Comparison of Fuzzy Inference System and Multiple Regression to Predict Synthetic Envelopes Clogging

Bakhtiar Karimi (Corresponding Author)

Ph.d Student, Irrigation and Drainage Engineering, Member of Young Researchers

Club of Islamic Azad University Branch of Kermanshah, Kermanshah, Iran

Tel/Fax: 98-261-224-1119   E-mail: Bakhtiar.Karimi@gmail.com

Farhad Mirzaei

Associated Professor, Department of Irrigation and Reclamation Engineering

Faculty of Agricultural Engineering and Technology, University of Tehran, Karaj, Iran

E-mail: fmirzaei@ut.ac.ir

Mohammad Javad Nahvinia

Graduate Student, Department of Irrigation and Reclamation Engineering

Faculty of Agricultural Engineering and Technology, University of Tehran, Karaj, Iran

E-mail: javad.nahvinia@gmail.com

Behnam Ababaei

Ph.d Student, Irrigation and Drainage Engineering, Member of Young Researchers

Club of Islamic Azad University, Science and Research Branch of Tehran, Iran

E-mail: behnam.ab@gmail.com

Abstract

Geo-synthetic materials are being used with acceptable performance in soil and water projects worldwide. Geotextiles are one of the categories of geo-synthetics being used in drainage systems. First generation of geotextiles used in the late 1950's as an alternative for gravel envelopes. In this research two methods (multiple regression and fuzzy interference system) evaluate to predict synthetic envelope clogging. In multiple regression method the correlation coefficients for PP450, PP700 and PP900 are 62.66%, 79.37% and 90.62%, respectively and results of fuzzy interference system and decision tree showed that this method have high potential in comparison with multiple regression and values of total classification accuracy for PP450, PP700 and PP900 are 98.6%, 97.3% and 98% respectively. Then final results of this research showed fuzzy interference systems by using decision tree have high potential to predict clogging in envelops.

Keywords: Prediction, Envelope clogging, Fuzzy logic, Multiple regressions

1. Introduction

Among popular products are synthetic materials. These materials are mainly manufactured from petrochemicals and oil industry. Geo-synthetic materials are being used having acceptable performance in soil and water projects worldwide. Geotextiles are one of the categories of geo-synthetics being used in drainage systems. First generation of geotextiles uses began in the late 1950's and they were used as an alternative for gravel envelopes.

In general, geotextiles are classified based on their type of polymers (polypropylene, polyethylene, and polyester), kind of threads (single thread, multi thread, cleft membrane) and their texture (textile, non-textile and lumpy textile). These properties, as in any material used in engineering projects, are depended on their production method and the properties of final product.

Karimi (2009) assessed the performance of three models of geotextiles (PP-450, PP-700 and PP-900) in comparison with common drain envelopes. They recommended PP-450 for Khorram Shahr drainage project by using required tests (flow variations, hydraulic conductivity, gradient ratio, hydraulic conductivity ratio).

Hasanoghli (1997) used an especial kind of geotextile filters as drain pipe which was made of wicker-like textile with polyester wraps and polypropylene hollow wefts. In this test, pipes were completely flexible and water could get into the pipe from small openings which were evenly distributed through out the pipe length. The size and distribution of these openings and of course the pipe water conductivity was controllable by changing the function of the geotextile. In such a situation, provided that the pipes performance were acceptable, there was no need for gravel envelopes and trenchless installation would be possible leading to major reduction in expenses.

Palmeira and Gardoni (2002) measured the effects of different pressure on hydraulics and physical properties of geotestiles and found that for drain pipes installed deeper and for increased soil pressure on pipe and envelope, the size and conductivity of filter openings would be smaller and a lower performance of drain pipe and filter would result. Fernando et al. (2006) conducted a study on biological clogging of geotextiles and mineral materials (for filtering agricultural waste waters in a 5-year period). This study revealed that both kinds of filters had some effects on water quality. Also, they found that geotextiles are more frugal in terms of economic and technical issues.

Soubaida et al. (2008) studied woven and non-woven geotextile filters using a pressure membrane and assessed the tensions on geotextiles. They concluded that tensions and strains on drain filters have a major effect on outflow rate and filter's hydraulic conductivity which is depended on the density of threads used in the envelope. Also, the resistance of these filters is related to surrounding soil particle size distribution.

Cho-seng et al. (2005) carried out an experiment in laboratory to assess the performance of a geotextile envelope and determine the values of outflow rate, hydraulic conductivity and gradient ratio. They found envelope clogging potential by placing sandy granular material on it. The results of this study showed that usage of particles smaller that the envelope opening spaces because the outflow values to be smaller. Using different size of soil particles around the drain envelopes increases outflow from drains and decreases the variations of gradient ratios.

Nejadyani (2007) conducted experiments to evaluate performance of a drainage synthetic envelope (PP450) in comparison with conventional envelope and showed synthetic envelopes have many design criteria, although, these envelopes have low performance in comparison with conventioanal envelopes and finaly envelope PP450 recommended for case study.

Azizi (2007) evaluate performance two synthetic envelopes in comparison with two conventional envelopes and with calculation of salinity values (before and after leaching), water table fluctuation and drainage system performance showed envelope PP450 have better performance and   recommended for case study.

2. Methods and Materials

2.1 Describtion of the study area and permeameter

The laboratory tests were conducted using soil samples from a drainage project under construction in the northern Khorram Shahr, south west of Iran (Figure 1). In this study, two soil samples were provided from regions 15 and 37 and from the depth of 1.65m in order to be more similar to the real depth of drains. The results of chemical and physical analysis showed that both soil samples are largely similar in terms of particle size distribution and texture and other experiments were done using the sample from region 15.

