TEST Engineering & Management

January-February 2020
ISSN: 0193-4120 Page No. 8355 - 8362

Urban Water Distribution Network Failure Prediction using Artificial Intelligence

Vinayak Patki, Professor, Dept. Civil Engg. N.K. Orchid College of Engineering and Technology, Solapur
Shrikant Jahagirdar, Professor, Dept. Civil Engg. N.K. Orchid College of Engineering and Technology, Solapur
Shriniwas Metan, Professor, Dept. Mechanical Engg. N.K. Orchid College of Engineering and Technology, Solapur
Shalvi Deshmukh, Student, Dept. Civil Engg. N.K. Orchid College of Engineering and Technology, Solapur

Article Info
Volume 82
Page Number: 8355 - 8362
Publication Issue:
January-February 2020

Article History
Article Received: *5 April 2019*
Revised: *18 Jun 2019*
Accepted: *24 October 2019*
Publication: *07 February 2020*

Abstract:
In this study Artificial Neural Network (ANN) and Adaptive Neuro Fuzzy Inference Systems (ANFIS) have been used to predict the failure trend of pipe network and to access the present condition of water distribution system. To predict the number of failures in pipelines different methods can be used. Data driven modeling is the most recent method adopted in different fields for prediction where historical data are available. Soft computing techniques like ANN and ANFIS have the potential of exploiting the tolerance for imprecision, uncertainty and partial truth to achieve tractability, robustness, low solution cost and better support with reality. This approach is tested and verified in a real life water distribution system in Trivandrum city. The case study demonstrates the entire process from data aggregation to model development. The work included the use of two ANN networks namely cascade and feed forward back propagation network for the prediction of water pipelines failure. Seven indicators were identified as input parameters in the water main failure. They include age of pipe, number of previous failures, pipe length, diameter, thickness, material and demand. The performance of the models is evaluated by using coefficient of correlation and mean absolute error. The study reveals that the predicted results will help the authorities to take decision regarding the repair and replacement of pipes in the distribution system.

Keywords: *Feed forward back propagation network, Cascade back propagation network, ANFIS, ANN, water distribution system.*

I. Introduction

Water distribution systems can be described as a network of pipes, valves and pumps which transport finished water to consumers. Even with the technological developments every year many cases are being reported in municipal distribution system. The pipe deterioration process needs to be understood completely to provide reliable and sustainable municipal distribution system at lowest cost. Now the mode of working strategy adopted by operators is just to respond to emergency breaks. Now a days due to shortage of funds and restricted resources with Municipal Corporation the underground infrastructure management using predictive model has become need of the day to optimize the maintenance cost. Well scheduled maintenance will certainly reduces the maintenance cost against unscheduled repairs. This can be achieved by predicting accurately the immediate need of repair work. Predictive modeling includes a collection of technologies that can be used to determine the likelihood of failure or failure rate for a pipe entity. Pipe failure is a part of infrastructure deterioration process. The lengths of time intervals and the mechanism of the failure are case dependent and extremely difficult to predict [1] classified the deterioration of pipes into two categories. The first is structural deterioration which happens when its capacity to withstand external stresses diminishes. The second is hydraulic capacity deterioration

January-February 2020
ISSN: 0193-4120 Page No. 8355 - 8362

resulting in diminished water carrying power of the pipeline system. This happens due to degradation of water quality and reduced structural resiliency in-case of severe internal corrosion. The area of failure and asset management in water supply systems has been in focus of both researchers and practitioners in water industry for the last. Depending on the timing of the failure management activities with respect to the failure itself, two types of pipe failure management strategies can be defined. One is proactive failure management and another is reactive failure management. When the pipe replacement decisions are made prior to failure to prevent the failure it is proactive failure management while in reactive failure management repair is performed only after the failure has occurred. In Kerala due to low cost of water and fairly high cost of proactive failure management technique, reactive failure management technique is used.

