World Applied Sciences Journal 18 (4): 528-537, 2012
ISSN 1818-4952

DOI: 10.5829/idosi.wasj.2012.18.04.1246

Estimating Water Losses in Water Distribution Networks Using a Hybrid of GA and Neuro-Fuzzy Models

[1]*Amir Jalalkamali* and [2]*Mehdi Eftekhari*

[1]Department of Water Engineering, Faculty of Engineering,
Islamic Azad University, Kerman Branch, Kerman, Iran
[2]Department of Computer Engineering, Shahid Bahonar University of Kerman, Kerman, Iran

Abstract: Modeling Leakage rate in water supply networks is important due to some problems namely: quality of service, the cost of system expansion and wasting energy resources. Monitoring the pressure in certain parts of network and then finding a relation between pressure changes and leakage rate is a quick way for detecting leakage. This relation is nonlinear and complex for modeling and there exist some problems related to mathematically modeling of it by hydraulic fundamentals. Thus, regarding to the abilities of soft-computing methods for modeling nonlinear processes, it seems they are very useful to apply in this field. Consequently, in this study, a procedure for developing fuzzy models is introduced that employs Genetic Algorithm (GA) and Adaptive Neuro-Fuzzy Inference System (ANFIS) for optimizing them in terms of accuracy and complexity. Two water supply networks located in Kerman Province (Iran) are considered as case studies in order to illustrate the efficiency of our modeling procedure. Simulation results apparently show, applying proposed method results in achieving compact and accurate fuzzy models for estimating leakage rate.

Key words: Fuzzy · Genetic Algorithm · Water leakage · Water supply networks

INTRODUCTION

Although there are many reasons for minimizing leakage in municipal water distribution networks, perhaps the most important one relates to the quality of service [1]. Additionally, during drought periods, a system with a high leakage index cannot be properly managed and may demand frequent service interruptions. Reducing leakage volume has the advantage of diminishing costly expansions of the system through hydraulic works (e.g., channels, treatment plants etc). Leakage is also costly in terms of energy losses [1] and causes a high environmental cost. In fact, a lot of time is required in order to detect the leakage of water in a widely distributed water supply network and therefore, a large amount of water goes waste.

Thus, due to the aforementioned reasons about the importance of managing water leakage, one way for quickly detecting leakage rate in a water supply network, is the monitoring of pressure in certain parts of the network. This work should be done by a number of sensors that are sensitive to pressure changes.

The relationship between pressure changes and the rate of leakage in water supply networks is non-linear and complex. Mathematical hydraulic models can be utilized for modeling the leakage rate in terms of pressure changes. The problems in developing these models are twofold: firstly, understanding of complex concepts related to the field of hydraulic, secondly, wasting a plenty of time for deriving a suitable mathematical relation.

The abilities of soft computing methods to discover complex relationships between variables of a nonlinear process as well as no need to deep understanding of the governing mathematical equations of that process have been led to widely use of these methods for modeling and solving non-linear regression problems [2, 3]. In the other hand soft computing approaches are model free methods that specially are employed when there is no exact governing mathematical equation for a nonlinear process. In recent years, successful applications of soft computing techniques in water distribution networks have been widely published. In the following a short review of these studies is given. Cheng *et al*. developed a methodology based on integration of seismic-based artificial neural

Corresponding Author: Amir Jalalkamali, Department of Water Engineering, Faculty of Engineering, Islamic Azad University, Kerman Branch, Kerman, Iran. Tel: +98-913-343-1541.

network model and a geographic information system (GIS) to assess water leakage [4]. Jalalkamali and Jalalkamali investigated the ability of a hybrid model of artificial neural network (ANN) to estimate leakage rate in water distribution network [5]. Nazif *et al*. used a genetic algorithm based optimization model to develop the optimal hourly water level variations in a storage tank for minimizing the leakage level in different seasons [6]. Mounce *et al*. presented the online application of artificial neural network and fuzzy inference systems for the detection of leakage in real water distribution system [7]. Koppel *et al*. showed that the Levenberg-Marquardt algorithm may successfully be used for calibration of a water distribution system model [8]. Asadiyani Yekta and Tabesh *et al*. used a methodology to calculate network leakage values using EPANET2.0 and considered the effects of leakage on the pressure of the demand nodes in water distribution systems [9, 10]. Araujo *et al*. and Burrows *et al*. presented a framework for leakage calculation in nodes. They considered leakage as a function of pressure at each node [11, 12]. Babel *et al*. described how the management of pressure can help reducing the leakage in the water distribution network [13]. Aksela *et al*.. used a method based on the self-organizing map for leakage detection in a water distribution network [14].Bremond *et al*. reported an improved formulation of the hydraulic network equations that incorporate pressure-dependent leakage in water distribution systems [15]. Perendeci *et al*. studied the potential of using Neuro-fuzzy models for an anaerobic wastewater treatment plant [16]. Among all of the soft computing techniques, ones that use “GAs” for designing accurate and compact fuzzy models, are called Genetic Fuzzy Systems and have been widely developed in recent years for engineering applications [2, 3 and 17]. A Genetic Fuzzy System can optimally exploit from the universal approximator property of fuzzy models as well as the search advantages of GAs for optimizing.

