



## Using Artificial Neural Networks to Predict Leakage Location in Water Distribution Networks

**Parisa kolyaee<sup>1</sup>, Hossein Mohammad vali samani<sup>2\*</sup>, Seyed Mohammad Mirhoseini Hazave<sup>3</sup>, Ehsanollah zeyghami<sup>4</sup>**

<sup>1</sup>PhD student, Department of Civil Engineering, Arak Branch, Islamic Azad University, Arak, Iran.

[pkolyaee@iau.ac.ir](mailto:pkolyaee@iau.ac.ir)

<sup>2\*</sup>Professor, Department of Civil Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran.

[Hossein.samani@gmail.com](mailto:Hossein.samani@gmail.com)

<sup>3</sup>Associate Professor, Department of Civil Engineering, Arak Branch, Islamic Azad University, Arak, Iran.

[m-mirhoseini@iau.arak.ac.ir](mailto:m-mirhoseini@iau.arak.ac.ir)

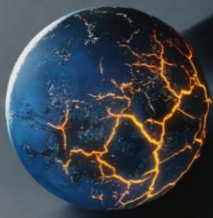
<sup>4</sup>Assistant Professor, Department of Civil Engineering, Arak Branch, Islamic Azad University, Arak, Iran.

[ezeighami@shahroodut.ac.ir](mailto:ezeighami@shahroodut.ac.ir)

### Abstract

One of the major types of water wastage in a distribution network is the leakage that occurs in pipes or other network components such as connections. Leaks in urban water supply networks impose many costs and workforce on governments and related organizations annually. If the cause of leakage is traced to the cracks and slits on various parts of a network, the amount of leakage from the location of the crack is said to be directly associated with the amount of pressure at that point. To trace leakage in water distribution networks, this study introduced a method based on hydraulics modeling and the inverse solution of flow equations to predict the location and amount of leakage in water distribution networks, by considering the values of the measured pressure in some network nodes. For this, a hydraulic model of the network under study was provided in the EPANET Hydraulic Analysis software, and the network analysis of various values and states of hypothetical leaks led to obtaining pressure values in the different nodes of the network via modeling water modeling under a steady state. Then, the artificial neural network, having been trained, was used to provide measured pressures in some network nodes at the testing hour as input data to the neural network to locate possible leaks in the water distribution network and to predict their approximate values. The findings were found to enjoy desirable accuracy.

**Keywords:** leakage, water distribution networks, hydraulic analysis, pressure, artificial neural network



## 1. Introduction

Leakage in a network occurs due to broken or cracked pipes, defective connections, and fittings, or due to failure to implement piping operations and connections. If the cause of leakage is traced to the network, cracks and slits in different components, and since the size of these cracks change with changing water pressure, i.e., as pressure increases, the cracks begin to widen and as pressure decreases, they get narrowed, one would say that the amount of leakage from the crack location is directly related to the amount of pressure at that point. Hence, any changes in water pressure may change the pace of water leakage from the network. Meanwhile, the IWA has presented the mechanism of estimating leakage over the course of past years using the experiences of various countries [1].

The lack of knowledge about leaks in network nodes causes difficulties when providing network hydraulic modeling because using demand values in the nodes would lead to unreal outflow values and pressure in the nodes if the leakage in the network is not taken into consideration. To solve this, attempts have been made to model leakage.

Araujo et al. (2003) and Burrows et al. (2007) have presented some equations for calculating leakage in nodes. According to these equations, node leaks have been calculated based on the number of subscribers in each node or in the form of a percentage of demands in each node [2,3].

The problems caused by leakage-detection processes include the lack of certain analytical equations to predict leaks in water distribution networks. Thus, by considering the capabilities of artificial neural networks, this article aimed to use this method to predict leakage using the results of the network hydraulic analysis model. The advantages of an artificial network include the approximation of any type of function, self-adaptability or updating, steadiness and easiness of work, and the absence of need to deeply perceive the relationships between the variables and the functions.

Damas et al. (2000) utilized an artificial network to control water supply systems. They used an artificial network to predict water needs to apply control and regulatory policies [6].

As well, Geem (2003) used artificial networks to develop a decision support system to assess pipes' status in the water distribution network [7].

Gibbs et al. (2002) used an artificial neural network to calculate the residual amount of chlorine in different points of the urban water distribution network [8].