II. Literature Review and Objectives of The Study

A. Literature Review

"Integrated GIS based management of water distribution network [2] presented a framework to manage urban water distribution networks based on both analytical and numerical modeling techniques. The work is based on data collected in Newyork city .The study is for finding a relationship between number of breaks and risk factor. It is done by means of Artificial Neural Network. ANN also helped to find a rough estimate of life cycle for each of the individual pipes in the network. According to ANN output the most influencing parameters responsible for failures are the number of previously observed breaks, material type, length and diameter of each pipe. Using fuzzy logic prioritization is done. GIS is used to represent the results in a convenient manner so that pipe managers can take a suitable decision. Sometimes if the breaks are in small quantity it will be better to go for rectification than replacement. But such an option is not available. [3] describes the development of NN for prediction of pipeline failure using a large database which is neither complete nor fully accurate. They have compared the results they obtained from ANN with shifted time power model (STPM) and shifted time exponential model (STEM). The study revealed that STEM gives very poor correlation (Cc-0.0097) whereas STPM showed a high coefficient of Correlation (Cc- 0.837) .The multilayer perceptron ANN architecture with input layer (6 neurons), two hidden layers (8 neurons) and output layer(one neuron) is being used. Conjugate gradient algorithm was used for error minimization. It was observed that the prediction by this model improves by 19% compared to STPM if year of construction and pipe age were used as input variables. ANN, ANFIS and multivariate regression approach was compared to model pipe failure rate [4]. The pipe data collected include 337 pipelines made up of asbestos which has diameters ranging from 80 to 300mm. The inputs to the model are length, diameter, age, pressure and depth of burial of pipe and the output is the failure rate. The developed ANN model was for 80% of the available data set and testing was carried on remaining dataset. The study was carried using tansigmoidal and linear activation function the study revealed that ANN algorithm with 2 hidden layers and with 10 neurons in the hidden layer gives better prediction. Then further ANN was compared with ANFIS. The ANFIS model was developed using Gaussion, Bell, triangular and trapezoidal shaped membership functions. The result of the ANFIS model showed unrealistic values and very sharp variations for break rates when diameters was increased from 80 to 200 mm. Whereas for larger length (200-500m) pipes the ANFIS model produces almost the same break rates. The study revealed that ANN and MLR methods produces lower break rates than the ANFIS results. When study was conducted using water pressure to find pipe break it showed that ANN predictions are more realistic compared to ANFIS and MLR predictions.

The Effectiveness of ANN and ANFIS model for prediction of pipe break was carried [5] for Benghazi city. The study was carried using diameter, length, thickness, material, age, soil, location, roughness of pipe, depth of installation and hydraulic loss as input variables and pipe outbreak as out variable. Total 410 pipe lines were taken for training multilayered Feed forward back propagation algorithm. In this study two hidden layer were used. In the hidden layer neurons were varied from 2-5.the neurons. Error between observed output and predicted output is compared. Other performance indices are not considered. The study revealed that ANN model outperforms ANFIS model.

B. OBJECTIVE OF THE STUDY

1. To compare ANN and ANFIS model for prediction of trend in pipeline failure.

2. To use Peroorkada water distribution network to prioritize the parameters to predict pipeline failure.

III. MATERIALS AND METHODS

A. DATA COLLECTION

The water distribution system of Kowdiar and Vellayambalam, Kerala was taken for study. In the study ares diameter of pipe varies from 400-700mm. In this study length of pipe, pipe age, pipe material, pipe diameter, location of pipe failure, number of failures for ten years(from 2000 to 2010) The figure 1 shows the Google image of the study area