In this study, a permeameter device according to the ASTM D-5101 standard was used for determining hydraulic conductivity and clogging potential of the combined soil-geotextile system. The main part of the system was a transparent cylinder made of Plexiglas with inner diameter of 100mm and wall thickness of about 5mm. At different levels, piezometers were installed to assess variability of hydraulic gradient through the soil sample and around the geotextile. This test was carried out for all the three types of geotextiles (PP-900, PP-700 and PP-450). A mesh steel plate (with openings of 4.76mm, mesh #4) was placed between the main part and foundation where geotextile sample was placed on. Piezometers were installed in sets of two at the same height with respect to the steel plate, i.e. Piezometer couples were at 25 and 75mm heights above the steel plate. Another piezometer was installed 143mm away from the steel plate, contrary to other piezometers; it was placed outside the soil sample. The inlet section was 162mm above the geotextile sample. In the topmost section of the device, an air valve was installed to be used for saturating the soil sample (Figure 2). The required steps were the installation of geotextile sample in place, filling the device with soil sample, making the soil-geotextile system saturated and finally, starting the test according to the standards.

After installation, by creating different hydraulic gradient (25, 50, 75 and 100 cm), the values of permeability and hydraulic conductivity of soil-geotextile system, outflow and hydraulic conductivity of the geotextile were

measured. At first, system was saturated by an upward flow in order to prevent the air from entering and to increase total head. After 24hr, the tests began

2.2 Fuzzy logic method

The step consists of five stages as shown in fig. 3 which are specially explained. The summary role of each procedure is described as followed (Mollazade et al., 2008):

Data Acquisition: this step consists of measurement of different hydraulic pressure heads, various hydraulic conductivity, outflow and values of piezometers.

Equation creation: according to pervious step data, using special equation the effect of two input parameters (outflow and pressure head) described by one special equation ($Z = h*q^{0.5}$) that shown the effects of two parameters where h is hydraulic head (cm) and q is drain outflow ($cm^3.sec^{-1}$).

Feature extraction: this step consists of extraction of the most important features calculated by using statistical feature parameter from special equation (Z).

Feature selection and classification model extraction: the J48 algorithm is used as a decision tree to choice the salient features from the whole feature set. In this section the data obtained from feature extraction procedure is divided into two data sets: training data and testing data. Training data is employed to make the model and testing data is for evaluation the model.

Clogging diagnosis: Fuzzy logic inference system is used to diagnose the clogging.

2.3 Decision tree

Firstly, according to available data sets three classes considered for clogging condition that by attention to envelope hydraulic conductivity was done these classification. Where if envelope hydraulic conductivity be less than 2.5 m/day then clogging condition is weak (Class-A) and if envelope hydraulic conductivity be between 2.5-5 m/day then clogging condition is medium (Class-B) and if envelope hydraulic conductivity be more than 5 m/day then clogging condition is good (Class-C) (Table 1). In this research J48 algorithm (A WEKA implementation of c4.5 Algorithm) was used to construct decision trees (Mollazade et al., 2008).

2.4 Multiple Regression models

Multiple Regression (MR) is a statistical technique that allows us to predict someone's score on one variable on the basis of their scores on several other variables. MR was used in order to learn more about the relationship between several independent or predictor variables and a dependent or criterion variable. The general form of these models is $y = b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_nx_n + c$ where $\{b_1, b_2, b_3, \dots b_n\}$ the regression coefficients are, and y is now written as a function of n independent variables ; x1, x2, x3, … xn. C is y-intercept (Milton et al., 1997; McClave et al., 1997; Ahmad Auli et al., 2009).

After accomplishing the tests, the values of hydraulic conductivity for soil-envelope system were estimated as average values between all the treatments as functions of hydraulic head and outflow values. Furthermore, other relations were proposed for gradient ratio. The relations between different parameters were analyzed by Sigma Plot software and proper relations were proposed. This was done for all the 3types of envelopes (3 geotextiles and). In order to evaluate the errors of the equations and the difference between measured and estimated values were analyzed using 5 different statistics (Homaee et al., 2002):

$$ME = MAX \quad \left| P_i - O_i \right|_{i=1}^{n}$$

$$RMSE = \left[\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n} \right]^{\frac{1}{2}}$$

$$CD = \frac{\sum_{i=1}^{n}\left(O_i - \bar{O}\right)^2}{\sum_{i=1}^{n}\left(P_i - \bar{O}\right)^2}$$

$$EF = \frac{\sum_{i=1}^{n}\left(O_i - \bar{O}\right)^2 - \sum_{i=1}^{n}\left(P_i - \bar{O}\right)^2}{\sum_{i=1}^{n}\left(O_i - \bar{O}\right)^2}$$

$$CRM = \frac{\sum_{i=1}^{n} O_i - \sum_{i=1}^{n} P_i}{\sum_{i=1}^{n} O_i}$$

Where P_i is estimated value, O_i is measured value; n is the number of treatments and $\overline{O}$ is the average of measured values. The minimum value of ME, RMSE and CD is zero and the maximum value of EF equals 1. EF and CRM can have negative values. The higher values of ME, the worse results of the model. CD shows dispersion between measured and estimated values. EF compares estimated values with the average of measured values. A negative value of EF indicates that the average of measured values has a better estimation in relation to the estimated values. If all the measured and estimated values are completely equal, then: CRM=0, EF=0, CD=1, RMSE=0, ME=0.