Our main intention in this study is to develop an algorithm in order to design compact and interpretable fuzzy models for the sake of estimating the leakage rate in the water distribution networks. Due to search abilities of GAs and advantages of using subtractive clustering and ANFIS for modeling nonlinear processes, a hybrid method is proposed in this paper, in which GA is utilized for optimizing the structure of a fuzzy model by means of adjusting subtractive clustering parameters and ANFIS is employed for fine tuning of fuzzy model parameters. The remainder of this paper is organized as following. After introduction, the problem of water leakage is described in section2. Then, section 3 discusses about the task of fuzzy modeling by subtractive clustering and the application of genetic algorithms for developing fuzzy models. Our proposed method is introduced in section4. In the fifth section simulation results are presented by tables and graphs. Finally, the sixth section concludes this research.

Water Leakage Problem: Water losses are inevitable in distribution systems and water institutions/utilities should strive to supply water efficiently and effectively by minimizing water losses [18, 19].

Water loss= Real losses + Apparent losses

Where real losses cover leakage from pipes, joints and fittings and leakage from reservoirs etc. Whereas apparent losses consist of unauthorized connections (theft and illegal) and metering errors. The major part of real losses is often due to leakage and is usually due to lack of maintenance or failure to renew and replace ageing systems. To measure leakage in a zone of an intermittent system all outlets should be plugged and the amount of water entering the zone during the assessment measured. This is a very laborious process [18].

Nowadays statistical based methods are being developed for detailed leakage assessment. A minimum nightly flow method is used to obtain leakage statistical information related to pressure. In order to do so, the existing flow in an isolated system (a system in which its input and output is controlled and is isolated from any other surrounding systems) is measured at the time of night with minimum demand known as minimum nightly flow method. Minimum nightly flow method consists of normal nightly consumption and leakage. The leakage can be calculated by subtracting the determined normal nightly consumption from the minimum nightly flow.

In this study, by setting pressure sensors on different locations of the main pipe and creating different leakages in different parts of the systems, the pressure variations related to leakages are recorded. Information generated from the leakage simulation is used by our proposed algorithm and the leakage rate is estimated regarding the pressure variations.

Fuzzy Modeling: The main contribution of fuzzy modeling theory is its ability to handle many practical problems that cannot be adequately represented by conventional methods. Fuzzy modeling of nonlinear systems has been the focus of many scientific researches [2, 3]. Takagi and

Sugeno (TS) [20] have proposed a search algorithm for a fuzzy controller and generalized their techniques to fuzzy identification. Jang [21] proposed a Neuro-Fuzzy architecture namely Adaptive Neuro-Fuzzy Inference System (ANFIS) and a learning procedure which combines fuzzy logic with neural networks for inference. ANFIS is capable of constructing input–output mapping accurately based on both human knowledge and stipulated input–output data pairs. One of the automated data-driven based methods for constructing the primary fuzzy models is Chiu's subtractive clustering [22] that is described briefly in following.

Subtractive Clustering: In order to obtain a set of rules and avoiding the problems inherent in grid partitioning based clustering techniques, (i.e. Rule Base (RB) explosion), subtractive clustering is usually utilized [2, 3]. This technique is employed since it allows a scatter input-output space partitioning [2, 3].

Subtractive clustering is, essentially, a modified form of the Mountain Method. Thus, let Z be the set of N data points obtained by concatenation of the inputs and output. In the algorithm, each point is seen as a potential cluster centre, for which some measure of potential is assigned according to equation (1):

$$p_i = \sum_{j=1}^{N} e^{-\alpha\left[z_i - z_j\right]^2} \tag{1}$$

Where $\alpha = 4 / r_a^2$ and $r_a > 0$ defines the neighborhood radius for each cluster center. Thus, the potential associated with each cluster depends on its distance to all the points, leading to clusters with high potential where neighborhood is dense. After calculating potential for each point, the one with higher potential is selected as the first cluster center. Let $z*_1$ be the center of the first group and $p*_1$ its potential. Then the potential for each z_i is reduced according to equation (2), especially for the points closer to the center of the cluster:

$$p_i = p_i - p_1^* e^{-\beta\left\|z_i - z_1^*\right\|^2} \tag{2}$$

Also $\beta = 4 / r_b^2$ and $r_b > 0$ represent the radius of the neighborhood for which significant potential reduction will occur. The radius for reduction of potential should be to some extend higher than the neighborhood radius to avoid closely spaced clusters. Typically, $r_b = 1.25_r$. Since the points closer to the cluster centre will have their potential strongly reduced, the probability for those points to be chosen as the next cluster is lower. This procedure (selecting centers and reducing potential) is carried out iteratively until stopping criteria is satisfied. Additionally two threshold levels are defined, one above which a point is selected for a cluster center and the other below which a point is rejected.

