Australian Journal of Basic and Applied Sciences, 6(1): 190-197, 2012
ISSN 1991-8178

Application of Fuzzy Decision Tree Analysis for Prediction Asphaltene Precipitation Due Natural Depletion; Case Study

[2]Mohammad Ebadi, [1]Mohammad Ali Ahmadi, [2]Kaveh Farhadi Hikoei, [2]Nima Rashidinia

[1]Department of Petroleum Engineering, Science and Research Branch, Islamic Azad University, Tehran, IRAN.
[2]Department of Petroleum Engineering, Ahwaz Faculty of Petroleum Engineering, Petroleum University of Technology, Ahwaz, IRAN.

Abstract: The purpose of this paper is to illustrate how Fuzzy Decision Tree (FDT), which is an automatic method of generating fuzzy rules, can predict the asphaltene precipitation due natural depletion of an under saturated Iranian light petroleum reservoir. Because of the special thermo dynamical conditions of the supposed reservoir, two very important variables consist of Temperature and Pressure, were selected as input factors. In order to develop the model of FDT, firstly, 275 series of data were gathered and divided to two main parts which 201 of them were utilized to build the model and the rest of them to test it. As the FDT method is strongly based on applying widely and effectively the concept of ambiguity and furthermore, to do this project more accurately and less dependent on experts' knowledge, it was decided to gain from piecewise linear membership functions (MFs) whose parameters have automatically been dedicated through calculating a very special method of possibility density function (pdf). When the process of developing the FDT was finished, there were five rules available to measure the rate of compatibility and flexibility of the model by applying the rules on testing set. The model result, 0.66 of R-square for testing set, shows that the FDT yields an acceptable result compared to other methods either practical or theoretical. In conclusion, according to the calculated result, it is possible to exploit this method for asphaltene precipitation prediction field wide.

Key words: Asphaltene, Fuzzy Logic, Decision Tree, Natural Depletion, Temperature, Pressure

INTRODUCTION

The precipitation of heavy organic materials such as asphaltene which is generally defined as a fraction of crude oil that is soluble in toluene or benzene and insoluble in low boiling alkenes such as n-pentane or n-heptane and usually dissolved in crude oil at reservoir initial condition (Kariznovi *et al*., 2006) is one of the most problematic phenomena in upstream oil industries which results in some technical, challenging and demanding problems such as: wettability alteration in reservoir rock, diffusivity reduction and ultimately affects oil production and economical efficiency (Speight, 1991; Galoppini and Tambini , 1994). In more details, the solid precipitation particles of asphaltene forming in reservoir porous media due to pressure depletion results in reducing reservoir rock permeability which definitely threaten the economic oil recovery or noticeably increase production cost (Zahedi *et al*., 2009). As a result, a numerous number of methods either practical or theoretical have been suggested to defeat this obstacle, causing many time and energy consuming problems in crude oil production and residual processing, through predicting the amount of precipitation in assistance with an exact model (Fazlali, 1999; Sözen *et al*., 2004a). Although these models have been a great source in order to do effectively reservoir project modeling or making an attempt to remove the subsequent skin, a fully satisfactory interpretation is still lacking (Sözen *et al*., 2004b). In other words, thanks to the complicated, sophisticated and ambiguous nature of asphaltene which is caused by a large number of parameters affecting its behavior, the conventional methods, commonly including thermo dynamical models (Andersen, and Speight, 1999; Pedersen *et al*., 1989), which have been recommended in order to deal with asphaltene and its consequent struggles are not enough satisfying. This is mainly because of the accuracy which must be considered during measuring different factors involved in calculating asphaltene precipitation. Recently, the new group of asphaltene modeling methods which are based on Artificial Intelligence (AI) approaches has enormously been used by experts. Furthermore, Fuzzy Logic (FL) as another method of AI is also capable of modeling asphaltene precipitation too. This process can be done through fuzzy rules which have been made with some linguistic terms. The supposed rules are generated either experts' knowledge based or automatically. In this research Fuzzy Decision Tree (FDT), as an automatic method, has been selected to generate fuzzy rules and the obtained results have been compared versus gained results from fuzzy logic.