3. Results and Discussion

3.1 Multiple Regression

By proposing different equations in Sigma Plot Software to estimate envelope hydraulic conductivity values, it was revealed that the relationship between these parameters and drain flow values are linear and the one between these parameters and hydraulic head values are in a parabolic form and this holds for all geotextiles (Table 2). Also, the results showed that by gradual increase of hydraulic head and soil particles transfer, envelope clogging arises and hydraulic conductivity decreases because soil particles are not still stabilized and can be moved easily. But at higher gradients, since soil particles are stabilized, clogging potential decreases substantially. The main reason for the increase of hydraulic conductivity values at higher hydraulic heads is increase of hydraulic gradient values. Again, it was concluded that envelope hydraulic conductivity values have the lowest values when hydraulic head value ranges between 40 and 60 cm and at higher head values, as hydraulic gradient increases, the soil particles (which caused clogging issue) move and hydraulic performance gets better (Fig 5). The results of correlation coefficient showed that multiple regressions have low precision to predict of envelope hydraulic conductivity and values of correlation coefficients for PP450, PP700 and PP900 are 62.66%, 79.37% and 90.62%, respectively (Table 2) however according Table 3 shows that all the error values were in an acceptable range.

3.2 Fuzzy logic

A membership function (MF) is a curve that defines how each point in the input space is mapped to a membership value (or degree of membership) between 0 and 1. Observing the values of the feature, based on which the branches of the decision tree are created for different conditions of the envelope, MFs for the corresponding features are defined that their results do not show due to their wide space. After defining membership functions and generating the "If Then" rules by J48 algorithm, the next step is to build the fuzzy inference engine classification (Mollazade et al., 2008). The fuzzy toolbox available in MATLAB software was used for producing fuzzy inference rules. Each rule was taken at a time and using membership functions and fuzzy operators the rules were entered the results of J48 algorithm are shown in Figs. 6 to 8. Decision trees define the relation between features and the situation of the clogging. Tracing a branch from the root node leads to a condition of the clogging and decoding the information available in a branch in the form of "if-then" statement gives the rules for classification using fuzzy for various conditions of envelopes. Hence the usefulness of the decision tree in forming the rules for fuzzy classification is established. The top node of decision tree is the best node for classification. The data sets of the features for each envelope have 150 samples. In each operating condition, two-thirds of samples are employed for training process and the remaining samples for testing purposes. The detailed descriptions of those data sets are given in Table 4 (Mollazade et al., 2008).

3.3 Fuzzy rules

According to Figs. 6-8, fuzzy rules were described with "If-Then" expressions. All rules (Figs. 9-11) are evaluated in parallel.

3.4 System Accuracy

The classification results are calculated using cross-validation evaluation where the data set to be evaluated is randomly partitioned so that in each condition 100 samples are used for training and 50 samples are used for testing. The confusion matrix for each condition is given in table 5 to 7. In confusion matrix, each cell contains the number of samples that was classified corresponding to actual algorithm outputs. The diagonal elements in the confusion matrix show the number of correctly classified instances (Mollazade et al., 2008).

The performance of the classifier can be checked by computing the statistical parameters such as sensitivity,

Specificity and total classification accuracy defined by (Mollazade et al., 2008):

• **Sensitivity**: number of true positive decisions/number of actually positive cases.

• **Specificity**: number of true negative decisions/number of actually negative cases.

• **Total classification accuracy**: number of correct decisions/total number of cases.

The values of statistical parameters are given in Table 8 to 10. Results show that the total classification accuracy for PP450, PP700, and PP900 conditions 98.6%, 97.3% and 98% respectively and results of fuzzy interference system and decision tree showed that this method have high potential in comparison with multiple regression.

4. Conclusion

The results of correlation coefficient showed that multiple regressions have low precision to predict of envelope hydraulic conductivity and values of correlation coefficients for PP450, PP700 and PP900 are 62.66%, 79.37% and 90.62%, respectively and results of fuzzy interference system and decision tree showed that this method have high potential in comparison with multiple regression and values of total classification accuracy for PP450, PP700 and PP900 are 98.6%, 97.3% and 98% respectively. Then final results of this research showed fuzzy interference systems by using decision tree have high potential to predict clogging in envelops.

References

Ahmad Auli, K., Parsinejad, M., & Rahmani, B. (2009). Estimation of saturation percentage of soil using multiple regression, ANN, and ANFIS techniques. *Journal of computer and Information science*, 2 (3), 127-136.

Azizi. J. (2007). Evaluation of synthetic drain filter performance compared with conventional gravel envelopes in drain tubes. *MSc Thesis*. Irrigation Engineering Dept, University of Ahvaz.

Cho- sen wu,Yung- Shang hong. (2005). Soil-nonwoven geotextile filtration behavior under contact withdrainage materials. *Geotextiles and Geomembranes*, 24, 1-10.

Fernando F. Junqueiraa, Anto nio R.L. Silvab, Ennio M. Palmeira. (2006). Performance of drainage systems incorporating geosynthetics and their effect on leachate properties. *Geotextiles and Geomembranes*, vol: 24, No.5.

Hassanoghli, A. (1997). The evaluation of geotextile drain pipe filters performance in soil (laboratory experiment). MSC thesis, college of Agriculture, *University of Tehran*. (In Persian).

Homaee, M., Direksen C., & Feddes R. A. (2002). Simulation of root water uptake: I. Non-uniform transient salinity using different macroscopic reduction functions, *Journal of Agricultural Water Management*, 57: 89-109.

Karimi, B. (2009). Evaluation of three synthetic drain envelope performances as compared with conventional gravel envelope in laboratory experiment. MSc thesis, *University of Tehran*. (In Persian).

Mollazade, K. Ahmadi, H. Omid, M. & Alimardani, R. (2008). An intelligent combined method based on power spectral density, decision trees and fuzzy logic for hydraulic pumps fault diagnosis. *International Journal of Intelligent Systems and Technologies*.