By the end of clustering, a set of fuzzy rules will be obtained. Each cluster represents a rule. However, since the clustering is carried out in a multidimensional space, the related fuzzy sets must be obtained. As each axis refers to a variable, the centers of the Membership Functions (MFs) (which are Gaussian in this case) are obtained by projecting the center of each cluster in the corresponding axis. The widths are obtained on the basis of the radius r_a for each dimension [22]. Therefore, number of produced clusters and subsequently the number of generated rules by subtractive clustering can be controlled via varying values of r_a for all dimensions (inputs and output). In this study, for each fuzzy model produced by subtractive clustering, these radii values are fine tuned by means of GA for the sake of reducing the number of rules (i.e. the structure of fuzzy model). Then, the parameters of this fuzzy model are fine tuned by ANFIS. ANFIS employs back-propagation algorithm and recursive least square method in order to adjust antecedent and consequent MFs parameters of fuzzy rules respectively. Details of tuning process by ANFIS are described in [22].

Genetic Fuzzy Systems: Generally speaking, developing Fuzzy Systems (FS) is a complex process and needs to employ some powerful optimization techniques [17]. On the other hand, various GAs are popular as global optimization methods, which are derivative free [17]. Therefore, in several recent researches, GAs have been utilized in order to develop FS for modeling, identification and classification tasks [2, 3, 17]. These systems are generally named Genetic Fuzzy Systems.

A FS is constituted from a Rule Base (RB) and some information about the Membership Functions (MFs) used in RB namely Data Base (DB). RB and DB are called Knowledge Base (KB). Fuzzy systems are generally categorized in two classes namely approximate and descriptive ([17]). This taxonomy is based on the nature of fuzzy models. The descriptive fuzzy model is essentially a qualitative expression of the system. Such models contain a KB in which the fuzzy sets giving meaning (semantics) to the linguistic labels are uniformly defined for all rules included in the RB. It constitutes a descriptive approach since the linguistic labels take the same meaning for all the fuzzy rules contained in the RB. In the approximate fuzzy model, a KB is considered for which each fuzzy rule presents its own meaning, i.e. the linguistic variables involved in the rules do not take, as their values, any linguistic label from a global term set. Approximate fuzzy models are more accurate than descriptive while descriptive models are more interpretable than accurate ones [17].

Recently, data-driven FSs have been widely used for modeling complex systems such as nonlinear identification, regression and classification tasks [2, 3]. The main and common problem of all methodologies for developing an FS is to find a suitable tradeoff between the accuracy of model and the complexity of it [20]. Correspondingly, according to Occam's razor parsimony principle, from all models that can describe a process accurately the simplest one is the best [23]. Therefore, in this study, genetic algorithm, subtractive clustering [3] and ANFIS are utilized in order to extract suitable approximate models in terms of accuracy and compactness. Both the structure and accuracy of fuzzy model are the subject of optimization by GA whereas ANFIS is employed for increasing the accuracy of model by fine tuning of its parameters. In order to generate the primary fuzzy model the subtractive clustering approach is employed to produce a Takagi-Sugeno-Kang Fuzzy Inference System (TSK FIS). Subtractive clustering has the advantage of avoiding the explosion of the RB and has been recently utilized in order to produce approximate fuzzy models for nonlinear tasks [2, 3]. Altering radii parameters in subtractive clustering can be led to different fuzzy models in terms of accuracy and complexity. Large values of these radii parameters result in producing more compact fuzzy models (i.e. fewer number of rules) while the accuracy is not suitable. Small values of radii parameters lead to generating many rules whereas the accuracy is desirable.

Proposed Method: A hybrid of canonical real-coded GA, subtractive clustering and ANFIS is utilized in order to produce suitable approximate fuzzy models in terms of accuracy and parsimony. The main procedure of modeling is an optimization task performed by GA where both the accuracy and compactness of fuzzy models are subjects of optimization simultaneously. Overall process of optimization by GA consists of Fitness assignment, Selection, Crossover and Mutation. Generating a fuzzy model based on subtractive clustering method is carried out in the fitness assignment part of GA. The process is shown in the Figure 1.

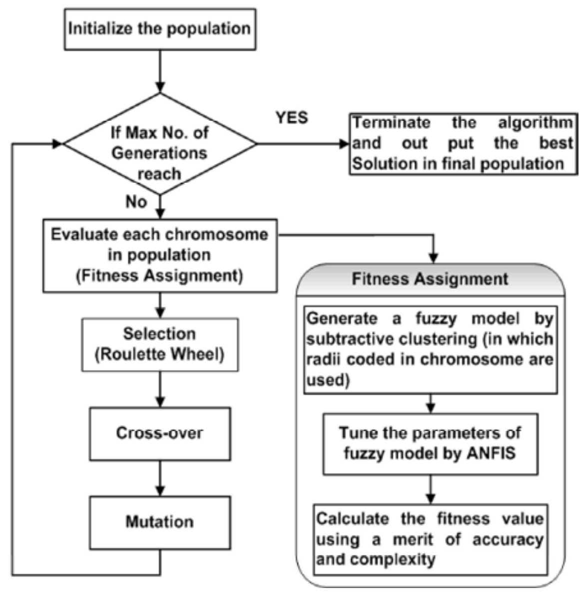

Fig. 1: Steps of modeling procedure.