Giustolisi (2008) presented a novel hydraulic simulation model under a steady state to determine pressure-dependent leakage, which offered a complete description of the network hydraulics, pressure-dependent demands and leakage at the pipe surface. This model was capable of simultaneously determining pressure-dependent demands and network leakage by aggregating all system consumption, including consumer consumption and leakage at pipe levels. The network simulation results were numerically tested using case studies in two small



*Received: 20-06-2024*

*Revised: 21-08-2024*

*Accepted: 06-09-2024*

and large networks, derived from real systems. The results confirmed the efficiency of the model [12].

Mashford et al. (2009) used SVM classifiers to predict leakage by analyzing measured pressure values and applied this method to predict the size and location of the leakage. They did a case study in the southeast of Melbourne [13].

Georgiou et al. (2014) suggested a leakage localization method based on pressure measurement and node correlation sensitivity analysis in a network calibrated in real networks. Since pressure sensors are quite expensive and there are only a few of them available, this study aimed to provide a sensitivity analysis and determine a network correlation to produce the best measurement nodes [14].

Ferrandez et al. (2015) provided a novel leakage localization method by employing a combination of models, pressure sensors, and classifiers [15]. The comparison of model predictions in hydraulic models in the network under study and the available real measurements obtained by sensors would yield pressure residual or differences. Finally, the recommended method was successfully tested by using the case study of a well-recognized Hanoi Network, consisting of a reservoir, 34 pipes and 31 connecting nodes [15].

Wachla et al. (2015) presented a drop-diameter approach-based method for the detection process. Accordingly, the leak localization by the SVM (support vector machine) in the proposed description models depends on the dynamics of observational variables and neuro-fuzzy state classifiers [16].

Ostapkowicz (2016) studied leakage detection in liquid transmission lines using simplified pressure analysis techniques by employing the minimum of standard and non-standard measure tools. This was made possible by two simple methods based on the pressure parameter and measurement of non-large liquid transmission pipelines subjected to steady-state conditions. The results confirmed the efficiency of the algorithms used for both methods [17].

Soldevila et al. (2017) presented a leakage localization technique in the water distribution network based on the Bayesian classification. Simulation results were provided for the case study of Hanoi. The results were also demonstrated for a real leakage scenario for the case study of Nova Icaria [18].

Baños (2018) developed a novel memetic algorithm to optimally design water distribution networks. To arrive at an accurate conclusion, five other methods, including neutral simulation, hybrid virtual simulation and tabu search, sparse search, genetic algorithm and binary linearly numerical programming, were adopted. The results from three water distribution networks indicated that the memetic algorithm outperformed other algorithms [19].

Paez et al. (2019) studied network theories and reliability to validate hybrid water distribution systems. This study used a novel method to produce hybrid distribution systems and to compare them with real-world systems, which would help create hybrid water distribution networks



based on the software provided by Mair et al. (2014). As for validation, five criteria were evaluated for network connectivity and system reliability (indicators of flexibility and network flexibility). In the end, an acceptable level of similarity between artificial and real sets in water distribution networks was obtained [20].

## 2. Methods and Procedure

### 2.1. Leakage modeling structure in the hydraulic analysis model

To predict leakage changes with controlling pressure, an equation is needed to show leakage at any point in the network as a function of pressure at that point. As for free leakage (outflow in the air), the theory of outflow-from-the-opening equation, assuming a constant cross-section, can be used as the following Equation 1:

$$(1) Q=KP^{0.5}$$

In Equation (1),  $Q$  is the outflow of the leakage of the opening on the pipe,  $P$  is the pressure at the leakage location and  $K$  is a constant coefficient, which depends on the shape and cross-section of the opening. As noted, this equation takes into consideration the amount of leakage proportional to the pressure square root at the location where the leakage has occurred. Some research has shown that the leakage outflow is significantly higher than what is seen by the opening outflow. Hence, Equation 2 is used as a general equation for leakage calculation:

$$(2) Q=CP^n$$

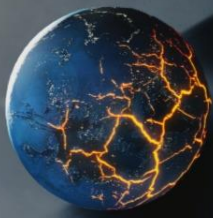
The  $C$  coefficient depends on each opening-specific characteristic. Various values have been proposed for the power of  $n$  in various sources and in different countries based on their distribution network status. Various methods have been introduced for assigning the leakage of hypothetical openings on the pipes to the nodes on the network and for calculating  $C$ . This article used this option as it applied the geysers feature of EPANET software.

