Johanyák, Zs. Cs., Kovács, Sz.: Fuzzy modeling of Petrophysical Properties Prediction Applying RBE-DSS and LESFRI, International Symposium on Logistics and Industrial Informatics (LINDI 2007), September 13-15, 2007, Wildau, Germany, pp.87-92.

Fuzzy modeling of Petrophysical Properties Prediction Applying RBE-DSS and LESFRI

Zsolt Csaba Johanyák[1], Szilveszter Kovács[2]
Institute of Information Technologies
Kecskemét College, GAMF Faculty
Kecskemét, Hungary
[1]johanyak.csaba@gamf.kefo.hu, [2]kovacs.szilveszter@gamf.kefo.hu

Abstract—**The appearance of Fuzzy Rule Interpolation (FRI) techniques in practical applications gains more and more importance for automatic identification of sparse fuzzy rule bases from given training data. This paper reports the generation of a fuzzy system, which models the relation between different oil well data aiming the prediction of petrophysical properties. The applied rule base generation method is RBE-DSS [8] and the fuzzy inference was performed by the technique LESFRI [6].**

Keywords-Fuzzy Rule Interpolation (FRI), Fuzzy Rule Base Identification, LESFRI, RBE-DSS

I. INTRODUCTION

The modeling of the functional relationship between input and output data represents a wide application area for fuzzy systems. The development of Fuzzy Rule Interpolation based Inference Techniques (FRITs) opened new horizons on this field for practical applications due to the reduced complexity and storage space demand as well as due to its ability to handle cases when there is no experimental information that would describe the expected output for all the possible inputs.

The prediction of petrophysical properties as part of reservoir evaluation is an important supporting tool in taking decisions on rentability of the exploration of a specific region. The collection of experimental data (borehole drilling, sampling and extensive laboratory analysis) is expensive. Therefore the FRIT based fuzzy modeling of petroleum well data could be very advantageous.

The rest of this paper is organized as follows. Section II. recalls the basic concepts of sparse rule bases and the inference technique LESFRI applied as FRIT. Section III. overviews briefly the backgrounds of the modeled phenomena. Section IV. presents the applied rule base generation method and the used performance index. The experimental results are discussed in section V.

II. FUZZY RULE INTERPOLATION BASED REASONING

A. Sparse Rule Base and FRI based reasoning

The classical fuzzy reasoning methods (e.g. Zadeh's [23], Mamdani's [14], Larsen's [13], Takagi-Sugeno's [15], etc.) require a full coverage of the input space by the rules of the rule base in order to ensure an acceptable output for each possible input (observation) of the system. This feature also called dense character of the rule base can be expressed by the condition

$$\min_{i=1}^{N}\left(\max_{j=1}^{n_i}\left(t\left(A_{ij},A_i^*\right)\right)\right)\geq \varepsilon,\ \forall A_i^*\subset U_i, \qquad (1)$$

where U_i is the i^{th} dimension of the antecedent space, A_i^* is the fuzzy set describing the observation in the i^{th} antecedent dimension, $0<\varepsilon\leq 1$ is an arbitrary constant, A_{ij} is the j^{th} linguistic term of the i^{th} antecedent dimension, *t* is an arbitrary t-norm, n_i is the number of the fuzzy sets in the i^{th} antecedent dimension and *N* is the number of antecedent dimensions.

The fulfillment of condition (1) results in an explosive increase of the number of rules. This higher system complexity implies increased memory demand and slows down the inference process.

Rule bases not fulfilling condition (1) are called sparse ones. For example fig. 1 illustrates the antecedent space of a fuzzy system having two input dimensions (A_1 and A_2). The rule base is sparse, it contains only three rules. The antecedent parts of the rules are represented by pyramids, which are defined by trapezoid shaped linguistic terms. In case of the observation *A** none of the rule antecedents intersect the pyramid representing the input of the system and therefore none of the classical fuzzy inference techniques can produce an acceptable conclusion.