Subtractive clustering method can be used for generating a TSK fuzzy model in which the number of rules (i.e. the number of clusters) can be determined through radii parameters dedicated into dimensions. As mentioned earlier in previous section, these radii are used for generating clusters. Each cluster represents a rule and regarding to the fact that clustering is carried out in multidimensional space, fuzzy sets for each rule must be obtained. The centers of MFs are obtained by projecting the center of each cluster in the corresponding dimension. The widths of MFs for each dimension are obtained on the basis of radius r_a which is considered for that dimension. Therefore, each chromosome in this study encodes radii values for all dimensions (inputs and outputs) of a fuzzy model. These radii of fuzzy model are then employed by subtractive clustering for generating a TSK FIS. Chromosome representation for GA is given below:

Where ra_i (for, $i = 1,\ldots NI$) is the radius for input number i and NI denotes the total number of inputs in fuzzy model. ra_{out} is the radius for the output of fuzzy model that is used by subtractive clustering as well as the input radii in order to obtain data clusters.

In order to evaluate a given model especially when data is scarce, Cross Validation (CV) is a suitable method [23]. Since CV avoids over-fitting and under-fitting problems and results in a tradeoff between the accuracy and complexity of a model. There are some merits namely information criteria that are good alternatives for computationally prohibitive CV method [23]. When an

Information Criterion (IC) is used, data is not split up in different parts. Rather, training is performed on the whole dataset. Some information criteria such as Akaike or Schwarz–Rissanen Criterion (SRC) may be used [3, 23]. In this work SRC is used for assigning the fitness value for each individual, which is given by:

$$SRC(Nm) = \ln(mse) + (\ln(N)/N).Nm \qquad (3)$$

Where *mse* is mean square error, *N* is the number of training samples and *Nm* is the number of model's parameters. In other words, *Nm* is a characteristic to determine the complexity of model. Let's, each rule in fuzzy models has *NI* antecedents and the number of rules generated by the subtractive clustering method for each fuzzy model to be *NR*. Subtractive clustering produces a TSK fuzzy model in which each antecedent of a rule has *NR* Gaussian MFs and each Gaussian MF has 2 parameters for tuning. Therefore, 2*NI* is the number of antecedent parameters for each rule. Moreover, *NI* +1 is the number of consequent parameters for each TSK fuzzy model rule. Consequently, the total number of parameters for each rule (antecedent and consequent parameters) is 2*NI* + *NI* +1. Thus, total number of parameters of model can be obtained as following:

$$Nm = NR.(2NI + NI + 1) = NR.(3NI + 1) \qquad (4)$$

As illustrated in Fig. (1), in fitness assignment part, subtractive clustering is firstly employed for generating an initial fuzzy model according to radii values coded in each chromosome. Then the ANFIS is utilized for the sake of adjusting MFs parameters of fuzzy model. The reason of using ANFIS instead of coding fuzzy model's parameters in GA chromosome is to diminish the computational efforts of GA for optimization. Finally, in order to evaluate each chromosome (given in Fig. 2), Eq(3) is utilized. This relation results in achieving a good tradeoff between accuracy and parsimony whilst computationally prohibitive CV is no longer required. The process of fitness assignment repeatedly continues to repeat three mentioned steps (i.e. Fuzzy model generation by clustering, tuning model's parameters by ANFIS and fitness value assignment by Eq(3) until all populations' members (chromosomes) are assigned a fitness value. The overall optimization process will go on until the termination condition is met (i.e. reaching to the Maximum number of generations).

Simulation and Results

Simulation Setup: The population size and generation numbers for GA are set to *ps* = 100 and *G* = 50, respectively. The 1-point crossover with the probability of 0.7 is employed. Classical mutation with probability 0.02 is used and selection method is the roulette wheel. Number of epochs and learning rate are set to 100 and 0.2 for ANFIS. Ranges of radii are considered to be in interval [0.1, 2]. Two water distribution networks from Kerman (one of Iranian provinces) are considered as case studies. After creating leakages in the networks, some data are taken for developing models of leakage by means of proposed method. In the following subsections case studies are described in more detail.

Case Studies: Kerman, a province in Iran, is located in a dry region with relatively low rainfall amounts. The long-term annual precipitation for the area has noticeably decreased from 150 to 100 (mm/year) during the 20 last years (1990-2011). The only source of drinking water supply in this province is practically limited through groundwater resources. The data set was collected by Iranian Ministry of Energy (IMOE). Since the networks of this province are too old, pipe network leakage is the main reason of fresh water loss. Therefore, two case studies are used in order to examine the efficiency and applicability of the proposed method in estimating water leakage in Kerman as shown in Fig 3.

ra_1	ra_2	ra_3	. . .	ra_{NI}	ra_{out}

Fig. 2: Chromosome representation for GA given in Fig.1.