Geysers are the installations connected to service lines which model the flow transmitting from a nuzzle or an opening that discharges into the atmosphere. The outflow transmitting from a geyser changes as a function of the pressure in the node. To model geysers in EPANET0.2 software, Equation 3 is used:

$$(3) Q_i=C_iP_i^N$$

Where  $Q_i$  is the  $i$ th geyser outflow,  $P_i$  is the pressure at the  $i$ th dispensing location,  $C_i$  is the flow intensity coefficient and  $N$  is the pressure power [4].

By controlling pressure at testing nodes and introducing them to the neural network following the training of the network, this article aimed to predict the nodes suspected of leakage and the approximate values of the leakage in each of those nodes. Hence, to perform the training process, the network requires a large number of pressure value sets in testing nodes and the leakage sets corresponding to each group of these pressures. To meet this goal, network modeling by hydraulic analysis software is required because providing this information under



ereal network is impossible. EPANET software was used to analyze the network hydraulics and calculate the pressure in the nodes. The following is suggested to model the network under study.

1. First, a general model of the network is developed by taking into consideration the information on the geometry, topographic, and geographical properties of the area under study.
2. Given the population covered by each node, the pressure-independent outflow or the subscribers' consumption could be determined at the time of the minimum night flow for all nodes on the network.
3. By assigning the pressure-independent outflow to each node, the hydraulic analysis model is implemented and the pressures of each and every node in the network are recorded.
4. Because consumption at any node includes two pressure-dependent consumption (leakage) and pressure-independent consumption (real consumption), additional consumption (as leakage outflow) can be added to the normal node consumption to simulate leakage in one node. By assigning the new outflow to the nodes in the network, the hydraulic analysis model is again implemented and the pressure of individual nodes is recorded. It should be noted that at this stage, each time the leakage occurs for one or several nodes, it is simultaneously simulated and the network analysis is made.

## 2.2. Artificial Neural Network (ANN) [14]

As stated, the intensity of the leakage from the network is directly related to the network pressure changes, and the amount of pressure at each point in the network is, by itself, a function of the total water amount (water consumed by the consumer and the leakage) collected from each node, the height of the node under study and the head in the reservoir. Thus, to properly manage the network, the user is required to take into consideration the various states of possible network changes to implement the hydraulic analysis model and to use its results in planning. As for the urban water distribution network, this article trained an artificial network model to help calculate the level of pressure and then the level of possible leakage in the network nodes, after entering the data of water consumption and its height in each node and of the water head in the reservoir. The artificial network model uses the EPANET2.0 hydraulic analysis model results to train the hypothetical leakage changes in the network nodes. This article used the multi-layered perceptron artificial neural network. The artificial neural network with a multi-layer perceptron structure consists of several layers (usually of three layers), with each layer made of some processing units called neurons. An artificial neuron can be a non-linear mathematical function; as a result, an artificial network, made of a group of these neurons, can also be a completely nonlinear and complex system. In a neural network, each neuron works independently, and the overall behavior of the network will be the result of the local behaviors of various neurons. This feature causes the local errors not to have a significant effect on the output. In other words, the neurons will, in a cooperating process, try to correct each other, thus increasing the system's durability. The parameters related to a neuron are:  $P$  representing the value of the input to the neuron,  $W$  the weight of each input,  $b$  the constant



Received: 20-06-2024

Revised: 21-08-2024

Accepted: 06-09-2024

coefficient of a unit of weight (the weight of a single input), usually called a bias, which, if added, will increase the network's flexibility,  $f$  the driving function and  $a$  the output of the neuron. The relationship between the neuron inputs and units can be represented by the following Equation 4:

$$(4) a = f(wp+b)$$

The sigmoid distribution function structure, usually used in the artificial neural network, is as Equation 5 below:

$$(5) Y_j = g(s_j) = 1/[1 + \exp(-s_j)]$$

Where  $s_j$  is a real number but the value of  $Y_j$  varies from zero to one.

Artificial network learning is the setting of network parameters, including  $w$  and  $b$  weight values. A network enjoys learning capability that manages to be efficient for new conditions if trained for a special situation, as a small change occurs under the environmental conditions of the network. The relationship between the output of each neuron (before entering the driving function) and the input values is as Equation 6:

$$(6) n_i = \sum_{j=1}^R (p_j w_{i,j} + b_i)$$

Where  $R$  is the total number of inputs and  $i$  is the number of neurons in each layer.