Nejadyani. M. (2007). Laboratory evaluation of synthetic envelope application in subsurface drainage and compared with conventional envelopes. *MSc Thesis*. Irrigation Engineering Dept, University of Ahvaz.

Palmeira, E. M., & Gardoni, G. (2002). Drainage and filtration properties of non-woven geotextiles under confinement using different experimental techniques. *Geotextiles and Geomembranes*, Elsevier, 20, 97-115.

Subaida, E. A., Chandrakaran, S. & Sankar, N. (2008). Experimental investigations on tensile and pullout behavior of woven coir geotextiles. *Journal of Geotextiles and Geomembranes*, 26(5), 384-392.

Table 1. Classification of envelope condition based on envelope hydraulic conductivity.

Row	Envelope Hydraulic Conductivity (m/day)	Drainage condition	Label of classification
1	$K_e<2.5$	weak	Class-A
2	$2.5<K_e<5$	medium	Class-B
3	$K_e>5$	good	Class-C

Table 2. Predictive relationship for evaluation of envelope hydraulic.

Envelopes	Predictive Models	R^2
PP450	$k_e^* = 1.0949 - 0.0195\,H + 3.2301\,Q + 0.0001\,H^2$	0.6266
PP700	$k_e = 0.2843 - 0.006\,H + 8.8344\,Q + 2.2824 \times 10^{-5}\,H^2$	0.7937
PP900	$k_e = 0.3237 - 0.0102\,H + 6.925\,Q + 7.1239 \times 10^{-5}\,H^2$	0.9062

* K_e is envelope hydraulic conductivity (m/day) H: hydraulic head (cm) Q: drain outflow ($cm^3.sec^{-1}$)

Table 3. Statistics for the proposed models.

Envelopes	Parameter	ME	RMSE	CD	EF	CRM
PP450	Ke	0.3369	0.7394	0.327	-1.0757	0.18058
PP700	Ke	0.0817	0.1133	1.2647	0.7932	0.0022
PP900	Ke	0.0256	0.0675	1.0774	0.9033	-0.0054

Table 4. Required data for envelopes.

Row	Envelope Type	Number of Training Data	Number of Training Data
1	PP450	100	50
2	PP700	100	50
3	PP900	100	50
Total data	-	300	150

Table 5. Confusion matrix for envelope PP450.

Condition	ClassA	ClassB	ClassC
ClassA	50	0	0
ClassB	2	48	0
ClassC	0	0	50

Table 6. Confusion matrix for envelope PP450.

Condition	ClassA	ClassB	ClassC
ClassA	47	3	0
ClassB	1	49	0
ClassC	0	0	50

Table 7. Confusion matrix for envelope PP450.

Condition	ClassA	ClassB	ClassC
ClassA	49	0	1
ClassB	2	48	0
ClassC	0	0	50

Table 8. The values of statistical parameters for envelope PP450.

Data Set	Sensitivity (%)	Specificity (%)	Total classification Accuracy (%)
ClassA	100	100	
ClassB	96	98	98.6
ClassC	100	100	

Table 9. The values of statistical parameters for envelope PP700

Data Set	Sensitivity (%)	Specificity (%)	Total classification accuracy (%)
ClassA	94	97	
ClassB	98	99	97.3
ClassC	100	100	

