
Ali Reza Mahmodian¹, Behrouz Yaghoubi²*, Fariborz Yosefvand²

¹ Ph.D. Candidate, Department of Water Engineering, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran
² Department of Water Engineering, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran; *behrouz.yaghoubi.h@gmail.com

ABSTRACT

Nowadays, a huge amount of natural resources such as gases and oil are exploited from offshore oil fields and transported by pipes located at seabed. The pipelines are exposed to waves and currents and scour may occur around them. Subsequently, stability of the pipes can be threatened, so estimation and simulation of scouring around the pipes are quite vital. In this study, a hybrid method for simulating the scour depth in the vicinity of submerged pipes was developed. In other words, the adaptive neuro-fuzzy inference system (ANFIS) and the differential algorithm were combined with each other to simulate the scour depth. In general, ANFIS is an artificial neural network acts based on the Takagi-Sugeno inference system. This model is a set of if-then rules which is able to approximate non-linear functions. In addition, the differential algorithm is a powerful evolutionary algorithm among optimization algorithms which have many applications in scientific fields. In this study, the Monte-Carlo simulation was employed for examining the ability of numerical models. To validate the modeling results, the k-fold cross validation approach was also utilized with k=6. Then, the parameters affecting the scour depth were detected and six ANFIS and hybrid models were developed for scour estimation. After that, the results of the mentioned models were examined and this analysis showed that the superior model predicts scour values in terms of all input parameters. This model has reasonable accuracy. For example, the values of R and RMSE for this model were calculated 0.974 and 0.079, respectively. Furthermore, the analysis of the modeling results indicated that the ratio of the pipe distance from the sedimentary bed to the pipe diameter (e/D) was identified as the most effective parameter.

1. Introduction

Generally, in coastal regions transmission lines located horizontally on the sea bed are used for operating from oil and gas resources. Due to the existence of transitional flows and tide waves, the possibility of scouring occurrence around submerged pipes must be considered. Erosion of beds beneath pipelines might lead to failure and damage. Therefore, taking effective measures for scour prediction and estimation are essential. Thus, several studies have been carried out by different researchers. Hansen et al. (1986) using the potential theory, analytically examined the scour in the vicinity of pipes located on sedimentary beds. He validated the analytical results with experimental data. Later, Sumer et al. (1988) by conducting an experimental study, evaluated the influence of transverse flows and pipe vibration on the scour pattern around pipelines located on live-beds. Also, Mao (1988) by studying the scour pattern in the vicinity of submerged pipes, stated that there are three types of vortexes forming around the mentioned pipes. Chiew (1993) investigated the influence of vanes attached to submerged pipe walls. He showed that the presence of the mentioned vanes increases the intensity and volume of the scour hole occurring in the vicinity of pipes up to 1.3 times the ordinary pipes.
Moncada et al. (1999) by carrying out an experimental study, examined the scour pattern around horizontal pipes located on sedimentary beds. According to the results, by increasing the Froude number causes to increase the depth and length of the scour hole as well. They also demonstrated that by increasing the diameter of sediment particles, the scour hole dimensions is reduced. Teh et al. (2003) experimentally studied the marine pipelines on unstable and liquefied seabed. By analyzing the experimental results, they proposed a method for designing pipes in such conditions. In addition, Dey and Singh (2008) experimentally studied the scour pattern under clear-water conditions around submerged pipes located on sedimentary beds. They studied the influence of various hydraulic and geometric conditions such as the pipe shape and the distance of the pipe from the bed. They also investigated the influence of the protected bed and stated that the scour depth for the unprotected bed is more than the protected bed. Wu and Chiew (2012) measured the 3D scour pattern in the vicinity of submerged pipes located on sedimentary beds under clear-water conditions by conducting an experiment. They examined the influence of different parameters on the scour hole and indicated that by increasing the Froude number and Shields parameters, the scouring process accelerates. Yang et al. (2012) experimentally studied the impact of the presence of rubber sheets and flow guidance vanes on the scour pattern around submerged pipes subjected to waves. They indicated that the use of rubber sheets with 1.5 times pipe dimensions is the best way for protecting sedimentary beds. Luan et al. (2015) in a numerical study simulated the effects of submerged pipe vibrations on the bed scour pattern in a two-dimensional way. They proved that by increasing the frequency and the range of vibrations, the scour hole dimensions increase as well. Currently, various neural networks and different neuro-fuzzy techniques are used in pattern cognition, prediction and estimation of complex hydrological and hydraulic phenomena. Etemad-Shahidi and Kazeminezhad (2011) by means of the M5’model, approximated the scour pattern around submerged pipes under clear-water and live-bed conditions. By reviewing the modeling results, they deduced that the Shields number is the most effective factor in the scour prediction under the mentioned conditions. Furthermore, they suggested a number of relationships for calculating scouring. Also, Najafzadeh et al. (2014) using the group method of data handling (GMDH) method, the ANFIS model, the tree model and empirical equations, predicted the scour pattern around submerged pipes located on sedimentary beds. By evaluating the results, they demonstrated that the GMDH method forecasts scour values with higher accuracy.