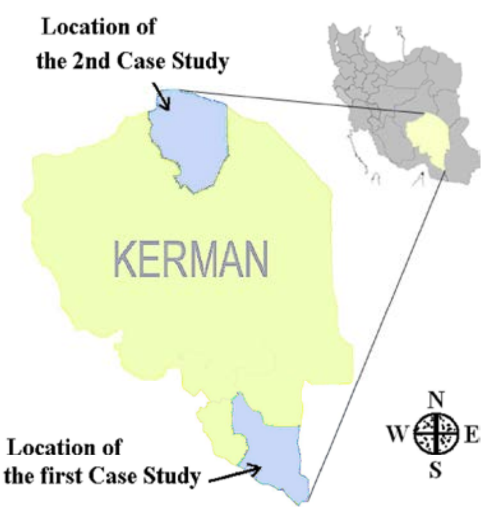

Fig. 3: Location of the case studies in Kerman, Iran.

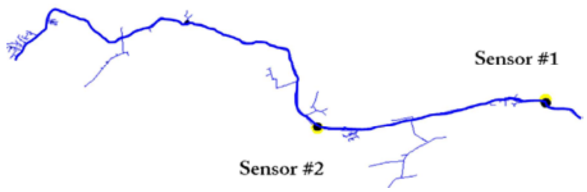

Fig. 4: Map of the first case study, Branching system

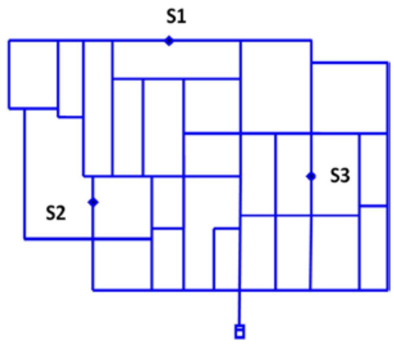

Fig. 5: Map of the second case study, Loop System

Case 1: The first case study is focused on a branching system that is located in the south of Kerman province as shown in Fig 3. The length of this network is 49 kilometers that covers several villages. Two pressure sensors are installed on the main pipe of the network as shown in Fig 4. Some random fractures in different zones of water distribution network are created and then the changes of pressure are recorded. Data consist of 220 samples in three columns. The first two columns are pressures measured in sensors 1 and 2 respectively. The values in these columns are fed as inputs to the fuzzy model. The third column is the leakage rate which is the target output. The fuzzy model should be able to estimate the values in this column. In this study testing and validation data set were considered together and due to the lack of data 80 percent of data is used for training.

Case 2: The second case study is focused on a loop system which is located in the North of Kerman province as shown in Fig 3. Number of network pipes are 133 that have been designed for a population of 4500 person. Three pressure sensors are installed in this work that their locations are shown on the network in Figure 5 by S1, S2 and S3 letters. Other nodes are depicted by dots on Fig 5. Also, 250 random fractures in different zones of water distribution network are created and then the changes of pressure are recorded. Data consist of 200 samples in four columns. The first three columns are pressure values measured in sensors 1, 2 and 3 respectively. The values in these columns are fed as inputs to the fuzzy model. The fourth column is the target output and the fuzzy model must predict the values in this column.

Table 1: The statistical parameters of recorded data in branching system

Data set	*Unit*	X_{mean}	S_x	C_{sx}	X_{max}	X_{min}
Pressure head in Sensor 1	*m*	17.97	3.72	-2.68	19.53	1.27
Pressure head in Sensor 2	*m*	20.56	8.48	-0.72	28.94	0.24
Leakage rate	*Lit/sec*	2.45	1.19	0.04	5.00	0.00

Table 2: The statistical parameters of recorded data in loop system

Data set	*Unit*	X_{mean}	S_x	C_{sx}	X_{max}	X_{min}
Pressure head in Sensor 1	*m*	35.64	3.76	-0.49	40.35	24.79
Pressure head in Sensor 2	*m*	52.60	3.78	-0.51	57.32	41.34
Pressure head in Sensor 3	*m*	67.85	3.52	-0.45	72.29	58.18
Leakage rate	*Lit/sec*	2.29	1.16	-0.14	5.00	0.00

The statistic characteristics of each recorded data are given in Table. 1 and Table. 2. In the Table (1) and (2) the X_{mean}, S_x, C_{sx}, X_{max} and X_{min} respectively denote the mean, standard deviation, skewness coefficient, maximum and minimum of recorded data for two case studies.

RESULTS AND DISCUSSIONS

Characteristics of obtained fuzzy models namely number of fuzzy rules, number of inputs and number of parameters are given below in Table (3) as well as some performance criteria for illustrating the efficiency of models. These criteria are Root Mean Square Error (RMSE) and R^2 (correlation coefficient). As illustrated in Table (3) our proposed method yields accurate and simple fuzzy models in terms of parameters.

In the theory of system identification and modeling, a good model is an accurate and simple model for training data [23]. The simplicity of obtained model by means of training data results in a good generalization of model when unseen data are presented in to model (i.e. a good performance on test data) [23]. Information given in Table (3) apparently shows the efficiency of our proposed algorithm for finding accurate and simple fuzzy models. Figure (6) show the distribution of antecedent MFs (i.e. pressures measured in sensor 1 and 2) for two obtained rules related to the first case study.