Corresponding Author: Petroleum University of Technology, Ahwaz Faculty of Petroleum Engineering, Kut Abdollah, Ahwaz, Iran,
+989126364936; E-mail: ahmadi6776@yahoo.com

Fuzzy logic:

Discussion about real and current problems of the world we live in is needed to consider accurately and very well the relevant ambiguity, vagueness and uncertainty **(**Liliana *et al*., Mohaghegh, 2000). This difficulty can be solved with a high level of quality by fuzzy logic which was firstly proposed by Lotfi A. Zadeh (1965). In the extended of the binary logic, Fuzzy Logic (FL), every object is a matter of degree which is dedicated by membership functions (MFs). The amount of degree for distinguished situations is always between 0 and 1(Ghafoori *et al*., 2008).

Methodology:

Determining Membership Functions:

In this research there are 3 parameters which are pressure (P), temperature (T) and Asphaltene precipitation. In order to apply fuzzy logic in this case, initially, these factors must be characterized suitably by flexible membership functions. To do this step, 74 of 275 series of data were randomly selected as the training part and the rest of them (201 series) were determined to test the model. Thanks to the having facilitation of a smooth transition between the real world and the fuzzy model (Medasani *et al*., 1998), It was decided to utilize piecewise linear MFs. he general functional form of a piecewise linear MF is given by:

$$\mu_j(x)\begin{cases} 0, & x \le a_{j,1,} \\ \dfrac{x-a_{j,1}}{a_{j,2}-a_{j,1}}, & a_{j,1} < x < a_{j,2,} \\ \dfrac{x-a_{j,2}}{a_{j,3}-a_{j,2}}, & a_{j,2} < x < a_{j,3,} \\ \dfrac{x-a_{j,3}}{a_{j,4}-a_{j,3}}, & a_{j,3} < x < a_{j,4,} \\ \dfrac{x-a_{j,4}}{a_{j,5}-a_{j,4}}, & a_{j,4} < x < a_{j,5,} \\ 0, & a_{j,5} < x, \end{cases} \qquad (2)$$

Fig. 1: General form of linear piecewise membership function.

Where [$\alpha_{j,1}$, $\alpha_{j,2}$, $\alpha_{j,3}$, $\alpha_{j,4}$, $\alpha_{j,5}$] are the values of designing parameters for the MF associated with the jth linguistic term of an attribute. Hence, the associated support for the jth linguistic term is indicated by the interval [$\alpha_{j,1}$, $\alpha_{j,5}$]. Routinely, the constructions of MFs are proposed by experts' knowledge, which can be different in reservoir engineering problems. To defeat this obstacle, it was decided to construct the MFs automatically (Malcolm *et al*., 2004). Thus, it was decided to use training data (201 series of gatherd data) to define the MFs and consequently, model fuzzy decision tree (FDT). Attributes were divided to 3 subsets. It was followed that three intervals were identified for an initial partitioning of each attribute, defined as Low (L), Medium (M) and High (H) levels. Using a simple equal-frequency partitioning (clustering) method was implemented in which the middle point between intervals in different clusters is defined as the initial interval boundary point (Quinlan, 1986).

Table 1: The relevant information on the three clusters for Asphaltene Precipitation

Description	Low(L)	Medium (M)	High (H)
Interval	$L \le 0.0386$ (67)	$0.0663 \le M \le 0.3493$ (67)	$0.4551 \le H$ (67)

When defining the intervals for each of attributes is over, it is essential to construct MFs. To do this step automatically, the proposed approach is to primarily make estimated distributions in the form of probability density functions (pdfs) for each of the intervals of the decision attribute; with the connection between probability distributions and MFs (through possibility distributions) which has been a drawing attention topic since the research of Zadeh (Medasani *et al*., 1998; Thompson and Tapia, 1990). Then, instead of holding full details linked to pdf expressions, in this paper only the centre of area (C-o-A) values for each pdf and the interval boundary values, which were found previously, are used as the key parameters to set MFs.(Malcolm *et al*., 2004)

In order to build the estimated distributions for each of the decision classes, the procedures of (Parzen, 1962; Thompson and Tapia 1990) was exactly followed as bellow and the ensuing results were gained.

$$pdf_j(x) = \frac{1}{\sqrt{2m_j\pi}\left(max(I_j)-min(I_j)\right)}\sum_{i=1}^{m_j} exp\left[-\frac{m_j}{2}\left(\frac{x-x_i}{max(I_j)-min(I_j)}\right)^2\right] \quad (2)$$

Where m_j is the number of values in I_j.

