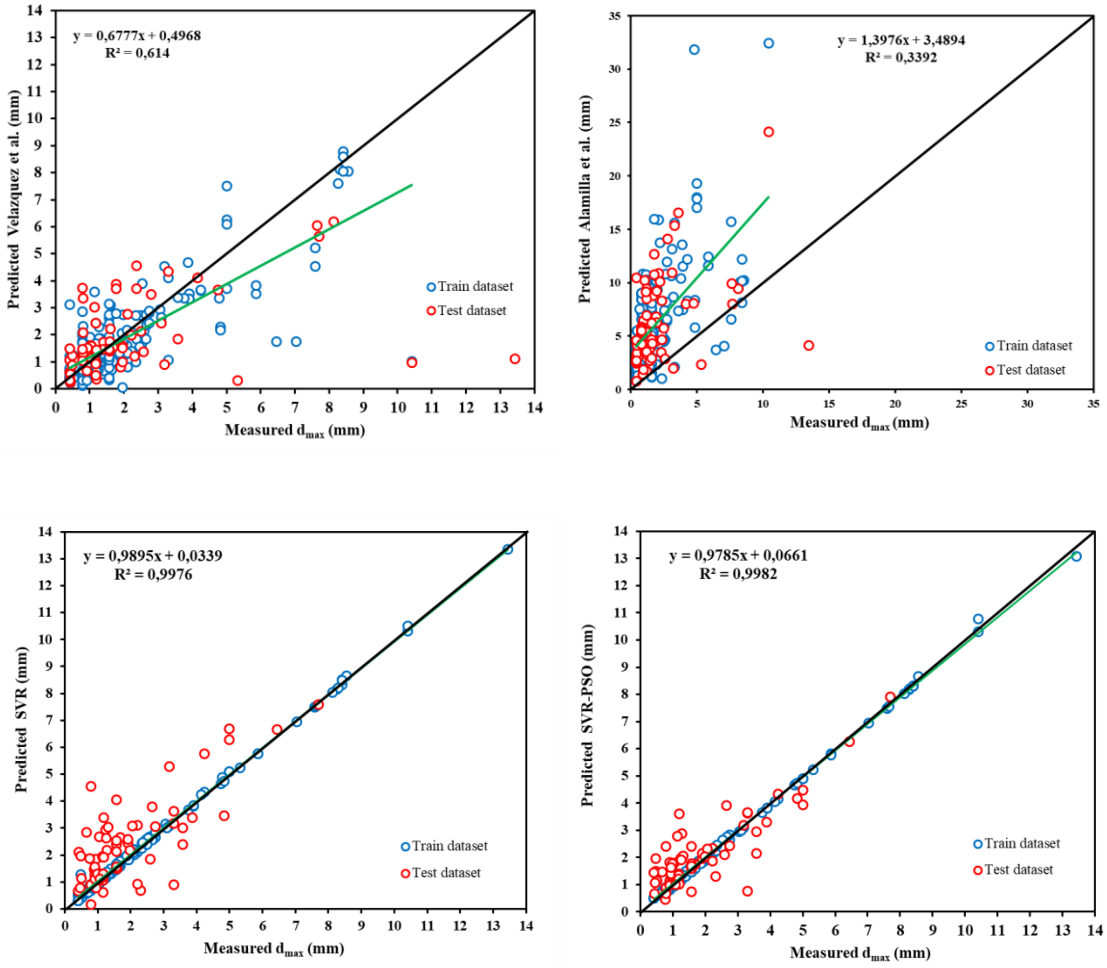


Fig. 2. *d/RMSE* factor for different empirical models and artificial intelligent-based hybrid SVR models in train and test phases

Figure. 3 depicts the scatterplots of the six models, i.e. SVR-FFA, SVR-GA, SVR-PSO, SVR, Velazquez et al. (2009) [14] and Alamilla et al. (2009) [15] based on both training and test datasets. Furthermore, the correlation (R^2) between predicted and measured/observed data points of maximum pitting corrosion depth is also provided in Figure. 3. In these figures, the green lines represent a linear relationship between the predicted ($\hat{d}_{max}$) and measured (d_{max}) values, while black lines refer to a unit slope.