As noted previously, the inputs to the artificial network in the present model are the pressures in the nodes where pressures have been measured and are called testing nodes, with the network outputs called the leakage in all the nodes. The following concerns with an example and predict a leak using the proposed method.

### 3. Example

This example is aimed at detecting leakage in a water distribution network consisting of 7 pipes and 9 nodes. Figure (9-6) exhibits a model of node and pipe arrays. This example assumes that Nodes 2, 3, 5 and 6 are the nodes accessible to the network and the pressure at these points has only been measured at the testing hour.

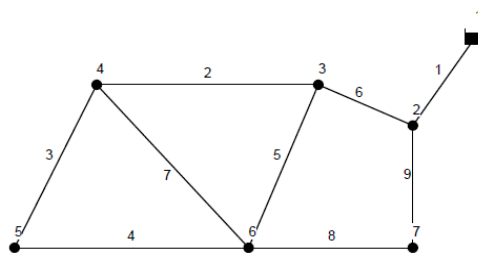


Figure 1: Geometric array of the network under study



### 3.1. Network's overall information

The data of the network under study are given in Tables 1 and 2.

**Table 1: Data of the node specifications in the network under study**

Node No.	Height (m)	Node No.	Height (m)
2	46.2	5	44.3
3	47.32	6	43.7
4	40.78	7	46.3
The piezometric level of Node 1 (reservoir)=78 m			

**Table 2: Data of pipes in the network under study**

Pipe No.	Source node	Target node	Length (m)	Diameter (mm)	Hazen-Williams coefficient
1	1	2	100	200	120
2	3	4	100	100	120
3	4	5	77	51	120
4	5	6	70	76	120
5	6	3	70	51	120
6	2	3	60	100	120
7	4	6	80	51	20
8	6	7	67	76	120
9	7	2	65	100	120

### 3.2. Data of normal network consumption at the testing hour

**Step 2:** Normal consumption values in each of the network nodes at the testing hour are given in the following table.



**Table 3: Consumption of the nodes of the network under study at the testing hour**

Node No.	Consumption per testing hour (l/s)
2	2.87
3	2.85
4	0.94
5	1.3
6	3.03
7	2.1

### 3.3. Pressures measured in testing nodes

Table 4 shows pressure values in the testing nodes of the network under study

**Table 4: Pressure in the testing nodes of the network under study at the testing hour**

Node No.	Pressure (per water meter)
2	31.61
3	29.73
5	31.17
6	32.48

### 3.4. Providing data required for training the neural network

The data required to train the network, consisting of 1398 data series, were provided by the various states of the EPANET software's simulation of the network. Pressure values in all network nodes were given to the network as the input vector and the corresponding leakage values as the target vector. To avoid the problem of overfitting from occurring, out of all data, 60% were used as training sets, 30% for validation, and 10% for testing the network.

## 4. Neural network details

Training data were provided in the prior stage. The next step was to develop a network, and in this case, a three-layered perceptron network made of 13 neurons in the first hidden layer and



Received: 20-06-2024

Revised: 21-08-2024

Accepted: 06-09-2024

15 neurons in the second hidden layer. The number of neurons in the output layers equals the number of components in the target vector, and in this case, 6. The used activity functions are *tansig* function for both hidden layers and the *purelin* function for the output layer, with the *trainlm* function and the backpropagation algorithm used to train the network. All these values and functions were selected out of all tested states after many times of developing a network of different architecture and functions of activity as well as different training and selecting the most optimal state. The criterion for selection was the least mean square errors (mean squared errors (MSE)), with the highest correlation coefficient  $r$  being between the output and the target vectors. The following gives a comparison of various architecture and their outputs, as well as the leakage coefficients obtained from the neural network, assuming the known pressure values.

**Table 5: Comparison of the efficiency of some neural networks studied by the example**

Network structure	Training algorithm	Correlation coefficient	MSE error
4-17-19-6	Trainlm	0.95962	0.0105
4-15-17-6	Trainlm	0.94712	0.0133
4-11-13-6	Trainlm	0.95081	0.0125
4-13-17-6	Trainlm	0.94603	0.0127
4-13-15-6	Trainlm	0.97208	0.00707
4-17-19-6	Trainlm	0.94374	0.0142
4-7-9-6	Trainlm	0.9213	0.0195
4-6-7-6	Trainlm	0.92158	0.0195
4-5-8-6	Trainlm	0.89915	0.025
4-7-11-6	Trainlm	0.92403	0.0193
4-13-6	Trainlm	0.9332	0.0164
4-16-6	Trainlm	0.91711	0.0206
4-20-6	Trainlm	0.91795	0.0198
4-17-6	trainscg	0.72069	0.0608
4-19-6	trainscg	0.72069	0.0608
4-15-17-6	trainscg	0.74876	0.0564