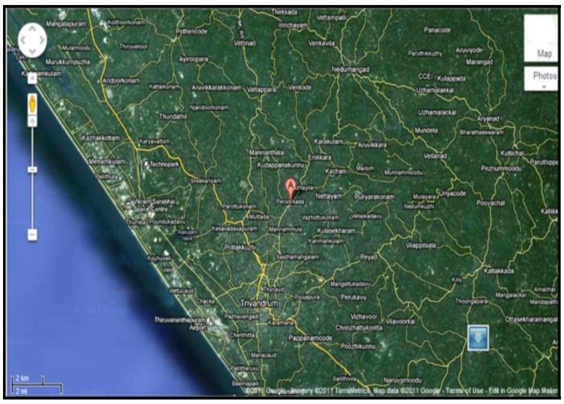

Fig.1 Study Area

B. STUDY AREA

Thiruvananthapuram is the southernmost district of the State of Kerala. It is bounded by Quilon District on the North, the Arabian Sea on the west, Tirunelveli and Kanyakumari Districts of Tamilnadu State on the East and South respectively. It lies between 8^0 17' & 8^0 47'North Latitudes and 76^0 41' & 77^0 16' East Longitudes. The district measures 57 km on the East-West axis and 60 km on the North-South axis. Thiruvananthapuram is the capital City of Kerala State, besides being the head-quarters of the Thiruvananthapuram District. Thiruvananthapuram city has a stable source of water supply. Karamana River is a perennial river in Kerala having large catchment area. This is surface water source fot Thiruvananthapuram city. This is a perennial river with a very large catchment area. For the development of vital storage, two dams are constructed on the upstream side of river. The major dam is at Peppara, 45 km north-east of Thiruvananthapuram City. The location of dam site is in the reserve forest and the reservoir formed has a natural backdrop of high mountains. The storage capacity is 70 million cubic metres. The other dam is at Aruvikkara, 25 Km downstream of the first reservoir.Its capacity is 2 million cubic metres. The entire head works of the Thiruvananthapuram Water Supply Scheme is located near the Aruvikkara dam. Thiruvananthapuram water supply scheme was started in1933 and is one of the oldest water supply schemes. It is fully possessed and presented by water authority of Kerala. The water supply line of Thiruvananthapuram is being laid for 183.7 sq .km. The service area of water supply scheme is divided into five zones viz. Low Level Zone, Observatory Zone, Manvila Zone, Peroorkada Zone and PTP Nagar Zone. The water treatment is installed at Aruvikkara, PTP Nagar and Vellayambalam. The average water consumption in Thiruvananthapuram is 175 lpcd. The total length of distribution system is 2280 km. This distribution system is planned to supply 24*7 water supply. Out of these zones, present study is carried out on Peroorkada zone. Peroorkada zone has eleven zones. The population of this zone is 1, 37,714. This zone has 13700 domestic water supply connections, 6 industrial and 180 commercial connections. The Peroorkada zone network has 16 loops with 99 nodes and 114 pipelines. The location map of this zone is shown in fig.2 and distribution layout for this zone is shown in fig.3 . Each node requires minimum head of 8m.

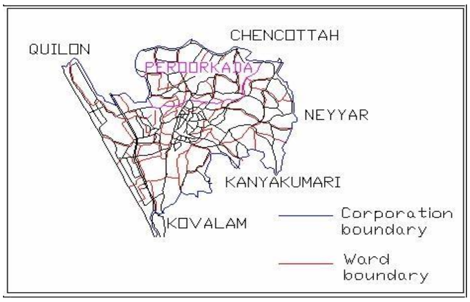

Fig. 2: Location map of Peroorkada zone

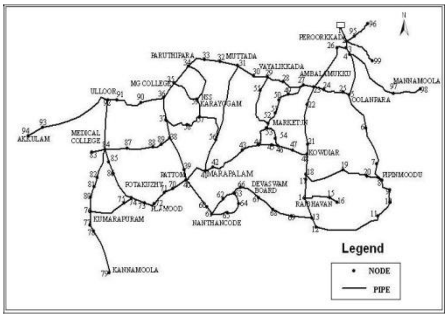

Fig. 3: Water Distribution Network

C. METHODOLOGY

The figure 4.4 shows the various steps involved in the present study

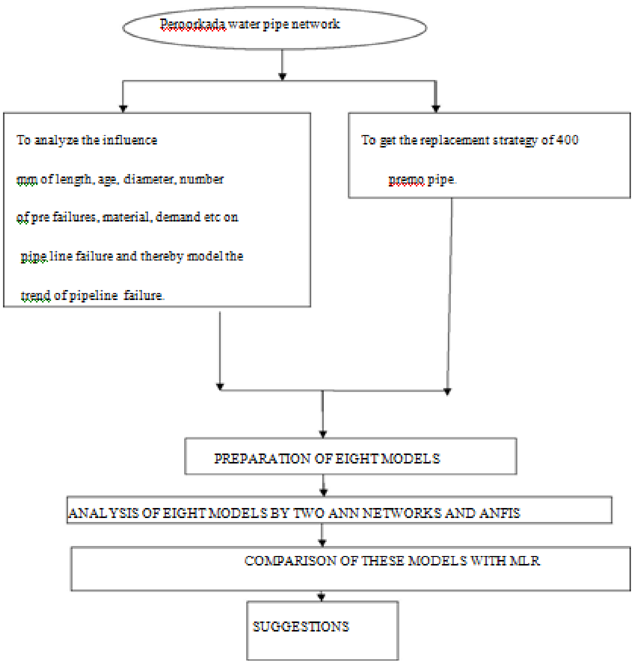

Fig 4: Project steps

To achieve the main purposes of the study i.e. to model the inclination of letdown after analyzing the factors responsible for failure such as length, age, diameter, demand, number of previous letdowns etc and to prioritize the network based on renewal the variation of each of the input factors with failure is considered (fig..5-fig.7).