Fuzzy systems based on sparse rule bases and applying

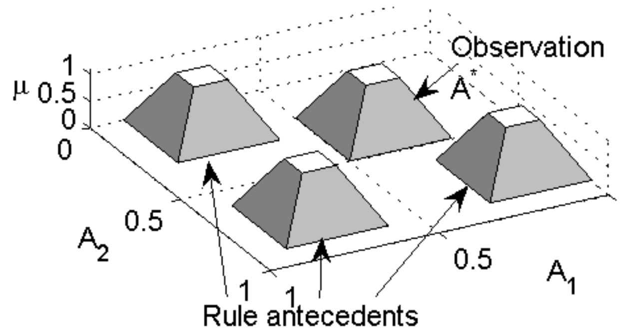

Figure 1. Antecedent space of a sparse rule base

Research partly supported by the Kecskemét College, GAMF Faculty grant no: 1KU08/2007.

FRITs can solve the problem of complexity explosion. Contrary to classical fuzzy reasoning methods they do not require the dense character of the rule base. They determine the conclusion taking into consideration two or more rules whose antecedent parts are in the neighborhood of the observation.

FRITs can be divided into two groups depending on whether they are producing the estimated conclusion directly or they are interpolating an intermediate rule first.

Relevant members of the first group are among others the linear rule interpolation (KH method) [11] proposed by Kóczy and Hirota, which is the first developed one, the MACI (Tikk and Baranyi) [18], the FIVE [10] introduced by Kovács and Kóczy, the IMUL proposed by Wong, Gedeon and Tikk [21], the method based on the conservation of the relative fuzziness suggested by Kóczy, Hirota and Gedeon [12] and the interpolative reasoning based on graduality introduced by Bouchon-Meunier, Marsala and Rifqi [2].

The methods belonging to the second group follow the concepts laid down by the generalized methodology (GM) defined by Baranyi et al. in [1]. Typical members of this group are e.g. the technique family proposed by Baranyi et al. in [1], the ST method [22] suggested by Yan, Mizumoto and Qiao and the IGRV [3] developed by Huang and Shen as well as the techniques LESFRI [6], FRIPOC [5] and VEIN [9] developed by Johanyák and Kovács.

B. LESFRI

In course of the fuzzy modeling we applied the method LESFRI [6] as fuzzy inference technique. It belongs to the group of two-step fuzzy rule interpolation techniques. In the first step it determines a new rule whose antecedent sets are situated in the same position as the sets describing the observation, i.e. their reference point are identical in each antecedent dimension.

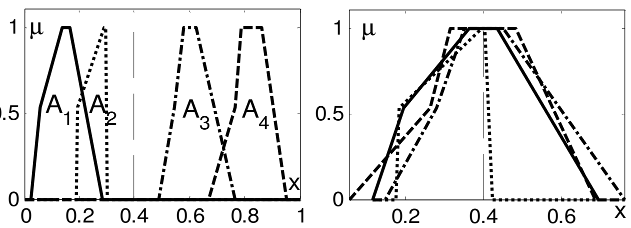

Figure 2. The original partition and the shifted linguistic terms

The shape of the antecedent and consequent linguistic terms is calculated by the set interpolation technique FEAT-LS, which is based on the concept of linguistic term shifting and polar cuts. Its key ideas are recalled in section *II.B.1*. The position of the consequent sets is determined by an adapted version of the Sherpard interpolation [16]. The shape of the conclusion is calculated by the method SURE-LS also based on the concept of least squares. It is presented shortly in section *II.B.2*.

1) FEAT-LS

Fuzzy Set Interpolation (FSI) aims the determination of a new linguistic term in a given point of a fuzzy partition called interpolation point. This means that the new fuzzy set is generated in such way that its reference point coincides with the interpolation point. In case of two-step FRI methods an FSI technique is used for the calculation of the antecedent and consequent sets of the new rule. Thus the interpolation point is either the reference point of the observation or the reference point of the consequence in the current dimension. An FSI technique works only with one partition. Therefore the calculations in the different dimensions in both the antecedent and consequent cases can be done separately.