As illustrated in Figure. 3, the best determination coefficient R^2 is provided by SVR-FFA model which is 0.9987 followed by 0.9985, 0.9982, 0.9976 for SVR-GA, SVR-PSO, and SVR respectively. This confirms the accuracy and robustness of the SVR techniques as tools for predicting the maximum depth of pitting corrosion in oil and gas pipelines. Obtained results of Velázquez et al. [14] empirical model is equal to R^2 = 0.614, which is lower compared to the one reported by Velazquez et al. (2009) (R^2=*0.87*). It is noted that Velazquez et al. (2009) eliminated a portion of samples from the original database as outliers, in which this is not the case in the present study. Therefore, the difference between the present study results (R^2=0.614) and

Velázquez et al (2009) results (R^2=*0.87*) return to this elimination of samples that were considered as outliers. In addition, it should be noted that SVR-GA and SVR-PSO results are close, where the maximum difference of R^2 is 0.003. The result obtained with Alamilla et al. (2009) empirical model is the lowest. The corresponding R^2 value is 0.3392, which indicates the incapability of the model against a larger database that contains more variety of data.



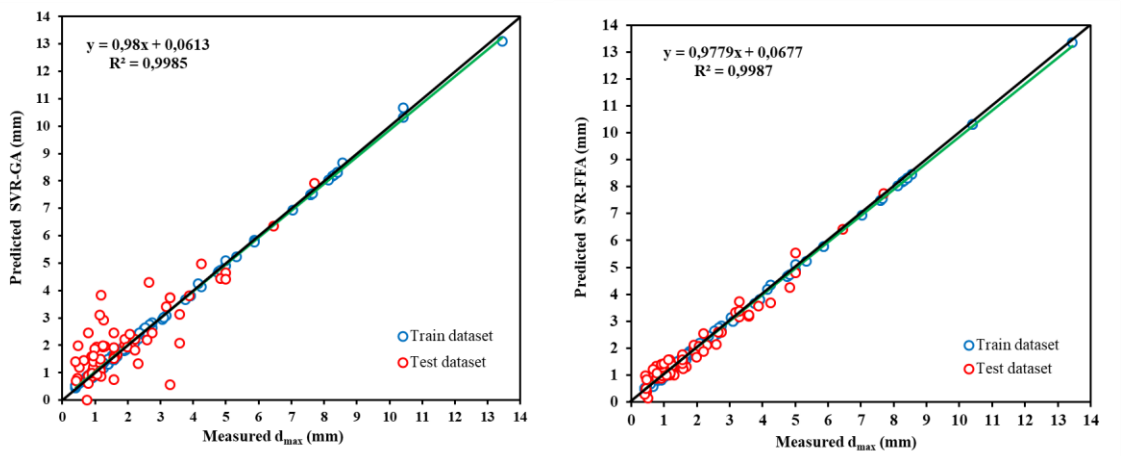


Fig. 3. The scattor plots presentation for the applied predictve models over the training and testing phase

Figure 4 illustrates the Taylor diagram for the test dataset, which is extracted using the hybrid models of SVR-PSO, SVR-FFA and SVR-GA, machine learning-based SVR models and empirical relations developed by Velazquez et al. (2009). Results of Figure. 4 indicated that the SVR and hybrid models are more accurate than the empirical model. However, the accuracies of the predicted data for maximum pitting corrosion depth using hybrid intelligent techniques showed that SVR is inaccurate modelling process compared to hybrid intelligent models of SVR-FFA, SVR-PSO, SVR-GA) due to the random selecting the SVR hyper-parameters. The best and the worst models among the AI schemes are founded to be the SVR-FFA and the SVR models, respectively. Following SVR-FFA results, the hybrid SVR model with PSO shows more accurate prediction results in comparison with SVR-GA and SVR. The optimization process for tuning the parameters of SVR using meta-heuristic algorithms is a robust strategy to improve the accuracy of SVR in complex engineering problems. The ability of the optimization method may differ from other engineering problems. Dataset for training models and parameters of optimization algorithms are the major parameters to provide the nonlinear response for real engineering problems. The

abilities of these three hybrid models can be investigated for other complex engineering problems in future.