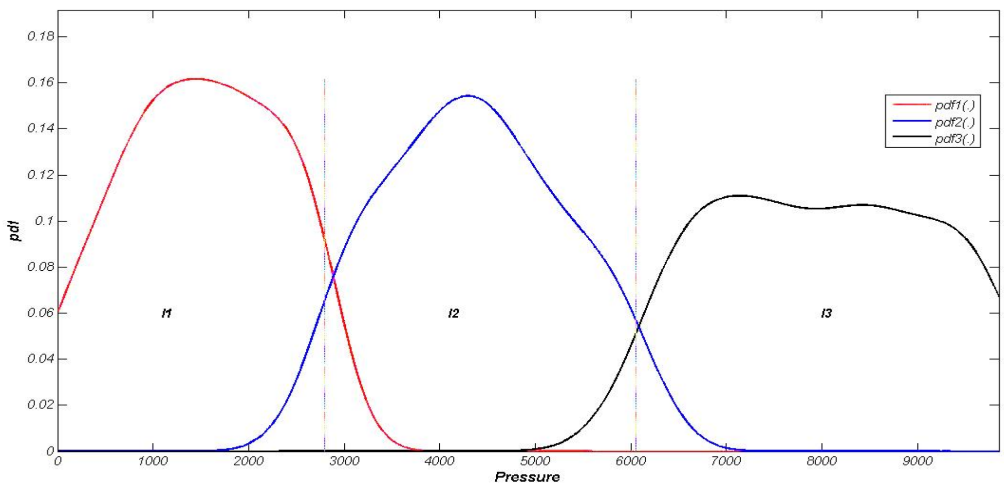

Fig. 2: Estimated distribution related to intervals I1, I2, I3 for Pressure.

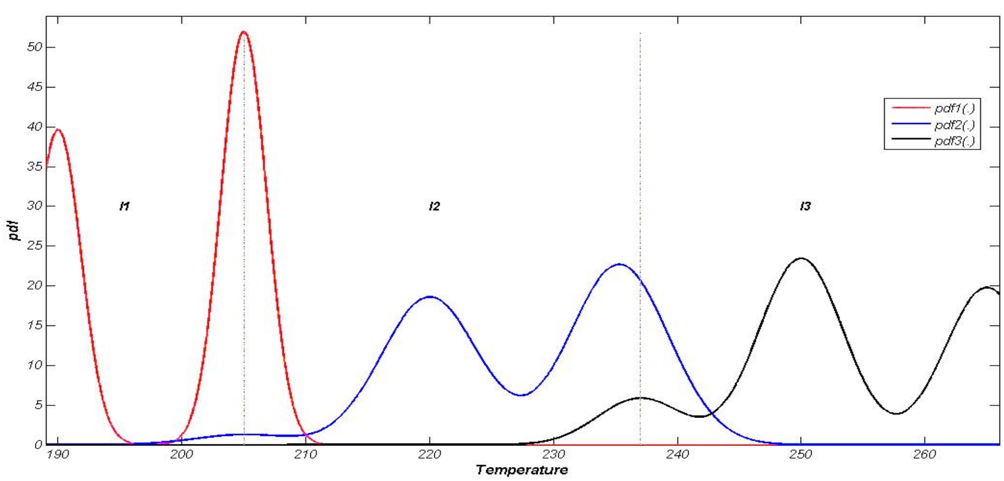

Fig. 3: Estimated distribution related to intervals I1, I2, I3 for Temperature.