On the one hand, natural resources such as oil and gases are transported by pipelines and after a while scour phenomenon may occur in the vicinity of such facilities. The scouring can damage the pipelines and threatened the stability of these infrastructures. On the other hand, many artificial intelligence studies have been carried out to model various engineering problems, with increasing such studies every day. The AI models such as Artificial Neural Networks (ANNs), Support Vector Machine (SVM), Gene Expression Programing (GEP) and Adaptive neuro fuzzy inference systems (ANFIS) are quite popular and practical. Generally, ANFIS network has been widely applied to model multifarious problems because it is a universal estimator and integrates both neural networks (NNs) and fuzzy logic principles in a single framework. By contrast, the model has some limitations. For instance, ANFIS uses different rules and numerous optimized parameters for modelling, so uncertainty of results is pretty high. Moreover, optimization algorithms like such as gradient descent and least-square approaches may be useful, however they may have some disadvantages including trapping in local optimum and high-volume of computations. To overcome these challenges, other optimization tools are utilized in different areas. For instance, genetic algorithm (Azimi et al 2017a), singular value decomposition (Khoshbin et al. 2016), differential evolution (Ebtehaj et al. 2017; Azimi et al. 2017b) and firefly algorithm (Azimi et al. 2018) were applied to optimize artificial intelligence models. Therefore, for the first time, differential evolution (DE) algorithm is applied for optimization of the ANFIS network in order to estimate scour around submerged pipelines in clear-water conditions. To do this, firstly, the applied numerical models are introduced. Secondly, the experimental model will be presented. Next, using the important input parameters, six ANFIS and ANFIS-DE are defined. Then, numerical models are analyzed and the most effective input parameter and the superior model are identified.

2. Material and methods
2.1. Adaptive Neuro-Fuzzy Inference System
This method was introduced by Yang (1993) for the first time. The structure of ANFIS comprises five layers including input nodes, base nodes, medium nodes, response nodes and output nodes which are directly connected (Figure1). Each node is function of adjustable or fixed parameters. The proper structure of the ANFIS technique is chosen proportional with output data. In the learning step, by correcting parameters of the degree of membership based on the acceptable error, input values get closer to actual values. The ANFIS technique uses neural network learning and fuzzy logic algorithms for conducting non-linear mapping into the space between inputs and


2
outputs. This algorithm also has a good ability in learning, construction and classification. In addition, this system is able to extract fuzzy rules from numerical data or the expert knowledge and creates a rule-base adaptively. Furthermore, this algorithm can adjust the transformation of mankind complex intelligence to the fuzzy system. The learning rule is based on the error back-propagation with a view to minimizing the error between the network output and the actual output. It worth noting that other evolutionary algorithms could be also used in such cases.