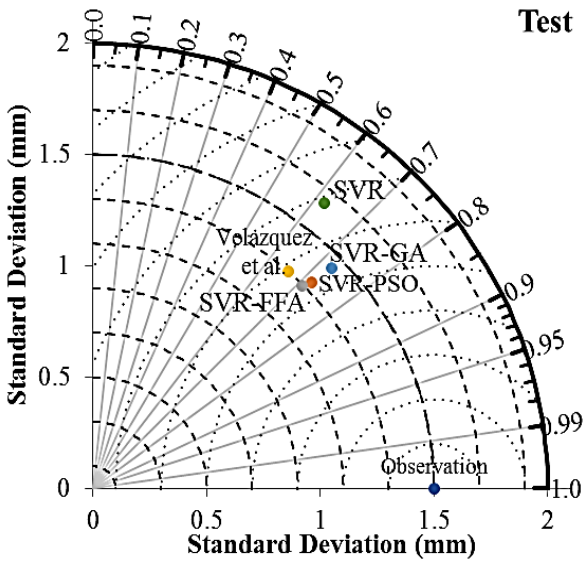


Fig. 4. Taylor diagrams for predicted test data using different models.

Histograms of the ratio between predicted and observed maximum depth of pitting corrosion ($d_{max\text{-}P}/d_{max\text{-}O}$) are plotted in Figure 5. It comprises the mean and standard deviation (SD) of results obtained from each model. When the mean value is close to the unit with a low standard deviation, it indicates that the model performs more accurate in predicting the real-field phenomena. Unlike empirical correlation developed by Velazquez et al. (2009) and Alamilla et al. (2009) models, the mean values of the proposed SVR-hybrid models are the closest to one while their standard deviations are the lowest with different results. The SVR-FFA model provides the lowest SD value (equal to 0.386) compared to the other hybrid models while the SVR-PSO outperforms SVR-GA. The Alamilla et al. (2009) model shows the lowest accuracy for approximating the maximum

pitting corrosion depth. Thus, for accurate and robust predictions of the maximum pitting depth for oil and gas pipelines, the hybrid models based SVR are suitable choices.

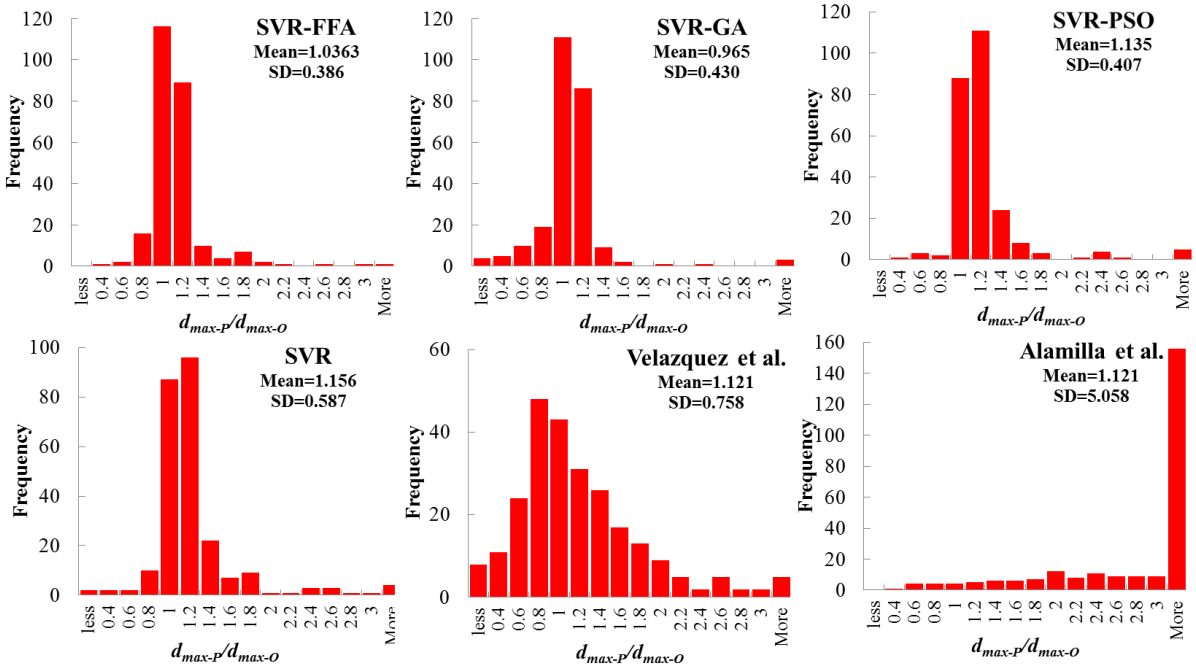


Fig. 5. Comparison of maximum pitting corrosion depth using predicted to observed ($d_{max\text{-}P}/d_{max\text{-}O}$) dataset.