Table 10. The values of statistical parameters for envelope PP900

Data Set	Sensitivity (%)	Specificity (%)	Total classification accuracy (%)
ClassA	98	99	
ClassB	96	98	98
ClassC	100	100	
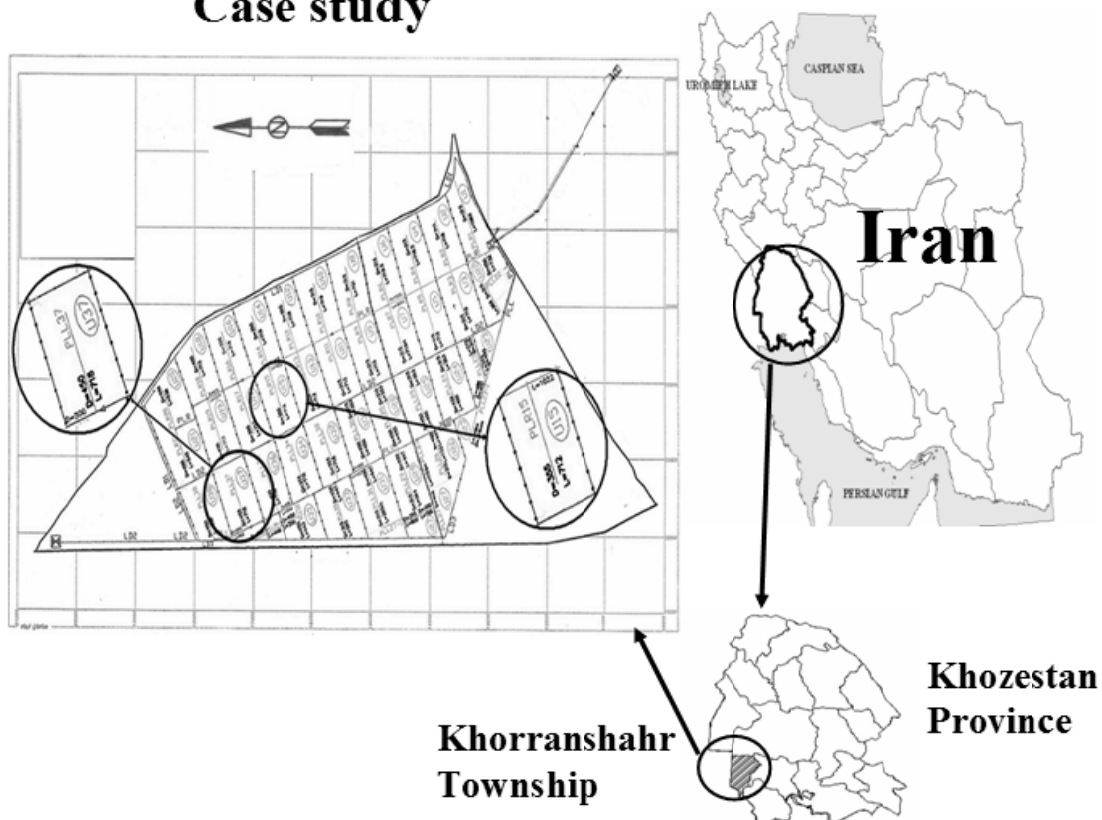
Figure 1. Situation of case study in north of Khuzestan province.
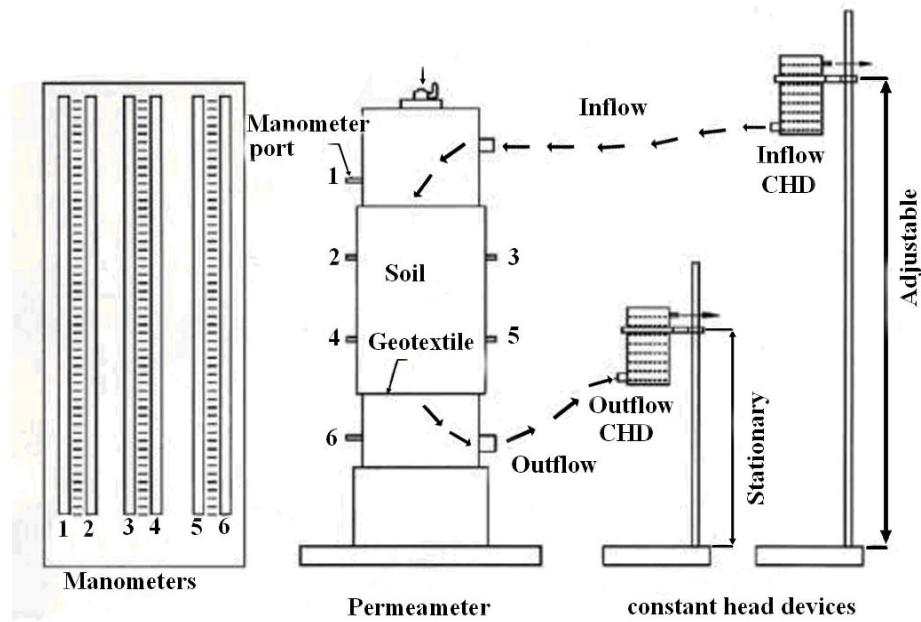
Figure 2. Permeameter device used in this experiment.
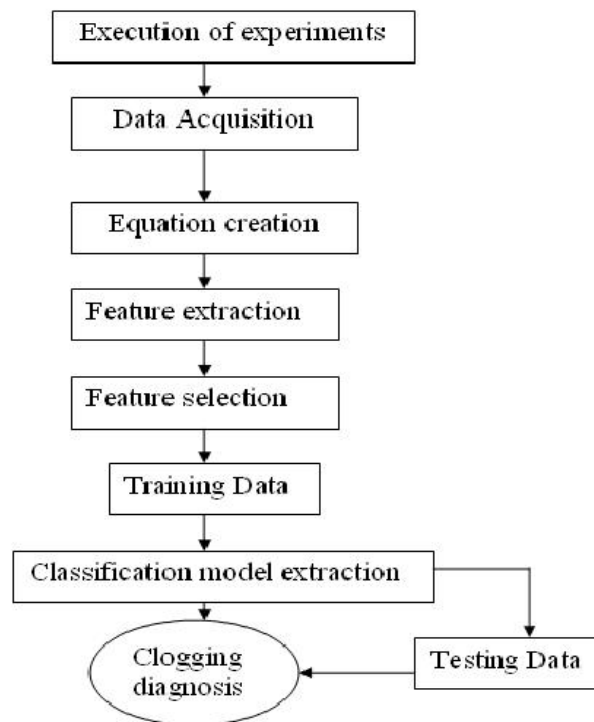
Figure 3. Proposed method based on Fuzzy logic inference system (Mollazade et al, 2008).
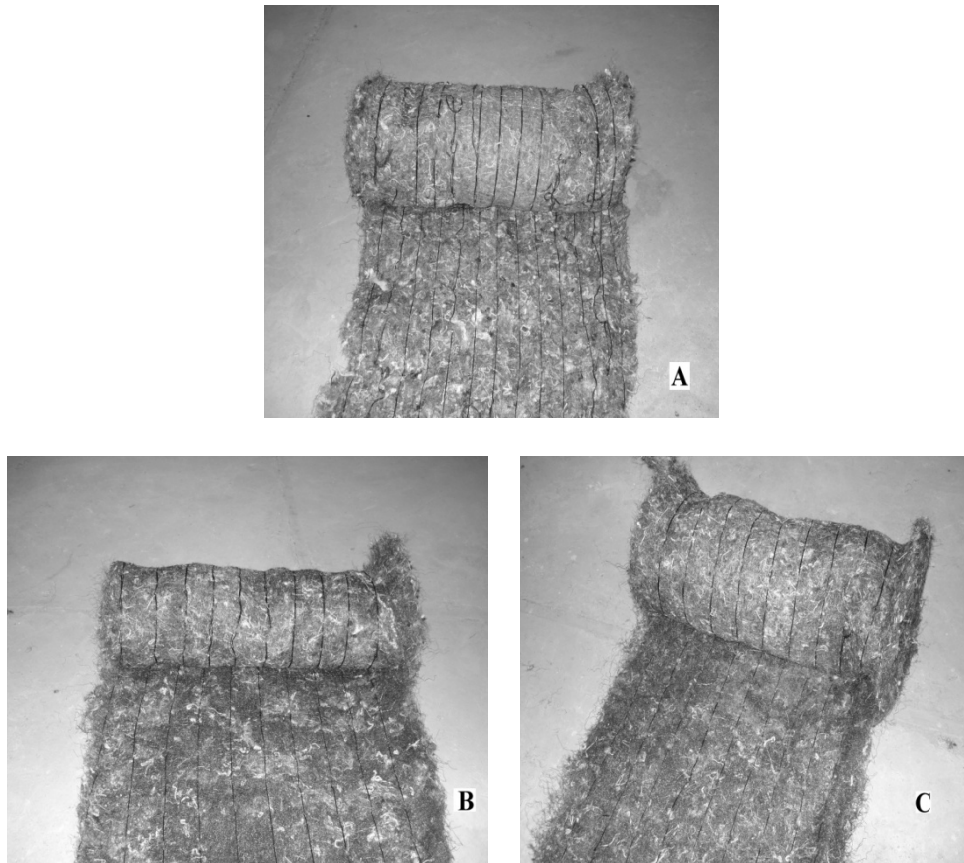
Figure 4. Used synthetic envelopes in this study A) envelope PP450, B) envelope PP700, C) envelope PP900.
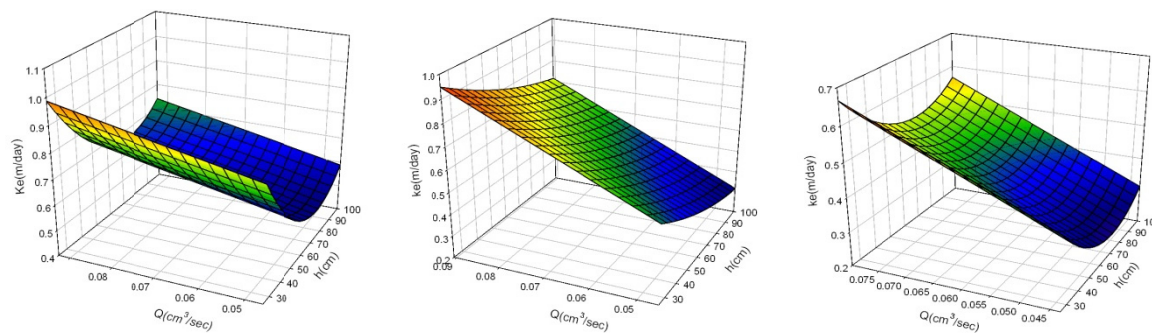
Figure 5. The estimated values for envelope hydraulic conductivity for envelopes pp450, pp700, and pp900 from left to right, respectively.
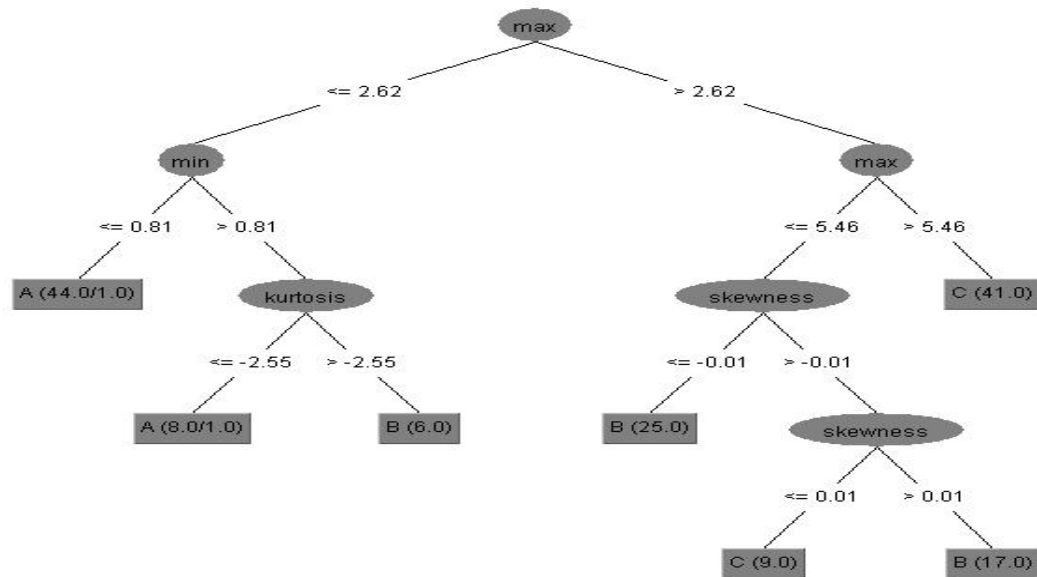
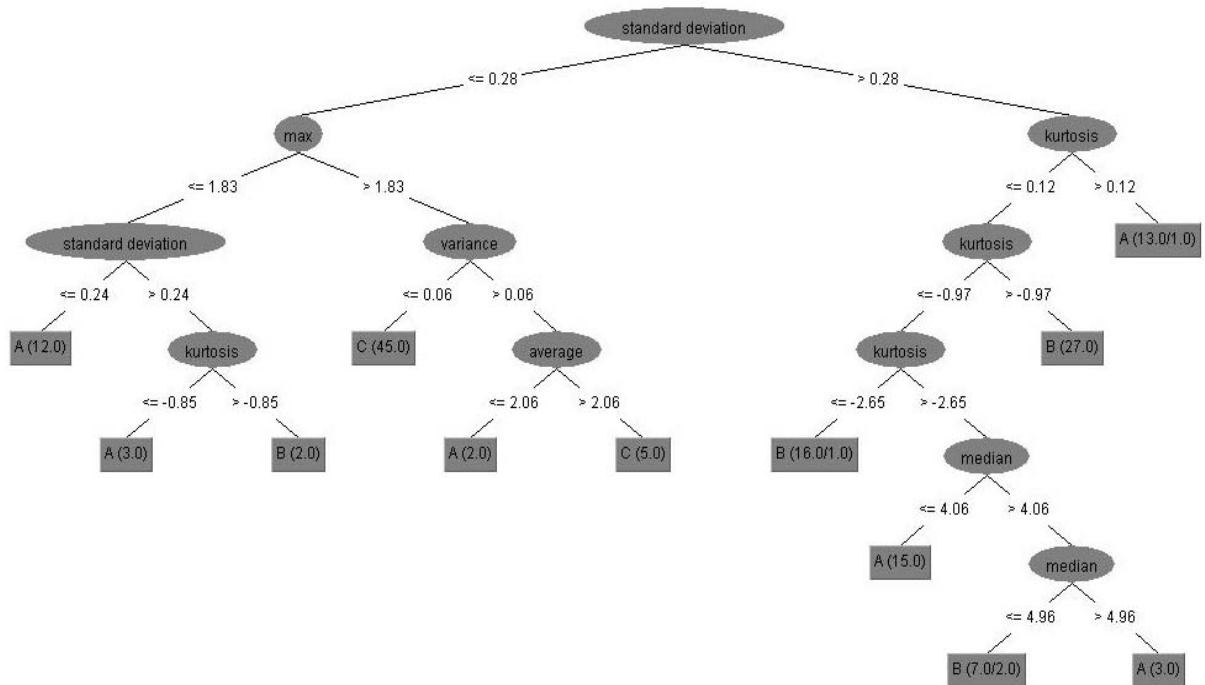Figure 6. Decision tree from J48 algorithm for envelope pp450.