The Fuzzy sEt interpolAtion Technique based on the method of weighted Least Squares (FEAT-LS) [6] was developed especially for the case when all sets of a partition belong to the same shape type and the characteristic (break) points are also situated at the same α-level. In such cases it seems to be a natural condition on the new linguistic term created in the interpolation point to suit this regularity as well.

As a first step all the sets of the partition are shifted horizontally in order to reach the coincidence between their reference points and the interpolation point. The left part of fig. 2 presents an example for a partition containing four linguistic terms and an interpolation point at $x^i=0.4$. The effect of the shifting is presented on the right part of the figure.

The effect of the shifting is not permanent. It is only used during the determination of the new set. Next the shape of the new linguistic term is calculated from the overlapped set shapes in a set form that belongs to the characteristic shape type of the partition (e.g. singleton, triangle, trapezoid, polygonal, etc).

The characteristic points of the shape are determined by the method of weighted least squares taking into consideration the corresponding characteristic points of the overlapped sets. The weighting expresses that the sets situated originally in closer neighborhood of the interpolation point should exercise a higher influence than those situated originally in farther regions of the partition.

2) SURE-LS

The revision method SURE-LS (Single rUle REasoning based on the method of Least Squares) [6] was developed for the case when all linguistic terms of a consequent partition belong to the same shape type (e.g. singleton, triangle, trapezoid, etc.) and all characteristic (break) points are situated at the same α-levels. It also means that the height of all sets is the same.

Thus the method seeks a special shape form with the having predetermined ordinate values of the characteristic points. Therefore only the abscissas of the characteristic points have to be calculated.

SURE-LS applies an α-cut based approach for this task. It uses a set of α-levels compiled together by taking into consideration the break-point levels of all antecedent dimensions and the current consequent partition. The calculations are done separately for the left and right flanks. On each side for each level it calculates the weighted average of the distances between the endpoints of the α-cuts of the rule antecedent and the observation set. The weighting makes possible to take into consideration the different antecedent dimensions (input state variables) with different influence.

The basic idea of the method is the conservation of the weighted average differences measured on the antecedent side. Applying these modifications on the consequent side usually results in a set of characteristic points that do not fit the default set shape type of the partition. Therefore the method of Least Squares is used in order to find the break-points of an acceptable conclusion.

III. Petrophysical Properties

In order to prove the practical applicability of the method pair RBE-DSS as rule base generation tool and LESFRI as fuzzy inference technique we chose a real world problem introduced in [20].

One of the key tasks in course of the analysis of petroleum well log data is the prediction of petrophysical properties corresponding to specific input data, i.e. depth values different from the original ones used by the experiments. Such properties are the porosity, permeability and volume of clay [20]. The expensive and time consuming character of the data collection from boreholes increases the significance of the prediction. The predicted values help taking decisions on rentability of the exploration of a specific region.

Our research task was to create a fuzzy model with low complexity that is applicable for the prediction of porosity (PHI) based on well log data described by three input variables. These are the gamma ray (GR), deep induction resistivity (ILD) and sonic travel time (DT).

IV. Fuzzy Modeling

1) *System Generation using RBE-DSS*

The rule base of the fuzzy model was generated by the method Rule Base Extension using Default Set Shapes (RBE-DSS) introduced in [8]. In the followings we recall briefly the basic ideas of the method.

The key idea of the method is that after defining a default set shapes for fuzzy linguistic terms for each input and output dimension separately, one creates two rules that fit (describe) the minimum and maximum output. The default set parameters are dependent on the range of the actual linguistic variable.

Next a tuning algorithm starts aiming the identification of the parameters of the initial fuzzy sets. This algorithm uses an iterative approach adjusting each parameter in several steps separately. The system is evaluated in each iteration step for different parameter values against a training data set and the parameter values ensuring the best performance index are kept for the next iteration.