The predicted to an observed maximum depth of pitting corrosion ($d_{max\text{-}P}/d_{max\text{-}O}$) ratios for different models are presented with respect to the time exposure in Figure 6 for the train and test datasets. It is shown from Figure. 6, that the ratio $d_{max\text{-}P}/d_{max\text{-}O}$ values are the nearest to one in training data for the machine learning approaches of, SVR, SVR-PSO and SVR-GA and SVR-FFA compared to the empirical models. By comparing the test data, the hybrid model of SVR-FFA is accurate modeling procedure than the models of SVR, SVR-GA and empirical formulations. In fact, the obtained regression models resulting from the adopted empirical formulations present a high percentage of errors compared to the developed intelligent models. The SVR-FFA provides the

most accurate results for time exposure with more than 30 years while its prediction results are sensitive within the range goes from 10 to 25 years.

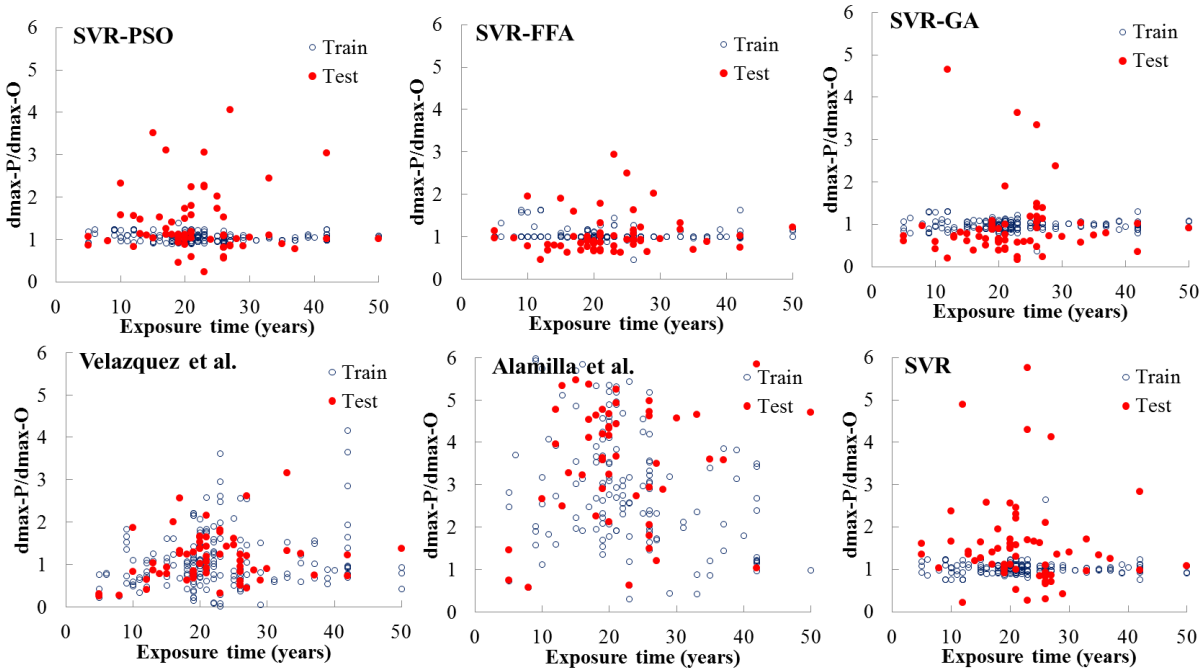


Fig. 6. Comparison of predicted to observed data points corresponding to exposure time for different models

Figure.7 presents the uncertainties of predicted models using the train and test data for maximum corrosion depth of pipelines. As seen, the empirical models are not robust schemes for the prediction thus it is required to improve the nonlinearities of these models as well as the machine learning approaches. By comparing the results of the best model of SVR-FFA, highly accurate and acceptable prediction results are obtained for depth larger than 5 mm while a high uncertainty is available for depth less than 3mm. This mean that the soil properties affecte on the corrosion defetc for long times exposure, which led to the increase of the corrosion depth.