Obtained rules for the first case study are given in Table (4) below. In these rules Low, Middle and High are linguistic variables assigned to MFs and are represented by **L**, **M** and **H** letters respectively. y_1 and y_2 are the outputs of fuzzy rules. As mentioned in section (3.2), in an approximate fuzzy model, a KB is considered for which each fuzzy rule presents its own meaning. Therefore, **L**, **M**

Table 3: Characteristics of obtained optimum fuzzy models.

				Train		Test	
Case study	No. of Inputs (*NI*)	No. of Rules (*NR*)	No. of Parameters (*Nm*)	RMSE	R^2	RMSE	R^2
Branching system	2	2	14	0.99	0.926	0.47	0.986
Loop system	3	3	30	0.41	0.997	0.43	0.997

Table 4: Obtained fuzzy rules for the first case study. Antecedent MFs H and M are presented in Fig 6

Rule No.1	If p_1 is H and p_2 is H then $y_1 = -1.5834\ p_1 + 0.1189\ p_2 + 1.5828$
Rule No.2	If p_1 is H and p_2 is M then $y_2 = 2.0510\ p_1 -0.6462\ p_2 -1.0829$

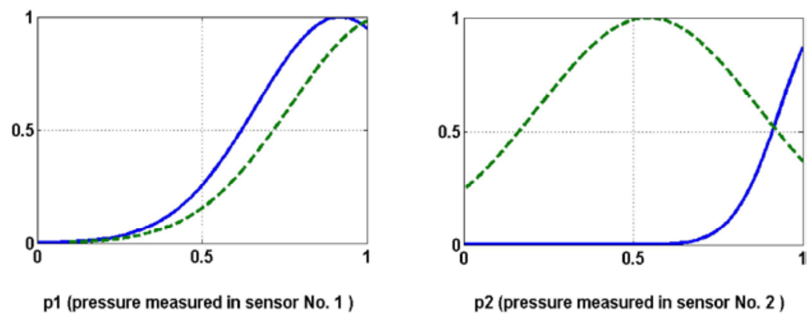

Fig. 6: The bold lines represent MFs related to antecedent of the first obtained rule. Dashed lines indicate MFs related to antecedent of the second obtained rule.

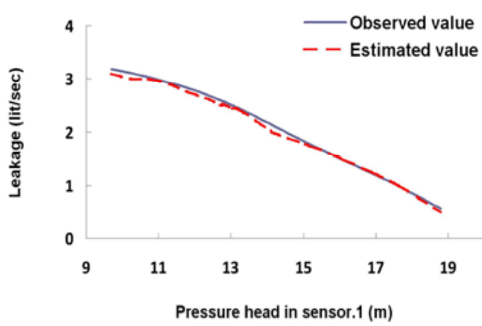

Fig. 7: Observed and estimated values of leakage rate in branch system versus the pressure values measured by sensor 1.

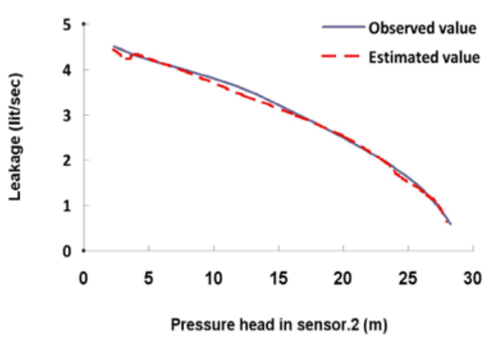

Fig. 8: Observed and estimated values of leakage rate in branch system versus the pressure values measured by sensor 2.

and **H** have not the same meaning for these two rules. For example, two different MFs depicted on input p_1 are different in parameters' values but both of them have the label **H** (one for the first rule and the other for second rule).

In addition to the Table (3) which indicates the overall accuracy and simplicity of obtained fuzzy models on train and test data for both cases, some graphs are presented in the following for more illustrations. Regarding to both case studies, these graphs depict the pressure head measured in each sensor versus the leakage rate for both actual and predicted leakage on test data. These figures are given in the following in order to visualizing the prediction accuracy. Figures (7) and (8) plot the leakage rate in branch system versus the values of pressure measured by sensor one and two respectively. In these figures the output of obtained fuzzy model are indicated by dashed lines and are compared with actual values of leakage rate (bold lines).

In the following, firstly, obtained fuzzy rules for the second case study are presented as well as antecedents MFs distributions. Figure (9) depicts the distribution of antecedent MFs for the three obtained rules. Bold, dashed and dotted lines are related to antecedent MFs of the first,second and third rules respectively. Rules are given in Table (5). Secondly, Figures (10-12) are depicted similar

Table 5: Three obtained fuzzy rules for the second case study. Antecedent MFs H and M are presented in Fig (9) above

Rule No.1	If p_1 is M and p_2 is M and p_3 is H then $y_1 = 3.6182\ p_1 - 11.1489\ p_2 + 8.4136\ p_3 + 1.7583$
Rule No.2	If p_1 is M and p_2 is M and p_3 is H then $y_2 = -3.9888\ p_1 + 11.3855\ p_2 - 9.7691\ p_3 + 0.2462$
Rule No.3	If p_1 is H and p_2 is H and p_3 is H then $y_3 = 0.0687\ p_1 - 2.1637\ p_2 - 0.4785\ p_3 - 3.2937$