If the decreasing velocity of the performance index of the system is too slow, i.e. it falls below a specified threshold after two consecutive iterations a new rule is generated. It is because the system tuning reached a local or global minimum of the performance index and the performance cannot ameliorate further by the applied parameter identification algorithm. The new rule introduces additional tuning possibilities. However, in some cases the performance index will increase temporarily after the insertion of the new rule into the rule base (see fig. 3).

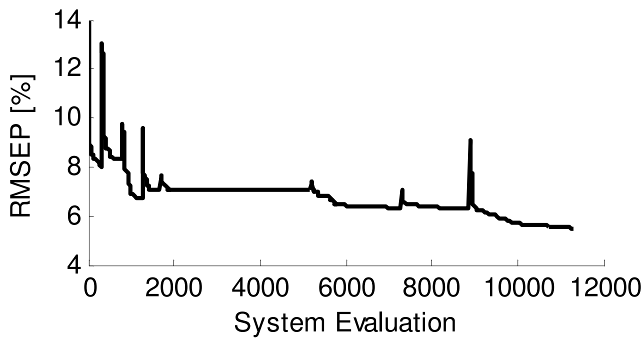

Figure 3. Variation of the performance index in course of the tuning

In order to create the new rule, one seeks for the calculated data point, which is the most differing one from its corresponding training point. The input and output values of this training point will be the reference points of the antecedent and consequent sets of the new rule. The shape of the new linguistic terms is determined by using default core and width values.

2) *Performance Index*

In course of the parameter identification process after each parameter adjustment the resulting parameter set is evaluated by calculating the system output for a collection of predefined input data, for which the expected output values are known. In order to compare the results obtained with different parameter sets a performance index is calculated after each system evaluation.

Our algorithm uses the relative value of the root mean square (quadratic mean) of the error (RMSEP) as performance index for the evaluation of the fuzzy system. We chose it owing to its easy interpretability. The value of the root mean square (quadratic mean) of the error [19] is calculated by

$$RMSE = \sqrt{\frac{\sum_{j=1}^{M}\left(y_j - \hat{y}_j\right)^2}{M}}, \qquad (2)$$

where M is the number of training data points, y_j is the output of the j^{th} data point and $\hat{y}_j$ is the output calculated by the system.

The relative value of *RMSE* to the range (*RMSEP*) expressed in percentage is determined by

$$RMSEP = \frac{RMSE}{DR} \cdot 100, \qquad (3)$$

where DR is the range of the output dimension. The application of *RMSEP* as performance index makes possible the parameter identification by our algorithm even in case of

Multiple Input Multiple Output (MIMO) systems. In that case the resulting performance index is calculated using the formula

$$RRMSEP = \sqrt{\sum_{l=1}^{n_{out}} RMSEP_l^2}\ , \tag{4}$$

where and n_{out} is the number of the output dimensions.

V. Experimental Results

For the sake of easier comparability to a known FRI Fuzzy model identification method (discussed in [20]) the same training and testing data sets were used as it was introduced in [20]. The training data set consisted of 71 data points and the testing data set consisted of 51 data points. The data were preprocessed and each variable was normalized to the unit interval.

The applied inference technique was LESFRI [6] combined with the COG defuzzification and we used RBE-DSS [4] for system generation and tuning. The antecedent and consequent partitions of the final system are presented on figures 4 and 5

The fuzzy system was generated using RMSEP (3) as performance index. In order to compare the results with those published in [20] we also evaluated the final system against the training and testing data set by the correlation factor (5), which was used in [20] as a prediction accuracy indicator.

$$R = \frac{\sum_{j=1}^{M}\left(y_j - \bar{y}\right)\cdot\left(\hat{y}_j - \bar{\hat{y}}\right)}{\sum_{j=1}^{M}\left(y_j - \bar{y}\right)^2 \cdot \sum_{j=1}^{M}\left(\hat{y}_j - \bar{\hat{y}}\right)^2} \tag{5}$$

In case of both (training and testing) data sets our system showed a slightly better performance. Table I. presents the correlation factor values obtained after the evaluation of our system and those published in [20].