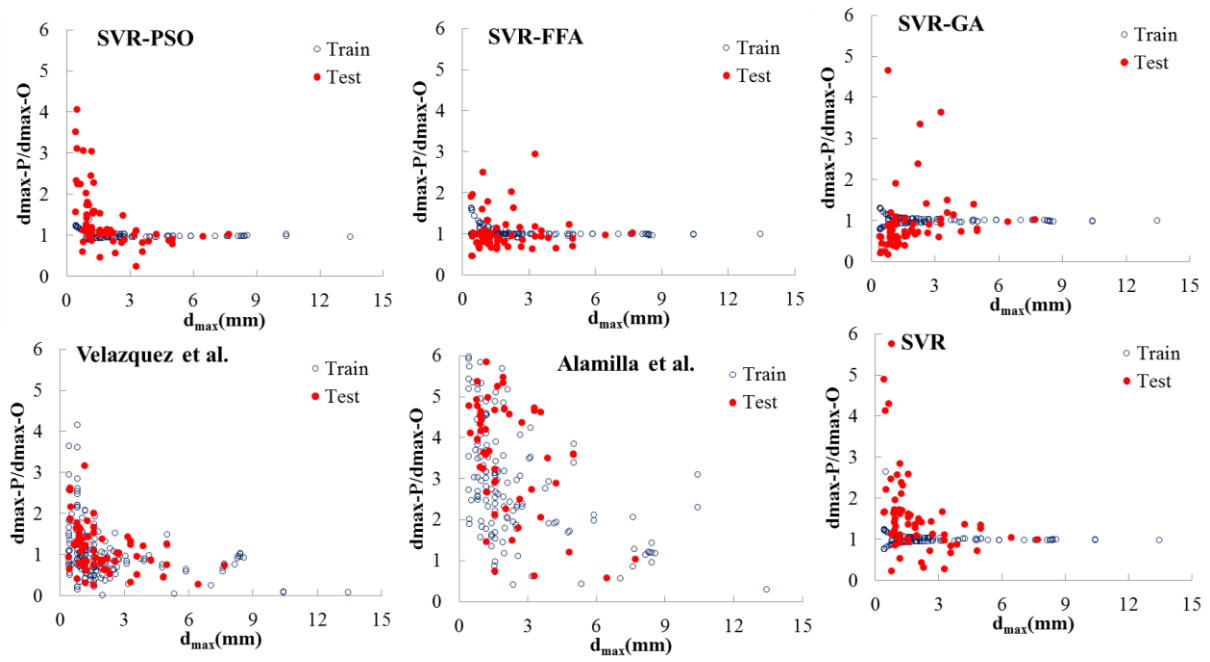


Fig. 7. Comparison of predicted to observed data points of maximum pitting corrosion depth for the proposed AI approaches and empirical models

6. CONCLUSIONS

Predicting maximum depth of pitting corrosion in oil and gas pipelines is significantly important in maintaining high reliability and security levels of those structures. In the present study, a new methodology based on using an artificial intelligence technique, namely SVR, has been developed to more accurately predict the depth of pitting corrosion in metallic pipes. Moreover, three meta-heuristic algorithms are used to enhance the accuracy of SVR capability in selecting optimal solutions for its hyper-parameters. Therefore, exploiting GA, PSO and FFA optimization algorithms for seeking global optimal hyper-parameters extended the SVR ability to move towards its best generalization capability. A real database of pitting corrosion depth has been adopted which

covers different factors affecting the pitting corrosion. The main conclusions from this study, therefore, are summarized as follows:

- Performance results based on statistical indicators show that all proposed models using hybrid AI (i.e. SVR-GA, SVR-PSO, and SVR-FFA) have high efficiency and accuracy in adapting when predicting the maximum depth of pitting corrosion in oil and gas pipelines;
- Comparative computational results based on adopting Velázquez et al (2009) database indicated that SVR-based hybrid models i.e. SVR, SVR-GA, SVR-PSO, and SVR-FFA outperform empirical ones. The improvement from Velazquez' model to the best-obtained hybrid model .i.e. SVR-FFA using *RSME* is about 94.23% and 29.72 % in train and test phases. Where SVR-FFA has the highest coefficient of determination ($R^2=0.9987$).
- The results of the present study indicate that assumptions adopted by empirical models may lead to the inaccurate performance of the models and low accuracy of the outcomes, in which the Alamilla et al. (2009) model shows the lowest performance among all models.

This study provides a guideline for using artificial intelligence techniques as an appropriate framework for modelling the depth of pitting corrosion in oil and gas pipelines in terms of accuracy and efficiency in performance. However, new machine learning techniques integrated with further optimizations algorithms should be investigated to enhance the accuracy of predicting the depth of pitting corrosion in oil and gas pipelines in the future.