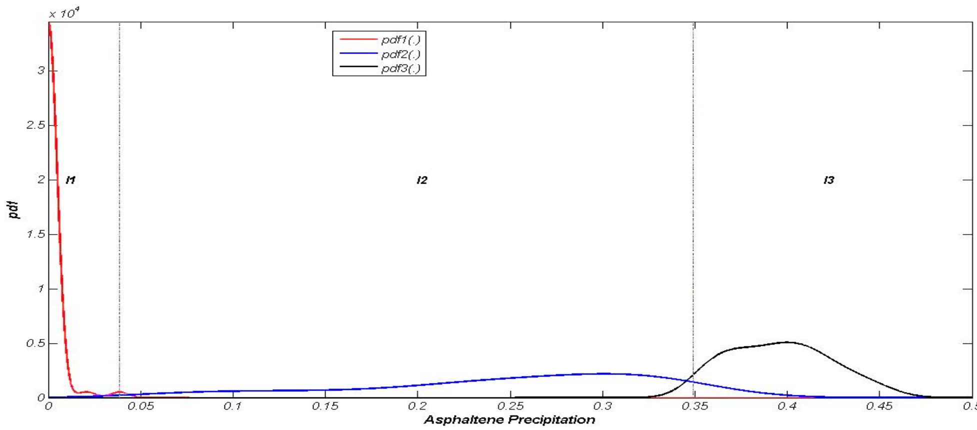

Fig. 4: Estimated distribution related to intervals I1, I2, I3 for Asphaltene Precipitation.

Consequently, based on the boundaries and C-o-As, the MFs were calculated.

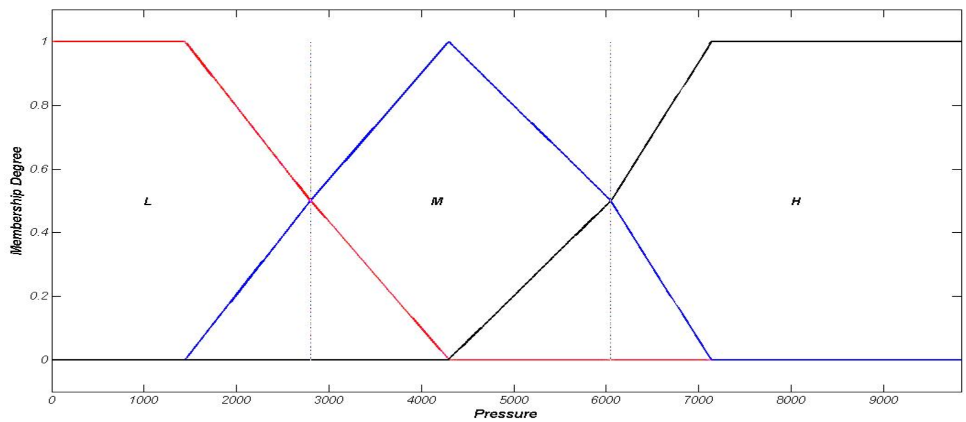

Fig. 5: Membership functions of the decision classes L, M, and H connected with Pressure.

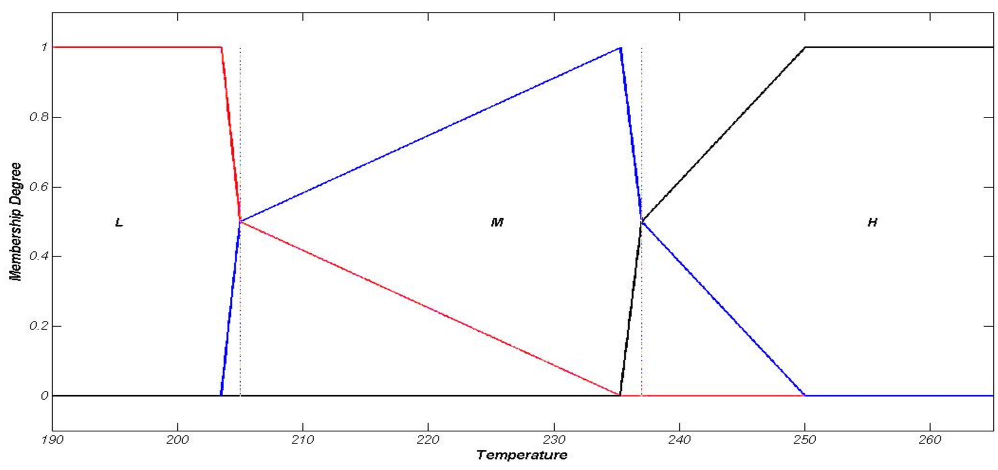

Fig. 6: Membership functions of the decision classes L, M, and H connected with Temperature.