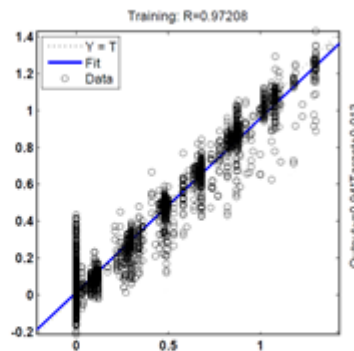
Received: 20-06-2024

Revised: 21-08-2024

Accepted: 06-09-2024

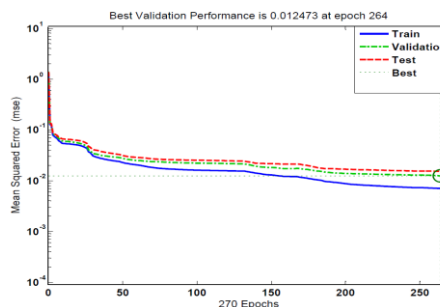
4-5-7-6	trainscg	0.73369	0.0583
4-5-7-6	trainrp	0.64037	0.0767
4-15-17-6	trainrp	0.75278	0.0563
4-17-19-6	trainrp	0.74743	0.0591
4-17-19-6	traingdx	0.2589	0.121

Coefficients of the leakage outflowing from the neural network



Coefficient of the leakage under study

**Figure 2: Comparison of modelled leakage coefficients and neural network outputs and the determination of the regression coefficient of the example under study**



**Figure 3: Changes in the network’s MSE errors after each time the iteration of the training process of the example**

Figure 3 exhibits the training of the network using the trainlm training algorithm. As noted, the MSE error of the network is seen decreasingly gradually, as the network training stops with the message of the Validation Stop, after 270 iterations. This indicates the increasing number of errors in the validation set, with the values of the weights and biases adjusting to the time when this error has been at a minimum rate. The closeness of the errors in the training set, the tests, and the validation suggest the model’s desirable results. The investigation of the C coefficients obtained from the network nodes helps to detect the nodes, and consequently, the pipes suspected of leakage. The comparison of the leakage coefficients derived from the neural



network and the real leakage coefficients confirms the efficiency of the neural networks in solving such problems.

**Table 6: Coefficients of the leakage modeled and coefficients derived from the neural network of the example at the testing hour**

Node No.	Coefficient of leakage derived from the neural network	Coefficient of leakage modelled by the software
2	0.0180	0
3	-0.0368	0
4	0.0124	0
5	0.5969	0.6
6	0.0154	0
7	0.0443	0

**Table 7: Nodes and pipes suspected of leakage**

Leaked nodes	Pipes suspected of leakage
5	3 & 4

## 5. Discussion and Conclusion

Due to the lack of analytical relations to predict the various states of leakage in water distribution networks and the inaccessibility of information in this regard, especially given the fact that water pipes are buried underground, this article aimed to use artificial neural networks to predict leakage. In the training stage, neural networks require many training data series with predetermined results. It is impossible, however, to provide a large volume of data for the real states of leakage in the network under study; for this, to meet this problem, the modeling and analysis of network hydraulics using the EPANET software were used. Because the information provided to train the neural network was derived from the hydraulic analysis of the network using the software, the accuracy of the data will change depending on the model's capability of simulating real conditions. It is clear that under no circumstances will this simulation be a hundred percent perfect and accurate, with the data and consequently the neural network results being subjected to some percentages of errors, even when the network performs



Received: 20-06-2024

Revised: 21-08-2024

Accepted: 06-09-2024

well. Since the goal of the leakage-detection process is to prioritize regions for inspection and restoration and leakage-repair operations, these percentages of errors will not have much effect on management decision-making.

The advantages of the proposed method include its simplicity and the lack of need for complex tests or any special or expensive means. To investigate leaks in a special network, network modeling and the development of a trained neural network could predict leaks for the long term, as some minor changes may be made to the programs under study, if necessary. Making these changes requires less time and less money. The accuracy of the results depends on the specifications of the network under study and the accuracy of the neural network accuracy. In sum, the reliability of the results will be enhanced if the accuracy of these characteristics is increased.