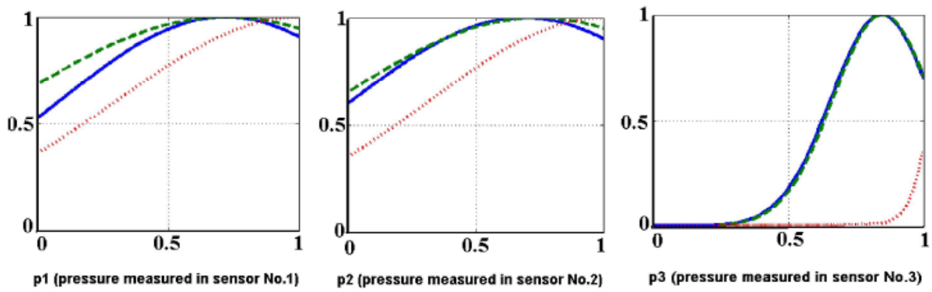

Fig. 9: The bold lines, dashed lines and dotted lines represent antecedent MFs distributions of the first, second and third obtained rules respectively.

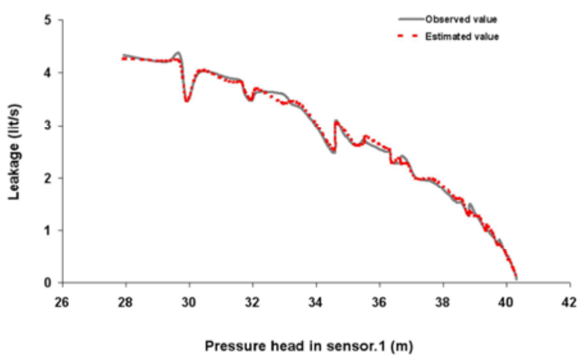

Fig. 10: Observed and estimated values of leakage rate in loop system versus the pressure values measured by sensor 1.

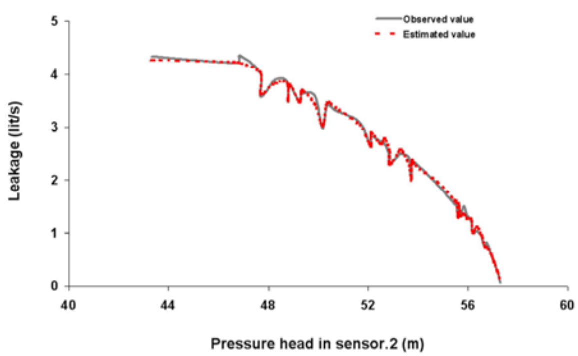

Fig. 11: Observed and estimated values of leakage rate in loop system versus the pressure values measured by sensor 2.

to 7, 8 ones in which the observed and estimated leakage rates (bold and dashed lines respectively) are visualized versus pressure measured in sensors 1,2 and 3 respectively.

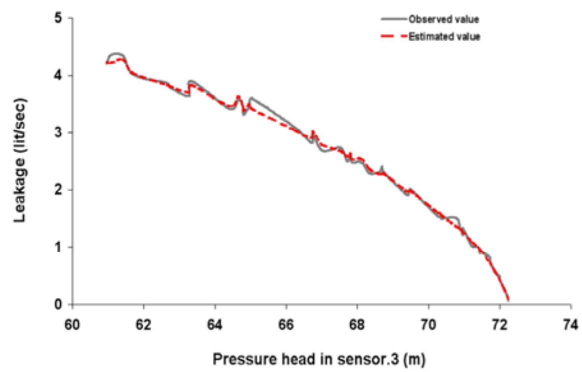

Fig. 12: Observed and estimated values of leakage rate in loop system versus the pressure values measured by sensor 3.

The obtained fuzzy model in this case is also an approximate fuzzy model in which the meaning of High for three rules is different. As shown in Figure (9), bold, dashed and dotted lines on the input p_3 are different MFs where all of them have the same label **H** in its own rule.

Concluding Remarks: Water leakage in distributed water supply networks, is an important problem due to wasting of energy resources and the quality of service in an urban area. Owing to the importance of this problem, a new hybrid algorithm is introduced in this study for the sake of estimating water leakage in terms of measured pressures at multiple positions of the network. Subtractive clustering and ANFIS are built into GA as embedded algorithms for constructing simple and accurate fuzzy models. In proposed method, GA is employed for optimizing the structure of a TSK FIS via subtractive clustering whereas ANFIS is utilized for fine tuning of this structurally optimized fuzzy model. Therefore, in fitness assignment part of GA, both accuracy and complexity of a fuzzy model are considered as subjects of optimization. Instead of coding fuzzy models' parameters in GA chromosome, ANFIS is employed for the sake of reducing the computational efforts of GA. Furthermore, the advantages of utilizing an information criterion alternatively for the CV are twofold: Firstly, the applicability of proposed procedure especially when the data is scarce for modeling, secondly, eliminating computational costs imposed by CV. Two types of distributed water supply networks namely branching and loop types in Kerman province are considered as case studies. These networks are used and the pressure and leakage rate are measured in some points of networks by means of some sensors. After that, these measured data are used as training and test data. Simulation results confirm the applicability of our proposed method in terms of producing simple fuzzy models for estimating the water leakage accurately. The method can be utilized in similar water supply networks.