TABLE I. Correlation Factor Values

Applied Method	Correlation Factor	
	Training data	***Testing data***
MACI [20]	0.917	0.865
RBE-DSS + LESFRI	0.934	0.890

Another advantage of our system is that the number of linguistic terms and rules is significantly reduced in comparison to [20]. For example while the system presented in [20] was based on 63 rules, our version contains only 9 rules. However, it should be mentioned as a drawback that the shape of the membership functions of the input and output partitions (presented on figures 4. and 5.) are not so nice and uniform as the triangle shaped ones introduced in [20].

This feature can be traced back to the applied tuning algorithm (see. Section *IV.1*) that adjusts the break-points of the linguistic terms used as parameters one-by-one.

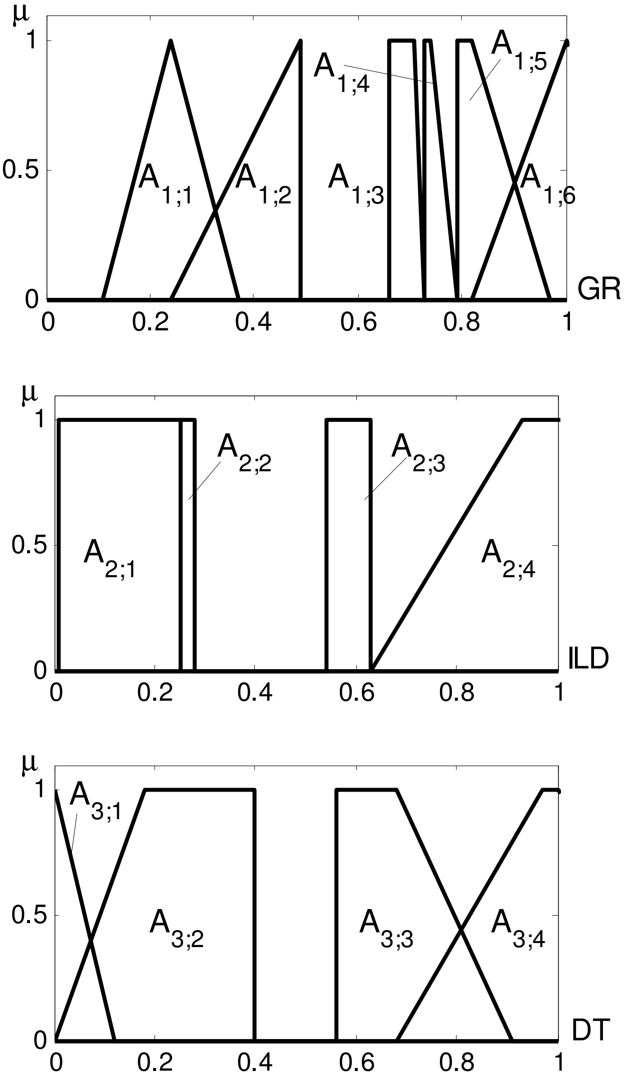

Figure 4. Membership functions of the linguistic terms participating in rule antecedents

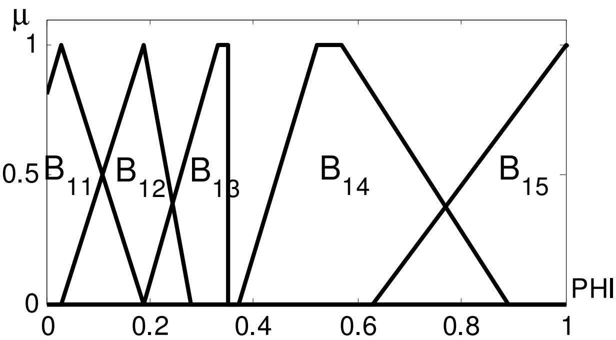

Figure 5. Output partition of the tuned system

We used trapezoidal linguistic terms owing to its description capability for the case of the most frequent used singleton, crisp, triangle and trapezoidal fuzzy set shape types.