REFERENCES

[1] M. Bagheri, Peng, S.-P. Zhu, B.S. Ben Mohamed El Amine, B. Keshtegar, Hybrid intelligent method for fuzzy reliability analysis of corroded X100 steel pipelines, Eng. Comput. (2020). doi:10.1007/s00366-020-00969-1.

[2] M.S. El-Abbasy, A. Senouci, T. Zayed, F. Mirahadi, L. Parvizsedghy, Artificial neural

network models for predicting condition of offshore oil and gas pipelines, Autom. Constr. 45 (2014) 50–65. doi:10.1016/j.autcon.2014.05.003.

[3] M.S. El-Abbasy, A. Senouci, T. Zayed, F. Mirahadi, L. Parvizsedghy, Condition Prediction Models for Oil and Gas Pipelines Using Regression Analysis, J. Constr. Eng. Manag. 140 (2014) 4014013. doi:10.1061/(ASCE)CO.1943-7862.0000838.

[4] M. El Amine Ben Seghier, B. Keshtegar, B. Elahmoune, Reliability analysis of low, mid and high-grade strength corroded pipes based on plastic flow theory using adaptive nonlinear conjugate map, Eng. Fail. Anal. (2018). doi:10.1016/j.engfailanal.2018.03.029.

[5] B. Keshtegar, M. El Amine Ben Seghier, S. Zhu, R. Abbassi, N. Trung, Journal of Loss Prevention in the Process Industries Reliability analysis of corroded pipelines : Novel adaptive conjugate first order reliability method, J. Loss Prev. Process Ind. 62 (2019) 103986. doi:10.1016/j.jlp.2019.103986.

[6] C. Lam, W. Zhou, Statistical analyses of incidents on onshore gas transmission pipelines based on PHMSA database, Int. J. Press. Vessel. Pip. 145 (2016) 29–40. doi:10.1016/j.ijpvp.2016.06.003.

[7] A. Valor, F. Caleyo, J.M. Hallen, J.C. Velazquez, Reliability assessment of buried pipelines based on different corrosion rate models, Corros. Sci. 66 (2013) 78–87. doi:10.1016/j.corsci.2012.09.005.

[8] A. Valor, L. Alfonso, F. Caleyo, J. Vidal, E. Perez-Baruch, J.M. Hallen, The negative binomial distribution as a model for external corrosion defect counts in buried pipelines, Corros. Sci. 101 (2015) 114–131. doi:10.1016/j.corsci.2015.09.009.

[9] A. Chen, F. Cao, X. Liao, W. Liu, L. Zheng, J. Zhang, C. Cao, Study of pitting corrosion on mild steel during wet-dry cycles by electrochemical noise analysis based on chaos

theory, Corros. Sci. 66 (2013) 183–195. doi:10.1016/j.corsci.2012.09.017.

[10] T. Wu, J. Xu, C. Sun, M. Yan, C. Yu, W. Ke, Microbiological corrosion of pipeline steel under yield stress in soil environment, Corros. Sci. 88 (2014) 291–305. doi:10.1016/j.corsci.2014.07.046.

[11] J.R. Scully, N.D. Budiansky, Y. Tiwary, A.S. Mikhailov, J.L. Hudson, An alternate explanation for the abrupt current increase at the pitting potential, Corros. Sci. 50 (2008) 316–324. doi:10.1016/j.corsci.2007.08.002.

[12] M. El Amine Ben Seghier, M. Bettayeb, J. Correia, A. De Jesus, R. Calçada, Structural reliability of corroded pipeline using the so-called Separable Monte Carlo method, J. Strain Anal. Eng. Des. (2018).

[13] S. Sim, I.S. Cole, Y.S. Choi, N. Birbilis, A review of the protection strategies against internal corrosion for the safe transport of supercritical CO2 via steel pipelines for CCS purposes, Int. J. Greenh. Gas Control. 29 (2014) 185–199. doi:10.1016/j.ijggc.2014.08.010.

[14] J.C. Velázquez, F. Caleyo, A. Valor, J.M. Hallen, Predictive Model for Pitting Corrosion in Buried Oil and Gas Pipelines, Corrosion. 65 (2009) 332–342. http://corrosionjournal.org/doi/abs/10.5006/1.3319138.

[15] J.L. Alamilla, E. Sosa, Stochastic modelling of corrosion damage propagation in active sites from field inspection data, Corros. Sci. 50 (2008) 1811–1819. doi:10.1016/j.corsci.2008.03.005.