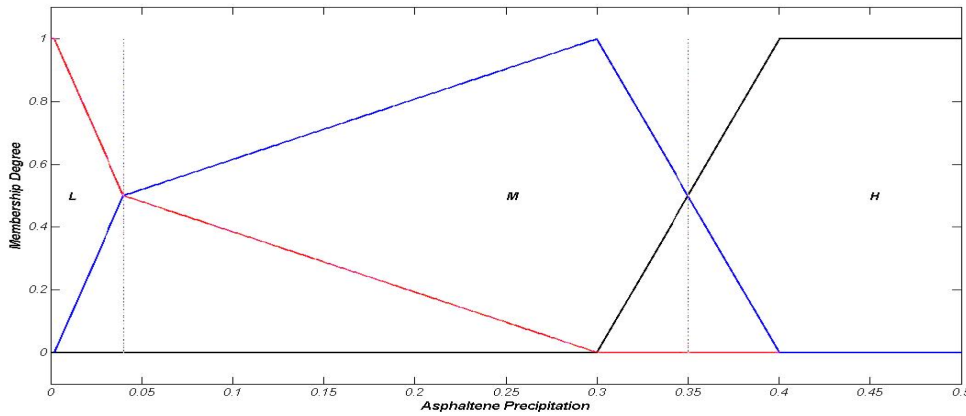

Fig. 7: Membership functions of the decision classes L, M, and H connected with Asphaltene Precipitation.

Description of The Inductive Fuzzy Decision Tree Method:

Applying this method leads to produce automatically series of fuzzy rules which are being characterized with linguistic terms of the classification classes. Its benefits can be summarized in minimizing classification ambiguity and measuring cognitive uncertainties when fuzzy evidences are present (Yuan and Shaw, 1995). A membership function μ(x) of a fuzzy variable Y defined on X, can be assumed as a possibility distribution of Y on X, that is π(x) = μ(x), for all x ⊠ X. The possibility measure $E_\alpha(Y)$ of ambiguity is defined as

$$E_\alpha(Y) = g(\pi) = \sum_{i=1}^{n} (\pi_i^* - \pi_{i+1}^*) \ \ln i \ ,$$

Where $\pi^* = \{\pi_1^*, \pi_2^*, \ldots, \pi_n^*\}$ is the permutation of the possibility distribution $\pi = \{\pi(x_1), \pi(x_2), \ldots, \pi(x_n)\}$ sorted so that $\pi_i^* \geq \pi_{i+1}^*$ for $i = 1, \ldots, n$, and $\pi_{n+1}^* = 0$(Zadeh, 1978; Higashi and Klir, 1983). The ambiguity of attribute *A* is then calculated as

$$E_\alpha(A) = \frac{1}{m}\sum_{i=1}^{m} E_\alpha\big(A(u_i)\big)$$

where $E_\alpha\big(A(u_i)\big) = g\left(\mu T_s(u_i)/max_{1\leq j\leq s}\left(\mu T_j(\mu_i)\right)\right)$, with T_j the linguistic scale used within an attribute for m samples. Overlapping exist when linguistic terms cover each other. The degree which *A* is a subset of *B* is measured through fuzzy subset hood *S(A,B)* and is given by (Kosko, 1986)

$$S(A,B) = \frac{\sum_{u\epsilon U} min\big(\mu_A(u), \mu_B(u)\big)}{\sum_{u\epsilon U} \mu_A(u)},$$

which all attributes are over the same set of objects.
The possibility of classifying an object to class C_i, based on the given fuzzy evidence *E* can be defined as