![Figure 1 - Architecture of ANFIS network](image)

Most fuzzy inference systems comprise three types including the Mamdani, Sugeno and Tsukamoto systems which in most cases the Mamdani system is employed, however the Sugeno system has a better performance and has the actual output. Thus, in this study, by programming in the MATLAB, the modeling is conducted by the ANFIS method using the Sugeno system. In addition, the hybrid learning algorithm which is a combination of the back-propagation and the least square method is used for learning and adapting with the fuzzy inference system. The membership function in the fuzzy inference system comprises several adjustable parameters. In order to achieve an optimized modeling, these parameters must be optimized. Thus, a more powerful algorithm is required for defining the values. There are many optimized algorithms which are able to enhance the performance of fuzzy systems. The differential evolution (DE) algorithm is one of these algorithms which is a suitable tool for optimization. This algorithm is able to minimize the error between the model output and the actual value of learning data.

### 2.2. Differential Evolution Algorithm

The differential evolution (DE) optimization algorithm provided by Storm R. and Price (1995). As a global search method, it is employed for optimizing membership function parameters in combination with ANFIS. The DE algorithm is a simple search approach yet with a powerful population based on the direct random search often used for selecting network parameters. The objective of DE is minimizing the objective function $f(\theta)$. Where, $\theta \in \mathbb{R}^D$ is the parameter vector. The DE algorithm produces a population of NP unique parameter vectors in order to reach the global optimization. In the generation $G$, the $i_{th}$ parameter vector is as follows:

$$\theta_{i,G} = [\theta_{i,G}^1, \theta_{i,G}^2, \ldots, \theta_{i,G}^D] \quad i = 1, \ldots, NP$$

The general procedure of the DE algorithm includes four steps as follows:

1. **First step: initial initialization**
   A set of NP unique parameter vectors $\theta_{i,G}$ is used to cover the parameter space so that the following equation is established:

   $$\theta_{min} = [\theta_{min}^1, \theta_{min}^2, \ldots, \theta_{min}^D] \quad \text{and} \quad \theta_{max} = [\theta_{max}^1, \theta_{max}^2, \ldots, \theta_{max}^D]$$

   are determined as minimum and maximum criteria, respectively.

2. **Second step: mutation**
   Based on the unique parameter vector $\theta_{i,G}$ in the current generation, the mutated vector $v_{i,G}$ is produced through a special mutated strategy. Many mutated strategies have been proposed for solving different problems. Here, four important strategies are expressed.

   **Strategy 1:** DE/rand/1
   $$v_{i,G} = \theta_{i,G} + F(\theta_{i,G} - \theta_{i,G})$$

   **Strategy 2:** DE/rand-to-best/2
   $$v_{i,G} = \theta_{i,G} + F(\theta_{best,G} - \theta_{i,G}) + F(\theta_{i,G} - \theta_{i,G})$$

   **Strategy 3:** DE/rand/2
   $$v_{i,G} = \theta_{i,G} + F(\theta_{i,G} - \theta_{i,G}) + F(\theta_{i,G} - \theta_{i,G})$$

   **Strategy 4:** DE/current-to-rand/1
   $$v_{i,G} = \theta_{i,G} + K(\theta_{i,G} - \theta_{i,G}) + F(\theta_{i,G} - \theta_{i,G})$$

   In all equations, $r_1, r_2, r_3, r_4, r_5, r_6$, integer indices are unique produced randomly in the domain on $[1, \ldots, NP]$ and are different from the index $i$. The positive reinforcement factor $F$ is used for controlling the scale of differential vectors and is usually chosen in the range of $0 \leq F \leq 2$. The control parameter $K$ is produced in the range of $0 \leq K \leq 1$.