[16] F. Caleyo, J.C. Velázquez, A. Valor, J.M. Hallen, Markov chain modelling of pitting corrosion in underground pipelines, Corros. Sci. 51 (2009) 2197–2207. doi:10.1016/j.corsci.2009.06.014.

[17] A. Valor, F. Caleyo, L. Alfonso, J.C. Velazquez, J.M. Hallen, Markov chain models for the

stochastic modeling of pitting corrosion, Math. Probl. Eng. 2013 (2013). doi:10.1155/2013/108386.

[18] A. Valor, F. Caleyo, L. Alfonso, D. Rivas, J.M. Hallen, Stochastic modeling of pitting corrosion: A new model for initiation and growth of multiple corrosion pits, Corros. Sci. 49 (2007) 559–579. doi:10.1016/j.corsci.2006.05.049.

[19] M. El Amine Ben Seghier, B. Mourad, B.E. et Al, Probabilistic approach evaluates reliability of pipelines with corrosion defects, OIL GAS J. 115 (2017) 64–68.

[20] Mohamed el Amine Ben Seghier, B. Keshtegar, J.A.F.O. Correia, G. Lesiuk, A.M.P. De Jesus, Reliability analysis based on hybrid algorithm of M5 model tree and Monte Carlo simulation for corroded pipelines : Case of study X60 Steel grade pipes, Eng. Fail. Anal. 97 (2019) 793–803. doi:10.1016/j.engfailanal.2019.01.061.

[21] M. Seghier, B. Keshtegar, J. Correia, A. De Jesus, G. Lesiuk, Structural Reliability Analysis of Corroded Pipeline made in X60 Steel Based M5 Model Tree Algorithm and Monte Carlo Simulation, Procedia Struct. Integr. 13 (2018) 1670–1675. doi:10.1016/j.prostr.2018.12.349.

[22] K.O. Akande, T.O. Owolabi, S.O. Olatunji, A.A. AbdulRaheem, A hybrid particle swarm optimization and support vector regression model for modelling permeability prediction of hydrocarbon reservoir, J. Pet. Sci. Eng. 150 (2017) 43–53. doi:10.1016/j.petrol.2016.11.033.

[23] B. Keshtegar, D. Meng, M. El Amine Ben Seghier, N. Trung, D. Tien, A hybrid sufficient performance measure approach to improve robustness and efficiency of reliability-based design optimization, Eng. Comput. (2020). doi:10.1007/s00366-019-00907-w.

[24] H. Jung, H. Jo, S. Kim, K. Lee, J. Choe, Geological model sampling using PCA-assisted

support vector machine for reliable channel reservoir characterization, J. Pet. Sci. Eng. (2018). doi:10.1016/j.petrol.2018.04.017.

[25] S.S. Ali, S. Nizamuddin, A. Abdulraheem, M.R. Hassan, M.E. Hossain, Hydraulic unit prediction using support vector machine, J. Pet. Sci. Eng. (2013). doi:10.1016/j.petrol.2013.09.005.

[26] E.A. El-Sebakhy, Forecasting PVT properties of crude oil systems based on support vector machines modeling scheme, J. Pet. Sci. Eng. (2009). doi:10.1016/j.petrol.2008.12.006.

[27] J. Hou, Z. quan Li, X. long Cao, X. wang Song, Integrating genetic algorithm and support vector machine for polymer flooding production performance prediction, J. Pet. Sci. Eng. (2009). doi:10.1016/j.petrol.2009.05.017.

[28] Y.F. Wen, C.Z. Cai, X.H. Liu, J.F. Pei, X.J. Zhu, T.T. Xiao, Corrosion rate prediction of 3C steel under different seawater environment by using support vector regression, Corros. Sci. 51 (2009) 349–355. doi:10.1016/j.corsci.2008.10.038.

[29] R.H. J.C. Xiu, L. Yong, No Title, Appl. Soft Comput. 8 (2008) 1222.

[30] C.L. a Xiaoli Zhang a, n, Peng Wang a, Dakai Liang b, Chunfeng Fan a, A soft self-repairing for FBG sensor network in SHM system based on PSO–SVR model reconstruction, Opt. Commun. (2015).

[31] V.N. Vapnik, Statistical Learning Theory, 1998. doi:10.2307/1271368.