$$\pi = (C_i|E) = \frac{S(E,C_i)}{max_j\, S(E, C_j)},$$

where $S(E, C_i)$ indicates the degree of truth for the classification rule (that is $E \Rightarrow C_i$). The classification ambiguity according to a single piece of evidence (a fuzzy value for an attribute) is defined as

$$G(E) = g\big(\pi(C|E)\big).$$

$G(P|F)$ is the classification ambiguity with fuzzy partitioning $P = \{E_1, \ldots, E_k\}$ on the fuzzy evidence F which is the weighted average of classification ambiguity with each subset of partition:

$$G(P|F) = \sum_{i=1}^{k} w(E_i|F)G(E_i \cap F),$$

where $G(E_i \cap F)$ is the classification ambiguity with fuzzy evidence $E_i \cap F$, and where $w(E_i|F)$ is the weight $(w(\cdot))$which performs the relative size of subset $E_i \cap F$ in F:

$$w(E_i|F) = \frac{\sum_{u \in U} min\left(\mu_{E_i}(u), \mu_F(u)\right)}{\sum_{j=1}^{k}\left(\sum_{u \in U} min\left(\mu_{E_j}(u), \mu_F(u)\right)\right)}.$$

In short, depend on the lowest level of ambiguity attributes are assigned to nodes. A node is terminated (becomes a leaf node), if degree of subset hood (based on the intersection of the nodes from the root) is higher than a truth value β (in this current paper it has been equated to 0.7) (Yuan and Shaw, 1995; Wang *et al*., 2000; Malcolm *et al*., 2004)

After calculating the mentioned method, the succeeding results obtained as shown below:

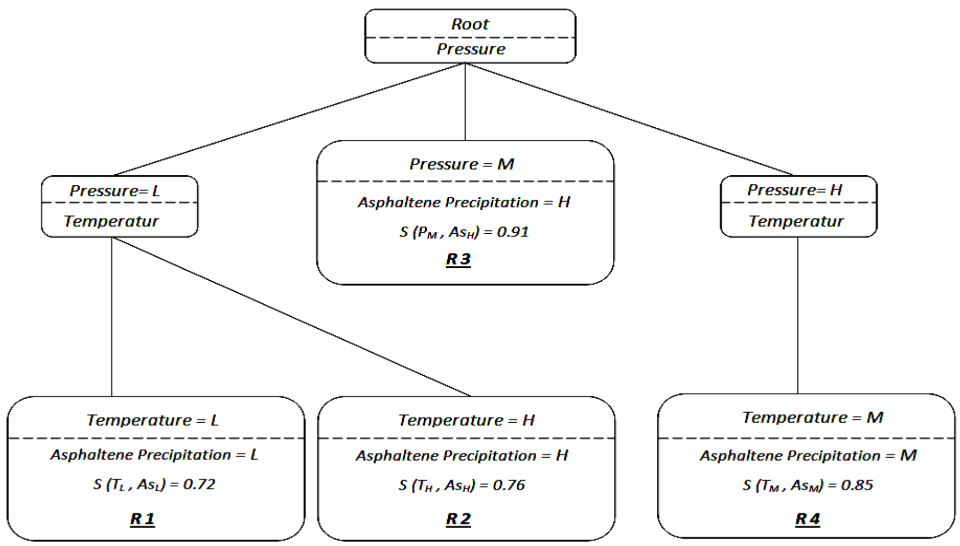

Fig. 8: Schematic structure of fuzzy decision for Asphaltene Precipitation.

Table 2: Class ambiguity values for each input attribute.

Attribute	Temperarure	Pressure
G(E)	0.8256	0.4894

After finishing the process of constructing the FDT and extracting rules, 74 series of data in the set of testing were feed to the model and R-square of 0.660 was calculaterd for the responses.