Figure 7. Decision tree from J48 algorithm for envelope pp700.
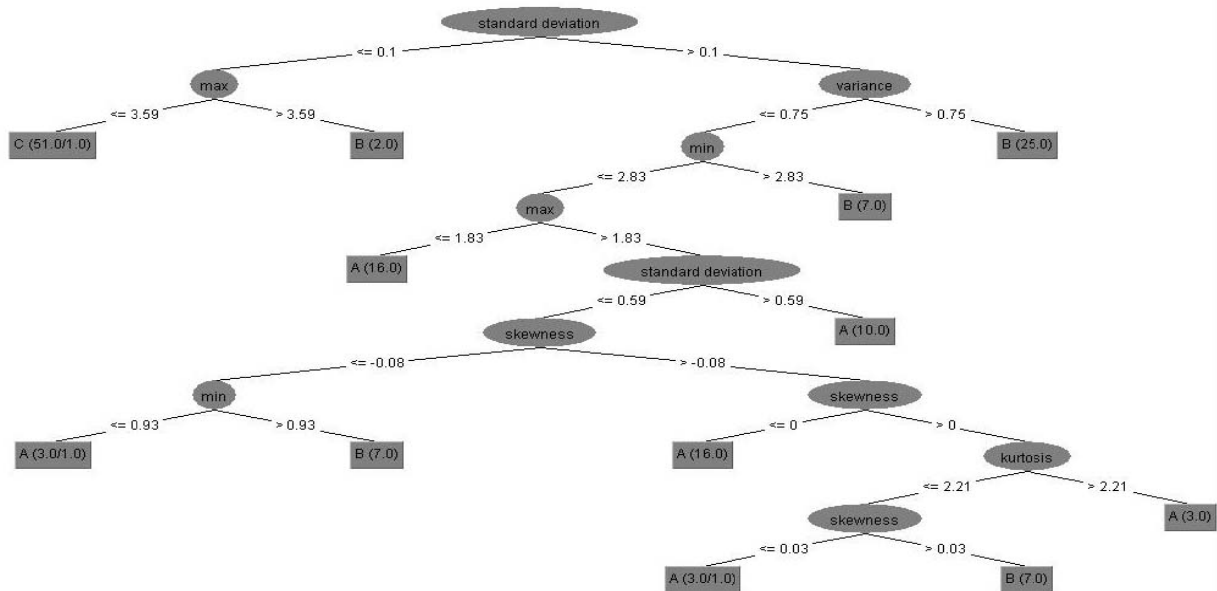
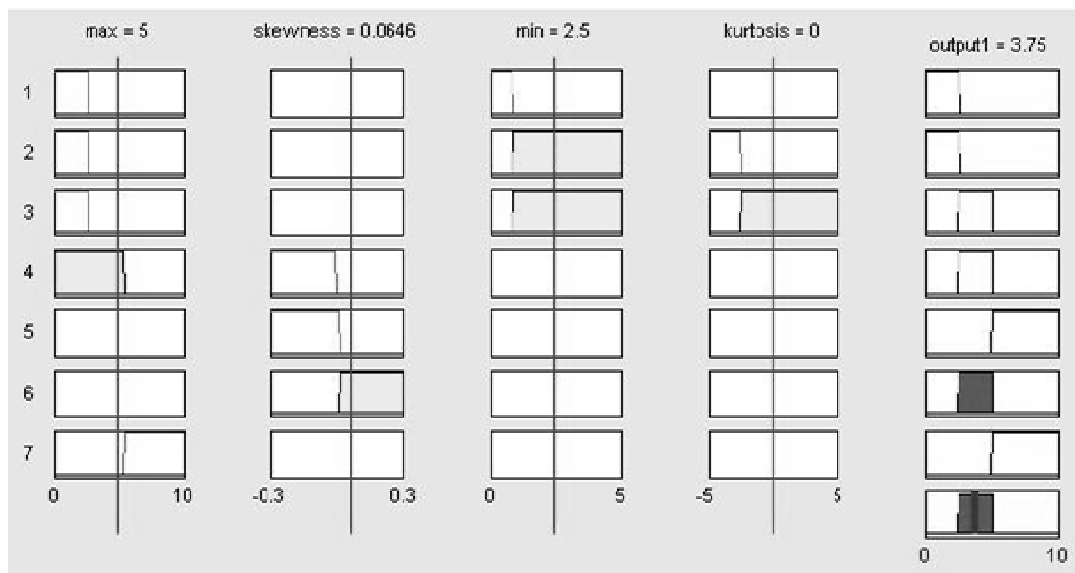Figure 8. Decision tree from J48 algorithm for envelope pp900.