REFERENCES

1. Farley, M. and S. Trow, 2003. Losses in water distribution networks, IWA Publishing.
2. Eftekhari, M., S.D. Katebi, M. Karimi and A.H. Jahanmiri, 2008. Eliciting Transparent Fuzzy Model using Differential Evolution, Applied Soft Computing, (8) 466-476.
3. Eftekhari, M. and S.D. Katebi, 2008. Extracting compact fuzzy rules for nonlinear system modeling using subtractive clustering, GA and unscented filter, Applied Mathematical Modeling, (32): 2634-2651.
4. Ho, C.I., M. Lin and S.L. Lo, 2009. Use of GIS-based hybrid artificial neural network to prioritize the order of pipe replacement in a water distribution network, Environmental Monitoring Assessment, pp: 1-13.
5. Jalalkamali, A. and N. Jalalkamali, 2011. Application of Hybrid Neural Modeling and Radial Basis Function Neural Network to Estimate Leakage Rate in Water Distribution Network, World Applied Sciences Journal, 15(3): 407-414.
6. Nazif, S., M. Karamouz, M. Tabesh and A. Moridi, 2010. Pressure management model for urban water distribution networks, Water resources management, (24): 437-458.
7. Mounce, S.R., J.B. Boxall and J. Machell, 2008. Online application of ANN and fuzzy logic system for burst detection, Proceeding of the 10th annual water distribution systems analysis conference, WDSA, pp: 735-746.
8. Koppel, T., A. Vassiljev, D. Lukjanov and I. Annus, 2008. Use of pressure dynamics for calibration of water distribution system and leakage detection, Proceeding of the 10th annual water distribution systems analysis conference, WDSA, pp: 704-715.
9. Asadiyani Yekta, A., 2005. A software tool to calculate the unaccounted for water in water distribution networks. MSc thesis, Civil Engineering Department, University of Tehran, Iran.
10. Tabesh, M., AH. Asadiyani Yekta and R. Burrows, 2009. An integrated model to evaluate losses in water distribution networks, Journal of Water Resources Management, (23): 477-492.
11. Araujo, L.S., S.T. Coelho and H.M. Ramos, 2003. Estimation of distributed pressure-dependent leakage and consumer demand in water supply network, In: Maksimovic C, Butler D, Memon A (eds), Advances in water supply management. Swets and Zeitlinger, Lisses, pp: 119-128.
12. Burrows, R., G. Mulried and M. Hayuti, 2003. Introduction of a fully dynamic representation of leakage into network modeling studies EPANET, In: Maksimovic C, Butler D, Memon A (eds), Advances in water supply management. Swets and Zeitlinger, Lisses, pp: 109-118.
13. Babel, M.S., Md. S. Islam and A. Das Gupta, 2009. Leakage management in low-pressure water distribution network of Bangkok, Water science and technology: Water Supply, 9(2): 141-147.
14. Aksela, K., M. Aksela and R. Vahala, 2009. Leakage detection in a real distribution network using a SOM, Urban Water Journal, 6(4): 279-289.
15. Bremond, B., P. Fabrie, E. Jaumouille, I. Mortazavi and O. Piller, 2009. Numerical simulation of a hydraulic saint-vanant type model with pressure demand leakage, Applied Mathematics Letters, 22(11): 1694-1699.
16. Perendeci, A., S. Arslan, A. Tanyolac and Serdar S. Celebi, 2007. Evaluation of input variables in adaptive-network-based fuzzy inference system modeling for an anaerobic wastewater treatment plant under unsteady state, Journal of Environmental Engineering, 133(7): 765-771.

17. Herrera, F., 2008. Genetic fuzzy systems: taxonomy, current research trends and prospects, Evol. Intel., 1: 27-46.
18. Colombo, A.F. and B.W. Karney, 2002. Energy and costs of leaky pipes: Toward comprehensive Picture. J. Water Resour. Plan. Manage., 128(6): 441-450.
19. Kailash Prasad, R. and S. Mathur, 2007. Groundwater flow and contaminant transport simulation with imprecise Parameters, Journal of Irrigation and Drainage Engineering, 133(1): 61-70.
20. Takagi, T. and M. Sugeno, 1985. Fuzzy identification of systems and its application to modeling and control, IEEE Trans. Syst. Man Cyber., 15(1): 116-132.
21. Jang, J.S.R., 1993. ANFIS: adaptive-network-based fuzzy inference systems, IEEE Trans. Syst. Man Cyber., (23): 665-685.
22. Chiu, S.L., 1994. Fuzzy model identification based on cluster estimation, J. Intell. Fuzzy Systems, (2): 267-278.
23. Nelles, O., 2001. Nonlinear System Identification, Springer-Verlag, Berlin Heidelberg.