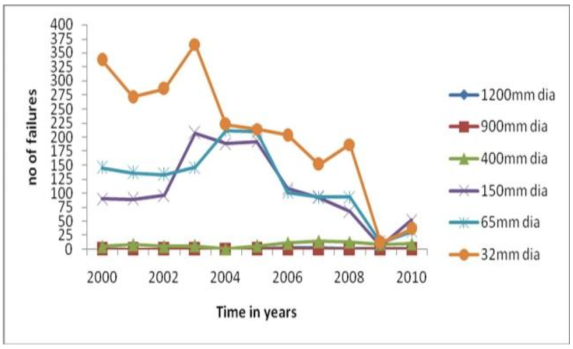

Fig.5: Variation of failures in different diameter with time

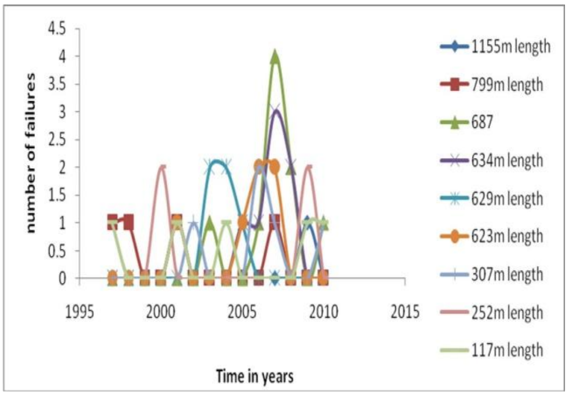

Fig. 6: Variation of failures in different length with year

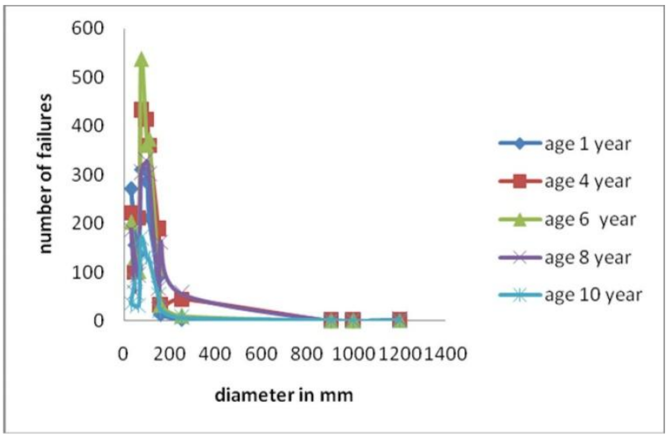

Fig .7 Variation of failures with age in different diameter pipes

TABLE I
The Models Used in the Study

A) ANN Modeling

Details of Model	Inputs Variables	Output Variable
M1	Past data of pipe Failures	Present status of pipe Failure
M2	Past data of pipe Failures and pipe diameter	Present status of pipe Failure
M3	Past data of pipe Failures, pipe age	Present status of pipe Failure
M4	Diameter of pipe,Past data of pipe Failures, pipe age	Present status of pipe Failure
M5	Past data of pipe Failures, pipe age Diameter, pipe material	Present status of pipe Failure
M6	Length, Number of failures up to 2009	Pipe Failures in 2010
M7	Pipe failure data up to 2009	Pipe Failures in 2010
M8	Pipe failure data up to 2009 ,Length of pipe line, water demand	Pipe Failures in 2010

The variation of each of the input factors with failure is nonlinear and it doesn't follow a regular pattern. So use of ANN and ANFIS is beneficial, as these soft computing techniques can model complex input output relationship faster and with better accuracy. Prior to analysis the first step in methodology is the preparation of models. The important factors responsible for failures obtained from preliminary analysis of data, enquiry with water authority officials and from literature review are age, length, number of pre failures, material and flow through pipe. Even though soil and pressure are two important factors affecting failures in water pipelines, in this particular area as soil is more or less same and pressure factor doesn't arise these two parameters are not considered. By varying the model inputs eight models are prepared. The first five models are for all diameter pipes in the network and the last three are especially for 400mm diameter pipes. Before preparing the models, the correlation between different parameters affecting failures with failures are looked into and found pre failures are affecting failures to the most. So in all models pre failures are considered as one of the inputs. In the first five models failures in the preceding year is considered and in the rest accumulated failures are considered since for diameter pipes pipe failures only data from 2000 to 2010 is available but for 400mm diameter pipes the data from installation period is available. All together eight models are prepared and is shown in table 1. After preparation of models, the next step is analysis of models by ANN networks and ANFIS.