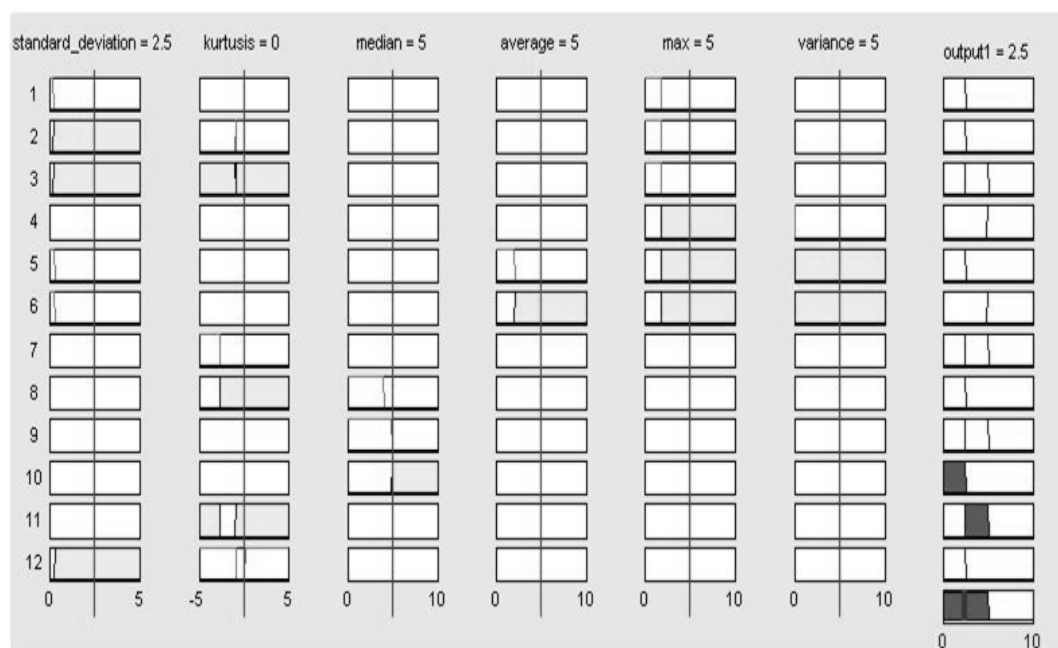Figure 9. Available Rules for one of the test data of envelope PP450.