## References

- [1] Lambert, A.O., McKenzie, R.D. (2002). "Practical Experience in Using The Infrastructure Index." International Water Data Comparisons Ltd, UK, Global water resources Ltd, Republic of south Africa.
- [2] Araujo, L.S., Coelho, S.T., and Ramos, H.M. (2003). "Estimation of Distributed Pressure-Dependent Leakage and Consumer Demand in Water Supply Networks." advances in water supply management- maksimovic, Butler & Memon (eds), Swets & Zeitlinger, Lisses, ISBN 9058096084.
- [3] Burrows, R., Mulried, G. and Hayuti, M. (2003). "Introduction of a Fully Dynamic Representation of leakage into Network Modeling Studies EPANET." advances in water supply management- maksimovic, Butler & Memon, eds., Swets & Zeitlinger, Lisses, ISBN 9058096084.
- [4] Damas, M.; Salmeron, M.; Ortega, J. (2000). "ANNs and GAs for Predictive Controlling of Water Supply Networks", Proceedings of the International Joint Conference on Neural Networks, 365-370.
- [5] Geem, Z.W. (2003). "Window-Based Decision Support System for the Water Pipe Condition Assessment using Artificial Neural Network." World Water and Environmental Resources Congress, 2027-2032.
- [6] Gibbs, M.S., Morgan, N., Maier, H.R., Dandy, G.C., Holmes, M., and J.B. Nixon. (2002). "Use of Artificial Neural Networks for Modelling Chlorine Residuals in Water Distribution Systems".
- [7] Karamouz, M., M.Tabari, M.R., Karachian, R. (2004). "Conjunctive use of surface and ground water resources: Application of genetic algorithms and neural networks." Conference on Critical transition in water and environmental resource management, Salt Lake city.
- [8] Karamouz, M., Razavi, S., Araghinejad, SH. (2004). "Application of artificial neural networks in flood estimation." Proceeding of 4th international conference on decision making in urban & civil engineering, Porto, Portugal.
- [9] O. Giustolisi, D. Savic, and Z. Kapelan, "Pressure-Driven Demand and Leakage Simulation for Water Distribution Networks," *J. Hydraul. Eng.*, vol. 134, no. 5, pp. 626–635, 2008.



Received: 20-06-2024

Revised: 21-08-2024

Accepted: 06-09-2024

- [10] J. Mashford, D. De Silva, and D. M. S. Burn, "An Approach to Leak Detection in Pipe Networks Using Analysis of Monitored Pressure Values by Support Vector Machine," in *Proc. Netw. Syst. Secur.*, 2009, no. 1, pp. 534–539.
- [11] C. M. Giorgio Bort, M. Righetti, and P. Bertola, "Methodology for leakage isolation using pressure sensitivity and correlation analysis in water distribution systems," in *Procedia Engineering*, 2014, vol. 89, pp. 1561–1568.
- [12] S. Lise Ferrandez-Gamot\*, Pierre Busson\*, Joaquim Blesa\*\*, "Leak Localization in Water Distribution Networks using Pressure Residuals and Classifiers," in *IFAC*, 2015, pp. 220–225.
- [13] W. Wachla, D and Przystalka, P and Moczulski, "A method of leakage location in water distribution networks using artificial neuro-fuzzy system," in *IFAC-PapersOnLine*, 2015, vol. 48, no. 21, pp. 1216–1223.
- [14] P. Ostapkowicz, "Leak detection in liquid transmission pipelines using simplified pressure analysis techniques employing a minimum of standard and non-standard measuring devices," *Eng. Struct.*, vol. 113, pp. 194–205, 2016.
- [15] Soldevila, A., Fernandez-Canti, R. M., Blesa, J., Tornil-Sin, S., & Puig, V. (2017). Leak localization in water distribution networks using Bayesian classifiers. *Journal of Process Control*, 55, 1-9.
- [16] R. Baños, C. Gil, J. Reça, F.G. Montoy, 2018, A memetic algorithm applied to the design of water distribution networks, *Applied Soft Computing*, Volume 10, Issue 1, January 2010, Pages 261-266
- [17] Paez, D., & Filion, Y. (2019). Use of network theory and reliability indexes for the validation of synthetic water distribution systems case studies. *Sustainable Energy Technologies and Assessments*, 24, 2-7.