Two ANN networks have been used in the study. Feed Forward Back Propagation (FFBP) algorithm and Cascade Forward Back Propagation (CFBP) algorithm are compared The modeling of two ANNs are coded using Matlab 2014a.

a. Preparation of dataset-For the first five models a total of 140 datasets are available and for the last three models a dataset of 117 values are available. Before giving to the model the dataset is normalized. In the case of model analyzed by giving material as input, the most failure prone material will be given a value equal to one and the others are given values considerably. The order of failure is as follows- AC>CI>GI>PVC>PSC

January-February 2020
ISSN: 0193-4120 Page No. 8355 - 8362

b. Training of ANN Models

In this study, three layered Feed Forward Back

M_5	0.925	14.51	0.898	17.426	0.859	36.07
M_6	0.987	0.23	0.977	0.247	0.88	0.62
M_7	0.951	0.282	0.951	0.282	0.95	0.17
M_8	0.99	0.154	0.99	0.064	0.983	0.179

Propagation (FFBP) algorithm and Cascade Forward Back Propagation (CFBP) algorithm are compared for prediction of pipe failure in the municipal distribution system. The training was carried using the fastest training Levenberg Marquardt algorithm. Tansigmoidal and linear activation functions have been used for training both the algorithms. The Mean Square Error (MSE) and Coefficient of Correlation (Cc) are used to check the performance of trained ANN models. The optimal numbers of neurons in the hidden layer is found out by conducting different trials.

c. Testing of Models

After training, the developed ANN models have been used to predict the pipe failure trends for various input variables by using testing sets. The predicted values are then compared with the observed values using various statistical performance criterions.

B) ANFIS modeling

The modeling of two artificial neural networks is coded using Matlab 2010a. Before giving data to ANFIS dataset is normalized as well as it is separated into testing and training. The training dataset is also kept under two files. One for keeping input training values and other output training values. The learning algorithm used is hybrid learning algorithm. Different trials are done by changing the number of membership functions, type of membership functions and number of cycles.

IV. Result and Discussion

A) Comparison of two ANN networks

From Table 2 and Table 3 it can be observed that CFBP algorithm outperforms FFBP algorithm for majority of developed models during testing even though the performance of FFBP and CFBP are

Dataset	**CFBP**		**FFBP**		**ANFIS**	
Models	**Cc**	**MAE**	**Cc**	**MAE**	**Cc**	**MAE**
M_1	0.93	14.5	0.906	17.861	0.93	16.4
M_2	0.929	13.162	0.952	15.389	0.93	13.16
M_3	0.923	18.178	0.907	22.975	0.91	37.46
M_4	0.931	14.808	0.924	35.148	0.86	49.49

almost similar in training.

TABLE II

Performance CFBP and FFBP algorithms in Training

B) Comparison of two ANN networks and ANFIS

If overall performance is considered CFBP ANN network and FFBP ANN network outfits ANFIS. Even though the performance of ANFIS is not better than two ANN networks, the ability of ANFIS to capture the learning it has obtained from training and apply it in unseen input parameters to find the desired output is high. This can be seen in Table 4 and Table 5.

TABLE III

Performance CFBP and FFBP algorithms in Testing

Dataset	**CFBP**		**FFBP**		**ANFIS**	
Models	**Cc**	**MAE**	**Cc**	**MAE**	**Cc**	**MAE**
M_1	0.92	36.8	0.911	32.082	0.92	34.9
M_2	0.911	35.37	0.911	35.557	0.94	30.9
M_3	0.917	36.67	0.938	29.967	0.98	18.31
M_4	0.916	34.877	0.941	22.22	0.96	22.8
M_5	0.894	30.678	0.898	33.388	0.999	3.422
M_6	0.99	0.107	0.987	0.114	0.99	0.05
M_7	0.983	0.162	0.988	0.142	0.99	0.12
M_8	0.999	0.086	0.983	0.15	0.998	0.051

TABLE IV

Performance of CFBP, FFBP algorithms and ANFIS in Training

Dataset	**CFBP**		**FFBP**	
Models	**CC**	**MAE**	**CC**	**MAE**
M1	0.92	36.8	0.911	32.082