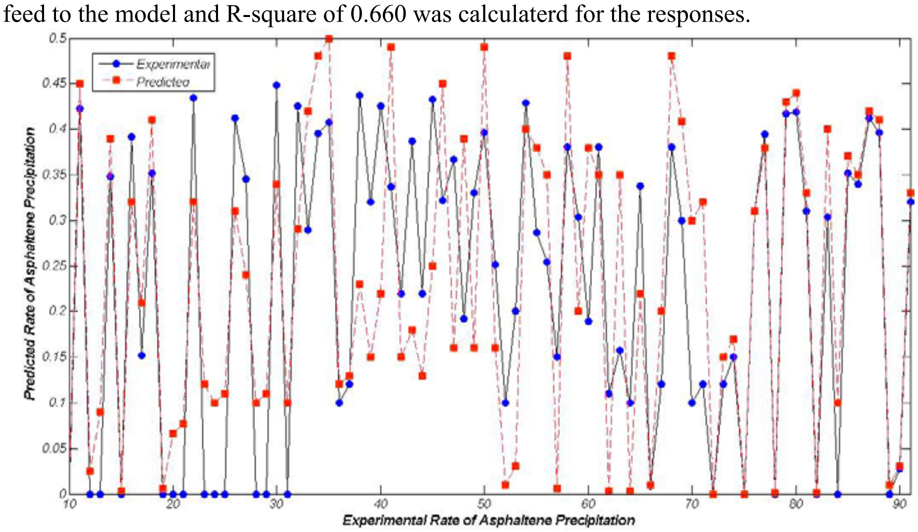

Fig. 9: The comparison between predicted and Experimental values of Asphaltene Precipitation for each sample of testing set.

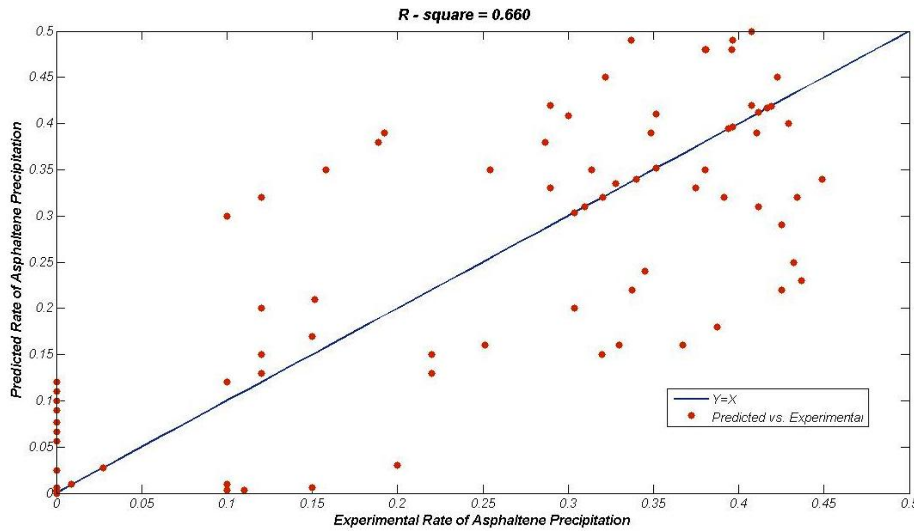

Fig. 10: The performance of predicted values of Asphaltene Precipitation vs. the measured

Conclusions:

Based on results obtained from this study following conclusions can be drawn:

However, it is acknowledged that in the current research the decision rules were restricted to three outcomes, identifying only low, medium and high fee levels. Clearly, if more intervals are specified more precision can be obtained with reference to the subsequent decision outcomes, depending on the particular application and the degree of clarity of linguistic interpretation required. Future research could examine the sensitivity of the results, by extending the analysis to petroleum applications, and in terms of increasing the classes (number of MFs) describing each attribute, together with the type of MF selected and the loss function associated with model errors.

This paper is concerned with the construction of a set of fuzzy rules using typical data employed in one of the northern Persian Gulf oil field. Central to the underlying philosophy towards this method has been the automation of the process to construct the fuzzy rules which limit the parameters which must be pre-specified by the expert (i.e. truth levels and number of membership functions).

REFERENCES

Andersen, S.I., JM. Speight, 1999. Thermodynamic models for asphaltene solubility and Precipitation. J. Petr. Sci. Eng, 22: 53-66.