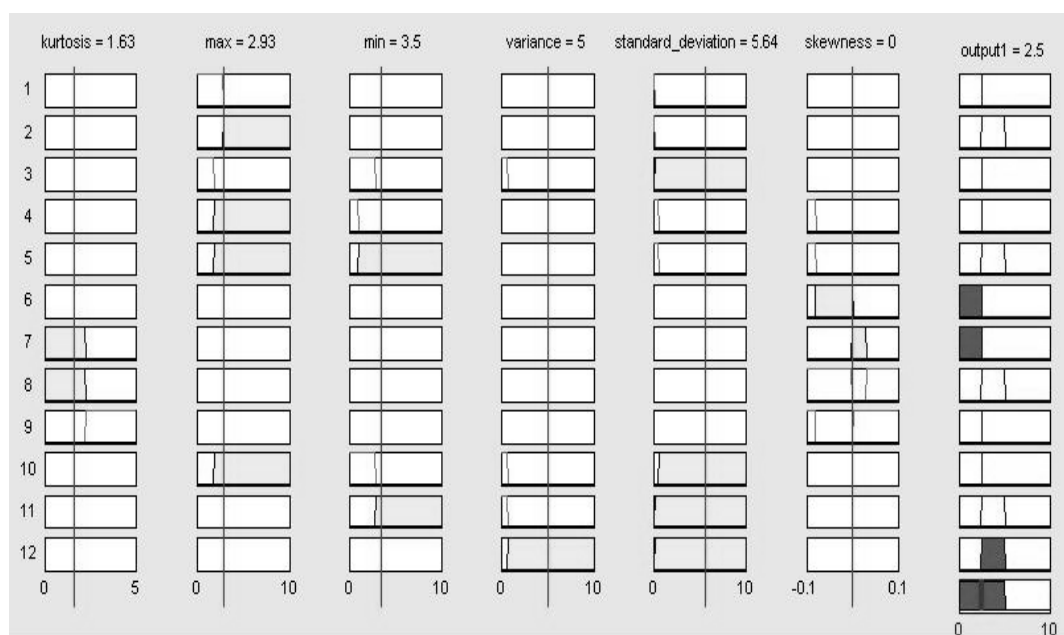Figure 10. Available Rules for one of the test data of envelope PP700

Figure 11. Available Rules for one of the test data of envelope PP900