The DE/rand/1 strategy is suitable for solving multi-purposes problems due to its powerful search ability, though its slow convergence speed. The DE/rand-to-best/2 strategy based on the best solution found so far, convergences fast in encountering with strategic issues and acts well. Although, during solving multi-purposes problems this strategy is stuck in the local optimization and it leads to fast convergence. Two strategies based on the differential vector (DE/rand-to-best/2 and DE/rand/2) have better mutation compared to a based on differential vector strategy, but they
have high computational cost. DE/current-to-rand/1 is an unchangeable rotation. As shown, this strategy is efficient in solving multi-purposes optimization problems.

Third step: reproduction

After production of all mutated vectors, the reproduction method is implemented for increasing variations of disruptive vectors. For each mutated vector \( v_{i,G} = [v_{i,G}^1, v_{i,G}^2, \ldots, v_{i,G}^d] \) in the generation \( G \), a test vector \( u_{i,G} = [u_{i,G}^1, u_{i,G}^2, \ldots, u_{i,G}^d] \) is created based on the following reproduction formula:

\[
u_{i,G} = \begin{cases} v_{i,G}^j & \text{if } \text{rand}_j \leq \text{CR} \\ \theta_{i,G}^j & \text{Otherwise} \end{cases}
\]

Where, CR is the reproduction rate rewritten from the mutated vector for a fraction of the parameter value and chooses a positive value in the range between 0 and 1. \( \text{rand}_j \) is the \( j \)th evolution of the uniform random number production with a result of [0, 1]. \( \text{rand} \) is a selected integer number from [1, \( d \)] and is used to ensure that is at least one parameter in \( u_{i,G} \) different from the objective vector \( \theta_{i,G} \). \( D \) is the dimension of the desired problem.

Fourth step: selection

For each objective function and the test vector related to it, the selection step is carried out by means of the fitness function. The vector with a value less than the objective function is kept as the population of the next generation. Step 2 to step 4 continues until the objective is satisfied or the maximum iteration is achieved.

2.4. Scouring around submerged pipes on sedimentary beds

In general, scouring around horizontal submerged pipes (\( S \)) is a function of the average velocity of the flow (\( U \)), the normal flow depth (\( y \)), the density of water (\( \rho \)), the density of sediment (\( \rho_s \)), the kinematic viscosity of water (\( \mu \)), the channel slope (\( S_0 \)), the channel width (\( B \)), the diameter of bed sediments (\( d_{50} \)), pipe diameter (\( D \)), the distance between the pipe and sedimentary bed before scouring (\( e \)), and gravitational acceleration (\( g \)) (Moncada and Aguirre, 1999):

\[
S = f \left( U, y, \rho, \rho_s, \mu, S_0, B, d_{50}, D, e, g \right)
\]

(8)

By conducting the dimensional analysis and introducing eight dimensionless group, Equation (8) is rewritten as follows:

\[
S / D = f \left( Fr, Re, \tau^*, y / D, D / d_{50}, e / D, S_0, y / B \right)
\]

(9)

Where, \( Fr = U / \sqrt{g \cdot y} \), \( Re = UD \cdot g \), and \( \tau^* = \rho_s \cdot (\rho / \rho_s - 1) \cdot d_{50} \) are the Froude number, the Reynolds number and the Shields dimensionless number due to the sediment transport. In Moncada and Aguirre (1999) study, bed slope (\( S_0 \)) is considered constant and the Reynolds number is placed in a range with no effect on the scour pattern. Thus, Equation (9) is written as follows:

\[
S / D = f \left( Fr, \tau^*, y / D, D / d_{50}, e / D \right)
\]

(10)