Figure 3 presents the variation of the performance index (RMSEP) in course of the tuning. The horizontal axis corresponds to the number of system evaluations.

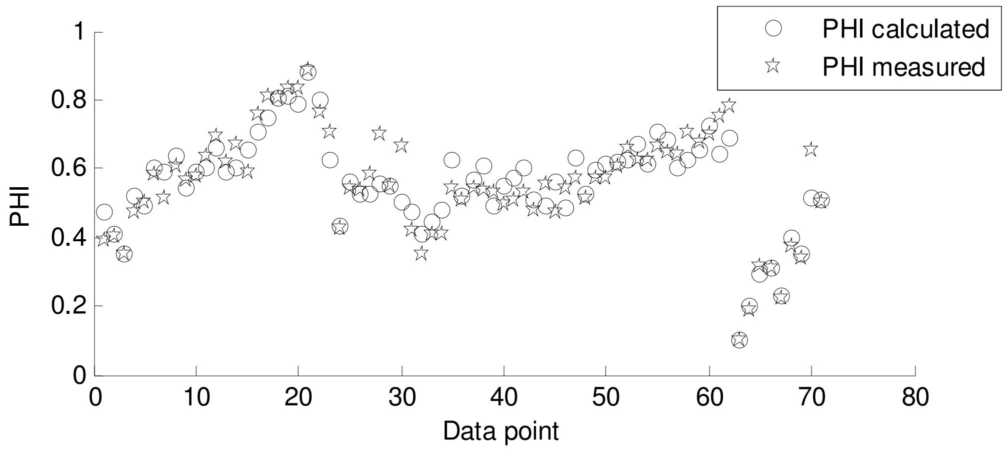

Figure 6. Measured and calculated data points obtained in case of the training data set

The generation of the seven new rules is indicated by good observable peak points.

Figure 6 presents the measured and calculated data points obtained in case of the training data set. Owing to the fact that the system has three input dimensions the calculated and measured data can be visualized only by a 2D plot where the horizontal axis represents the ordinal number of the data points and the vertical axis corresponds to the calculated and measured output values.

VI. CONCLUSIONS

This paper presented the implementation details of a fuzzy system that was generated automatically from available training data by the help of the method RBE-DSS and the inference technique LESFRI. The system was also validated again a testing data set.

The results of the experiments proved the practical applicability of the chosen methods showing better prediction accuracy than reported before in the literature. The calculations were carried out by the help of the Fuzzy Rule Interpolation and Fuzzy Rule Generation Matlab ToolBoxes available at [24].

REFERENCES

[1] P. Baranyi, L. T. Kóczy and T. D. Gedeon, "A Generalized Concept for Fuzzy Rule Interpolation," In IEEE Transaction On Fuzzy Systems, ISSN 1063-6706, Vol. 12, No. 6, 2004. pp 820-837.

[2] B. Bouchon-Meunier, C. Marsala and M. Rifqi, "Interpolative reasoning based on graduality," In Proc. FUZZ-IEEE'2000, 2000, pp. 483-487.

[3] Z. H. Huang, and Q. Shen, "Fuzzy interpolation with generalized representative values," In Proceedings of the UK Workshop on Computational Intelligence, pp. 161-171, 2004.

[4] Z. C. Johanyák and S. Kovács, "The effect of different fuzzy partition parameterization strategies in gradient descent parameter identification," 4th International Symposium on Applied Computational Intelligence and Informatics (SACI 2007), May 17-18, 2007 Timisoara, Romania, pp. 141-146.

[5] Z. C. Johanyák and S. Kovács, "Fuzzy Rule Interpolation Based on Polar Cuts," Computational Intelligence, Theory and Applications, Springer Berlin Heidelberg, 2006, pp. 499-511.

[6] Z. C. Johanyák and S. Kovács, "Fuzzy Rule Interpolation by the Least Squares Method," 7th International Symposium of Hungarian Researchers on Computational Intelligence (HUCI 2006), November 24-25, 2006 Budapest, pp. 495-506.