Fazlali, A., 1999. The asphaltene precipitation in crude oil of Iran, PhD Thesis , Amir Kabir University, Iran,

Galoppini, G., M. Tambini, 1994. Asphaltene Deposition Monitoring and Removal Treatments: An Experience Deep Wells. Soc Pet Eng; (SPE27622)

Ghafoori, M.R., M. Roostaeian and V.A. Sajjadian, 2008. "A State-of-Art Permeability Modeling Using Fuzzy Logic in a Heterogeneous Carbonate (An Iranian Carbonate Reservoir Case Study)," IPTC 12019 presented at the International Petroleum Technology Conference held in Kuala Lumpur, Malaysia, 3-5 December 2008.

Higashi, M., G.J. Klir, 1983. Measures of uncertainty and information based on possibility distributions. International Journal of General Systems, 9: 43-58.

iliana, P.M., R.G. Hughes and M.L. Wiggins, "Identification and Characterization of Naturally Fractured Reservoirs Using Conventional Well Logs," The University of Oklahoma.

Kariznovi, M., H. Nourzieh, M. Jamialahmadi and A. Shahrabadi, 2008. "Optimization of Asphaltene Deposition and Adsorption Parameter in Porous Media Search" paper SPE 114037 presented at the SPE 2008 Western Regional and Pacific Section AAPG Joint Meeting held in Bakersfield, California, U.S.A., 31 March-2 April 2008.

Kosko, B., 1986. Fuzzy entropy and conditioning. Information Science, 30: 165-74.

Malcolm, J., Beynon, Michael J. Peel, 2004. Yu-Cheng Tang. The application of fuzzy decision tree analysis in an exposition of the antecedents of audit fees. The International Journal of Management Science, 32: 231-244.

Medasani, S., J. Kim, R. Krishnapuram, 1998. An overview of membership function generation techniques for pattern recognition. International Journal of Approximate Reasoning, 19: 391–417.
Medasani, S., J. Kim, R. Krishnapuram, 1998. An overview of membership function generation techniques for pattern recognition. International Journal of Approximate Reasoning, 19: 391–417.
Mohaghegh, S., 2000. "Virtual Intelligence Applications in Petroleum Engineering: Part 3-Fuzzy Logic," JPT (November 2000) 82.

Parzen, E., 1962. On estimation of a probability density function mode. Annals of Mathematical Statistics., 33: 1065-76.

Pedersen, K.S., A. Fredenslund, P. Thomassen, 1989. Properties of Oils and Natural Gases. Houston: Gulf Publishing.

Quinlan, J.R., 1986. Induction of decision trees. Machine Learning, 1: 81-106.

Sözen, Adnan, Özalp, Mehmet, Arcakioglu, Erol, 2004a. Formulation based on artificial neural network of thermodynamic properties of ozone friendly refrigerant/absorbent couples. Appl. Therm., Eng. 25.

Sözen, Adnan, Özalp, Mehmet, Arcakioglu, Erol, 2004b. Investigation of thermodynamic properties of refrigerant/absorbent couples using artificial neural networks. Chemical Engineering and Processing, 43: 1253-1264.

Speight, J.G., 1991. The chemistry and technology of petroleum. 2nd ed. New York: Marcel Dekker; 1991.

Thompson, J.R., R.A. Tapia, 1990. Nonparametric function estimation, modeling, and simulation. Philadelphia: Society for Industrial and Applied Mathematics.

Wang, X., B. Chen, G. Qian, F. Ye, 2000. On the optimization of fuzzy decision trees. Fuzzy Sets and Systems, 112: 117-25.

Yuan, Y., M.J. Shaw, 1995. Induction of fuzzy decision trees. Fuzzy Sets and Systems, 69: 125-39.
Yuan, Y., M.J. Shaw, 1995. Induction of fuzzy decision trees. Fuzzy Sets and Systems, 69: 125-39.
Zadeh, L.A., 1965. "Fuzzy Sets," Information and Control., 8: 338.

Zadeh, L.A., 1978. Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems, 1: 3-28.

Zahedi, G., A.R. Fazlali, S.M. Hosseini, G.R. Pazuki, L. Sheikhattar, 2009. Prediction of asphaltene precipitation in crude oil. Journal of Petroleum Science and Engineering, 68: 218-222.