TEST Engineering & Management

January-February 2020
ISSN: 0193-4120 Page No. 8355 - 8362

M2	0.911	35.37	0.911	35.557
M3	0.917	36.67	0.938	29.967
M4	0.916	34.877	0.941	22.22
M5	0.894	30.678	0.898	33.388
M6	0.99	0.107	0.987	0.114
M7	0.983	0.162	0.988	0.142
M8	0.999	0.086	0.983	0.15

TABLE V

Performance of CFBP, FFBP algorithms and ANFIS in Testing

Dataset(D)	**CFBP**		**FFBP**	
Models	**CC**	**MAE**	**CC**	**MAE**
M_1	0.93	14.5	0.906	17.8617
M_2	0.929	13.162	0.952	15.389
M_3	0.923	18.178	0.907	22.975
M_4	0.931	14.808	0.924	35.148
M_5	0.925	14.51	0.898	17.426
M_6	0.987	0.23	0.977	0.247
M_7	0.951	0.282	0.951	0.282
M_8	0.99	0.154	0.99	0.064

C) Prioritization of network based on renewal

Prioritization of network is done only for network containing 400mm diameter pipes since complete network details' pertaining to this diameter is available. Table 6 shows the order of replacement.

TABLE VI

Sequential Order for Pipe replacement based on the Prediction

Link	**Length(m)**	**Order of Prioratisation**
4-5	687	1
6-7	634	1
5-6	623	2
30-31	322	2
3-4	117	3
22-23	307	3
40-41	448	3
42-43	629	3
71-72	252	3

V. CONCLUSION

The analysis of the data reveals the following conclusions.

1. Among the five input parameters, the number of previous failures and diameter are the most important parameters affecting failure. By increasing the diameter of pipes failure problems can be reduced to a bit. This is understood from the preliminary data analysis.
2. Along with diameter, number of previous failures in last year, age when material is also considered the mean absolute error value decreases. So material change is also an adoptable factor in reducing failures. In the case of material, nowadays PVC pipes are used since it is less expensive. But this is also prone to leakage. MDPE (medium density polyethylene)pipes are a better option to prevent leakage to a greater extend. MDPE pipes came in coils and could be laid in a virtually joint free manner. This they reasoned would drastically reduce the chances of such pipes springing leaks. Since these pipes are flexible, they would be less prone to bursting if exposed to sudden loads.
3. By using ANN and ANFIS, number of failures in each pipeline in future can be known in almost accurate manner. This helps in prioritization of network for replacement based on failure. This can improve the efficiency in timely replacement of pipes and early acquisition of funds for replacement. In majority of cases the delay in repairing leakage is caused due to insufficient funds.
4. Normalized data input to ANN networks outfits raw data given to the network.
5. Different length of testing dataset is considered for each of the models. In majority of models 10% dataset is performing better than others. It is always better to consider majority of data in training rather than testing.

January-February 2020
ISSN: 0193-4120 Page No. 8355 - 8362

6. Even though performance of ANFIS in training is not better than ANN networks it outstands ANN networks in testing.

7. The previous literatures have explored the feed forward network in predicting failures in network. The cascade feed forward back propagation network is a better option than feed forward back propagation network since it gives better correlation coefficient and less mean absolute error value compared to feed forward network

VI. Reference

1. Kleiner Y. and BalvantRajani (2002). "Forecasting variations and trends in water main breaks". Journal of Infrastructure Systems, Vol. 8, 122-131
2. Symeon Christodoulou et al. (2006). "Integrated GIS based management of water distribution networks". Joint International Conference On Computing And Decision
3. Making In Civil And Building Engineering, Montreal, Canada, 858-865.
4. D. Achim,F. Ghotb, K J Mcmanus(2007). "Prediction of Water Pipe Asset Life Using Neural Networks".Journal of Infrastructure Systems, 13(1), 26–30.
5. M. Tabesh, J. Soltani,R. Farmani, D. A. Savic(2009). Assessing Pipe Failure Rate and Mechanical Reliability of Water Distribution Networks Using Data Driven Modelling. J. of Hydroinformatics, 11(1), 1-17
6. Abdelwahab M. Bubtiena et al. (2011). "Application of Artificial Neural Networks in Modeling Water Networks".7th International Colloquium on Signal Processing and its Applications, 50-57.