.[7] Z. C. Johanyák, D. Tikk, S. Kovács and K. W. Wong, "Fuzzy Rule Interpolation Matlab Toolbox – FRI Toolbox," Proc. of the IEEE World Congress on Computational Intelligence (WCCI'06), 15th Int. Conf. on Fuzzy Systems (FUZZ-IEEE'06), July 16--21, 2006, Vancouver, BC, Canada, pages 1427-1433, Omnipress.

[8] Z. C. Johanyák and S. Kovács, "Sparse Fuzzy System Generation by Rule Base Extension," 11th IEEE International Conference of Intelligent Engineering Systems (INES 2007), June 29 – July 1, 2007, Budapest, (in press)

[9] Z. C. Johanyák and S. Kovács, "Vague Environment-based Two-step Fuzzy Rule Interpolation Method," 5th Slovakian-Hungarian Joint Symposium on Applied Machine Intelligence and Informatics (SAMI 2007), January 25-26, 2007 Poprad, Slovakia, pp. 189-200.

[10] S. Kovács and L. T.Kóczy, "Application of an approximate fuzzy logic controller in an AGV steering system, path tracking and collision avoidance strategy," Fuzzy Set Theory and Applications , In Tatra Mountains Mathematical Publications, Mathematical Institute Slovak Academy of Sciences, vol.16, Bratislava, Slovakia, 1999, pp. 456-467.

[11] L. T. Kóczy and K. Hirota, "Approximate reasoning by linear rule interpolation and general approximation. International Journal of Approximative Reasoning," 9:197–225, 1993.

[12] L. T. Kóczy, K. Hirota and T. D. Gedeon, "Fuzzy rule interpolation by the conservation of relative fuzziness," Technical Report TR 97/2. Hirota Lab, Dept. of Comp. Int. and Sys. Sci., Tokyo Inst. of Techn., Yokohama, 1997.

[13] P. M. Larsen, "Industrial application of fuzzy logic control," Int. J. of Man Machine Studies, 12(4):3-10, 1980.

[14] E. H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," Int. J. of Man Machine Studies, 7:1-13, 1975.

[15] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," IEEE Trans. on SMC, 15:116-132, 1985.

[16] D. Shepard, "A two dimensional interpolation function for irregularly spaced data," Proc. 23rd ACM Internat. Conf., (1968) 517-524.

[17] M. Sugeno and T. Yasukawa, “A fuzzy-logic-based approach to qualitative modeling,” IEEE Transactions on Fuzzy Systems, vol. 1, pp. 7-31, 1993.

[18] D. Tikk and P. Baranyi, “Comprehensive analysis of a new fuzzy rule interpolation method,” In IEEE Trans. Fuzzy Syst., vol. 8, pp. 281-296, June 2000.

[19] E. W. Weisstein, “Root-Mean-Square,” From MathWorld--A Wolfram Web Resource. http://mathworld.wolfram.com/Root-Mean-Square.html

[20] K. W. Wong and T.D. Gedeon, ”Petrophysical Properties Prediction Using Self-generating Fuzzy Rules Inference System with Modified Alpha-cut Based Fuzzy Interpolation,” Proceedings of The Seventh International Conference of Neural Information Processing ICONIP 2000, November 2000, Korea, pp. 1088 - 1092.

[21] K. W. Wong, T. D. Gedeon and D. Tikk, “An improved multidimensional α-cut based fuzzy interpolation technique,” In Proc. Int. Conf. Artificial Intelligence in Science and Technology (AISAT’2000) , Hobart, Australia, 2000, pp. 29–32.

[22] S. Yan, M. Mizumoto and W. Z. Qiao, “An Improvement to Kóczy and Hirota's Interpolative Reasoning in Sparse Fuzzy Rule Bases,” In International Journal of Approximate Reasoning, 1996, Vol. 15, pp. 185-201.

[23] L. A. Zadeh, “Outline of a new approach to the analysis of complex systems and decision processes,” IEEE Trans. on SMC, 3:28-44, 1973.

[24] http://fri.gamf.hu