[32] P.S.L. R.J. Kuo , Taiwanese export trade forecasting using firefly algorithm based K-means algorithm and SVR with wavelet transform, Comput. Ind. Eng. (2016) 153–161.

[33] S.-J. Yeh, C.-Y., Huang, C.-W., & Lee, A multiple-kernel support vector regression approach for stock market price forecasting, Expert Syst. Appl. 3 (2011) 2177–2186.

[34] R.J. Kuo, P.S. Li, Taiwanese export trade forecasting using firefly algorithm based K-means

algorithm and SVR with wavelet transform, Comput. Ind. Eng. 99 (2016) 153–161. doi:10.1016/j.cie.2016.07.012.

[35] C.J. Lu, Sales forecasting of computer products based on variable selection scheme and support vector regression, Neurocomputing. 128 (2014) 491–499. doi:10.1016/j.neucom.2013.08.012.

[36] W.C. Hong, Y. Dong, F. Zheng, C.Y. Lai, Forecasting urban traffic flow by SVR with continuous ACO, Appl. Math. Model. 35 (2011) 1282–1291. doi:10.1016/j.apm.2010.09.005.

[37] T. Kavzoglu, I. Colkesen, A kernel functions analysis for support vector machines for land cover classification, Int. J. Appl. Earth Obs. Geoinf. 11 (2009) 352–359. doi:10.1016/j.jag.2009.06.002.

[38] Z. Yaseen, O. Kisi, V. Demir, Enhancing Long-Term Streamflow Forecasting and Predicting using Periodicity Data Component: Application of Artificial Intelligence, Water Resour. Manag. (2016). doi:10.1007/s11269-016-1408-5.

[39] Z.M. Yaseen, M.T. Tran, S. Kim, T. Bakhshpoori, R.C. Deo, Shear strength prediction of steel fiber reinforced concrete beam using hybrid intelligence models: A new approach, Eng. Struct. 177 (2018) 244–255. doi:10.1016/j.engstruct.2018.09.074.

[40] J. Kennedy, R. Eberhart, Particle swarm optimization, Neural Networks, 1995. Proceedings., IEEE Int. Conf. 4 (1995) 1942–1948 vol.4. doi:10.1109/ICNN.1995.488968.

[41] M. Kumar, M. Husian, N. Upreti, D. Gupta, Genetic Algorithm: Review and Application, Int. J. Inf. Technol. Knowl. Manag. 2 (2010) 451–454.

[42] S. Liu, H. Tai, Q. Ding, D. Li, L. Xu, Y. Wei, A hybrid approach of support vector regression with genetic algorithm optimization for aquaculture water quality prediction,

Math. Comput. Model. 58 (2013). doi:10.1016/j.mcm.2011.11.021.

[43] X. Yang, Firefly algorithm, stochastic test functions and design optimization, Int. J. Bio Insp. Comput. 2 (2010) 78–84.

[44] M.A. Ghorbani, R.C. Deo, Z.M. Yaseen, M.H. Kashani, Pan evaporation prediction using a hybrid multilayer perceptron-firefly algorithm (MLP-FFA) model : case study in North Iran, (2017). doi:10.1007/s00704-017-2244-0.

[45] M.A. Ghorbani, R.C. Deo, V. Karimi, Z.M. Yaseen, O. Terzi, Implementation of a hybrid MLP-FFA model for water level prediction of Lake Egirdir, Turkey, Stoch. Environ. Res. Risk Assess. (2017) 1–15. doi:10.1007/s00477-017-1474-0.

[46] H. Tao, S.O. Sulaiman, Z.M. Yaseen, H. Asadi, S.G. Meshram, M.A. Ghorbani, What Is the Potential of Integrating Phase Space Reconstruction with SVM-FFA Data-Intelligence Model? Application of Rainfall Forecasting over Regional Scale, Water Resour. Manag. (2018). doi:10.1007/s11269-018-2028-z.

[47] B. Keshtegara, M. el A. Ben Seghier, Modified response surface method basis harmony search to predict the burst pressure of corroded pipelines, Eng. Fail. Anal. 89 (2018) 177–199. doi:10.1016/j.engfailanal.2018.02.016.

[48] J.C. Velázquez, F. Caleyo, A. Valor, J.M. Hallen, Technical Note: Field study - Pitting corrosion of underground pipelines related to local soil and pipe characteristics, Corrosion. 66 (2010) 0160011–0160015. doi:10.5006/1.3318290.