The dimensionless parameters of Equation (10) are used as the input parameters of different ANFIS-DE models. In Figure 3, combinations of the Equation (10) parameters for six ANFIS-DE models are illustrated.
In this study, the Monte Carlo simulations are used for examining the abilities of the numerical models. These simulations are a broad classification of computational algorithms using random sampling for calculating numerical results. The Monte Carlo methods are usually implemented for simulating physical and mathematical systems which are not solvable by means of other methods. Furthermore, the k-fold cross validation method is utilized for examining the proficiency of the mentioned models. In the k-fold cross validation method, the main sample is divided into k sub-samples with the same size randomly. Among k sub-samples, one sub-sample is used as the validation data and the remaining (k-1) as the test data of the model. Then, the method repeats k times so that each k sub-sample is used exactly once as the validation data once. The results obtained from the mentioned k layers are averaged and provided as an approximation. The advantage of this method is the random repetition of sub-samples in the test and learning process for all observations. In this paper, the k value is assumed 6. The dealing of the k-fold cross validation method with the experimental data in the test and learning conditions are depicted in Figure 4.

Figure 3- Combinations of input parameters for six ANFIS-DE models

In this study, the correlation coefficient (R), the root mean square error (RMSE) and the scatter index (SI) are used for examining the accuracy of different numerical models:

\[
R = \frac{\sum_{i=1}^{n}(d_i/D)_{\text{Observed}} \cdot (d_i/D)_{\text{Predicted}} - (d_i/D)_{\text{Observed}}^2}{\sqrt{\sum_{i=1}^{n}(d_i/D)_{\text{Observed}}^2 \cdot \sum_{i=1}^{n}(d_i/D)_{\text{Predicted}}^2}}
\]

\[
RMSE = \frac{1}{n} \sum_{i=1}^{n} (d_i/D)_{\text{Predicted}}^2 - (d_i/D)_{\text{Observed}}^2
\]

\[
SI = \frac{RMSE}{(d_i/D)_{\text{Observed}}}
\]

Where, \((d_i/D)_{\text{Observed}}\), \((d_i/D)_{\text{Predicted}}\), \((d_i/D)_{\text{Observed}}\), and n are experimental scour, predicted scour, average experimental scour and the number of experimental measurements, respectively.

As discussed, in this study, six models with different combinations of input parameters are introduced. Model 1 is a function of all inputs, while Model 2 to Model 6 simulate scour values in the vicinity of horizontal submerged pipes by a combination of four input parameters. First, six ANFIS models are investigated. In Figure 5, the variations of the correlation coefficient (R) versus RMSE and SI for the ANFIS models are shown. The scatter plots for different ANFIS models are shown in Figure 6. The ANFIS (1) model is a function of all input parameters. The R value for this model is equal to 0.968. Meanwhile, the values of RMSE and SI for this model are estimated 0.087 and 0.099, respectively. In the following, the models with four input parameters are evaluated. For example, the values of R and RMSE for the ANFIS (2) model are obtained 0.968 and 0.086, respectively. This model is a function of the dimensionless parameters \(\tau^*, y/D, D/d_{50}, e/D\). In other words, the influence of the Froude number is eliminated for scour simulation around submerged pipes. According to the modeling results, the results of Model 1 and Model 2 are very close to each other. In other words, elimination of the influence of the Froude number has no significant impact on the modeling accuracy. ANFIS (3) is a function of \(Fr, y/D, D/d_{50}, e/D\). For this model, the influence of the dimensionless Shields number is neglected due to the sediment transport. For this model, the value of RMSE and R are considered 0.100 and 0.958, respectively. Also, the value of SI for this model is 0.114. In order to estimate the scour depth around horizontal submerged pipes by ANFIS (4), the
influence of $y/D$ is removed. Furthermore, the values of $SI$ and $RMSE$ for the mentioned model are 0.967 and 0.089, respectively. In the following, the accuracy of ANFIS (5) is evaluated. This model estimates the objective function in terms of $Fr$, $\tau^*$, $y/D$, $e/D$. The influence of the parameter $D/d_0$ is neglected for the mentioned model. The values of $RMSE$ and $R$ for the ANFIS (5) mode are calculated 0.116 and 0.939, respectively. Meanwhile, among all models with four inputs the ANFIS (6) model has the lowest accuracy and the maximum error value. In other words, the values of $R$ and $RMSE$ for this model are obtained 0.896 and 0.116, respectively. However, the $SI$ value for the mentioned model is calculated equal to 0.167.

![Figure 5- Variations of R versus RMSE and SI for different ANFIS models](image-url)
Next, the ANFIS-DE models are investigated. In Figure 7, the variations of R versus RMSE and SI for the mentioned models are shown. Also, the scatter plots for these models are depicted in Figure 8. Similar to the ANFIS models, in the hybrid model the ANFIS-DE (1) model has also the highest accuracy among all models. For example, the values of R and RMSE for the mentioned models are calculated 0.974 and 0.079, respectively. For this model, the SI value is estimated 0.090. Also, for the ANFIS-DE (2) model, the R value is equal to 0.973 and the RMSE index is equal to 0.081. Among all ANFIS-DE models with four input parameters, ANFIS-DE (3) has the highest accuracy. For the mentioned model, the values of RMSE and R are estimated 0.974 and 0.079, respectively. Additionally, scatter index for ANFIS-DE (3) is almost 0.090. In the following, the accuracy of the ANFIS-DE (4) model is examined. For this model, the values of R, SI and RMSE are 0.966, 0.103 and 0.090, respectively. For the model, the effect of the $r'$ is eliminated and ANFIS-DE (4) simulates the target parameter as a function of other input parameters. In addition, for the ANFIS-DE (5) model, the values of R and RMSE are 0.952 and 0.104, respectively. Also, for the mentioned model the value of SI is calculated 0.118. It should be noted that the influence of $d/D_{50}$ is removed for ANFIS-DE (5). Among all meta-heuristics models, the ANFIS-DE (6) model has the lowest accuracy and the maximum error value. For the mentioned model, the values of R and RMSE are computed 0.903 and 0.144, respectively. Furthermore, SI index for this model is obtained 0.163. For estimation the scour function using ANFIS-DE (6), the effect of $e/D$ is ignored.

As discussed, the accuracy of the ANFIS-DE models corresponding with the ANFIS models is higher which shows that the hybrid algorithm is optimized. In addition, among all numerical models, the ANFIS-DE (1) has the highest correlation with the experimental values. This model predicts scour values.
in terms of all input parameters. Thus, the ANFIS-DE (1) is detected as the superior model. Furthermore, by eliminating the parameter $e/D$, the modeling accuracy is significantly reduced. So, this parameter is identified as the most effective factor in scour estimation in the vicinity of submerged pipes.

Figure 7 - Variations of R versus RMSE and SI for different ANFIS-DE models
4. Discussion and Conclusion

Today, the transfer of oil derivations and gas condensates produced in the coastal regions is performed through pipelines. In general, pipelines are passed above sedimentary beds and due to the presence of coastal flows and waves the possibility of scouring must be considered. In this study, first, the parameters affecting the scour depth around submerged pipes were identified. Then, a hybrid model was developed using the ANFIS model and the differential evolution (DE) algorithm (ANFIS-DE). Based on the input parameters, six ANFIS models and six ANFIS-DE models were introduced. By analyzing the modeling results, the superior model for each of the ANFIS and hybrid models was detected. The superior models predicted the scour values with reasonable accuracy. For example, the value of the correlation coefficient (R) for the ANFIS and hybrid models were calculated 0.968 and 0.974, respectively. Both models estimated the scour values using a combination of all input parameters. It should be noted that the hybrid algorithm modeled the objective function values with higher accuracy compared to the ANFIS model. In other words, the DE algorithm optimized the ANFIS model. In addition, the analysis of the modeling results indicated that the ratio of the distance between the submerged pipe and the sedimentary bed to the pipe diameter is the most effective factor in simulating the scour depth by the mentioned algorithm.

5. References

2. Sumer, B.M., Mao, Y. and Fredsøe, J., (1988), Interaction between vibrating pipe and erodible bed,