Fuzzy Rule Interpolation for Multidimensional Input Space with Petroleum Engineering Application

Kok Wai Wong and Tamás D. Gedeon
School of Information Technology
Murdoch University
South St, Murdoch
Western Australia 6150
Email: { kwong | tgedeon } @murdoch.edu.au

Abstract
Fuzzy rule based systems have been very popular in many engineering applications. In petroleum engineering, fuzzy rules are normally constructed using some fuzzy rule extraction techniques to establish the petrophysical properties prediction model. However, when generating fuzzy rules from the available information, it may result in a sparse fuzzy rule base. The use of more than one input variable is also common in petroleum engineering. This paper examines the application of fuzzy interpolation to resolve the problems using sparse fuzzy rule bases, and perform analysis of fuzzy interpolation in multidimensional input space.

1. Introduction

In petroleum reservoir modeling, boreholes are drilled at different locations around the region. Well logging instruments are lowered into the borehole to collect data at different depths known as *well log data*. Well logging instruments used in the measurement of well log data fall broadly into three categories: electrical, nuclear and acoustic [1]. Examples are Gamma Ray (GR), Resistivity (RT), Spontaneous Potential (SP), Neutron Density (NPHI) and Sonic interval transit time (DT). There are over fifty different types of logging tools available for different requirements. Beside the well log data, samples from various depths are also obtained and undergo extensive laboratory analysis. This laboratory analysis data is known as *core data*. In well log analysis, the objective is to establish an accurate interpretation model for the prediction of petrophysical characteristics such as porosity, permeability and volume of clay for uncored depths and boreholes around the region [2,3]. Such information is essential to the determination of the economic viability of a particular well or reservoir to be explored.

Recently, fuzzy set theory that is capable of handling vagueness and uncertainty in the core data are also used [4]. A fuzzy set allows for the degree of membership of an item in a set to be any real number between 0 and 1. This allows human observations, expressions and expertise to be modeled more closely. Once the fuzzy sets have been defined, it is possible to use them in constructing rules for fuzzy expert systems and in performing fuzzy inference. This approach seems to be suitable to permeability determination as it allows the incorporation of intelligent and human knowledge to deal with each individual case. However, the extraction of fuzzy rules from the data can be difficult for analysts with little experience. This could be a major drawback for use in permeability determination. If a fuzzy rule extraction technique is made available, then fuzzy systems can still be used for permeability determination [5,6].

Normally the information embedded in the available core data is not enough to cover the whole population. With the use of fuzzy extraction techniques, the fuzzy rules generated from these core data form a sparse fuzzy rule base. Classical fuzzy reasoning methods cannot be used to handle a sparse fuzzy rule base. This is due to the lack of an inference mechanism in the case when observations find no fuzzy rule to fire in uncored depths or wells around the region [7]. This is undesirable when using a fuzzy interpretation model. If more than half the input instances in the prediction well cannot find any rule to fire, this interpretation model is considered useless.

This paper examines the practical use of fuzzy rule interpolation for multidimensional input spaces. The fuzzy interpolation techniques that are examined in this paper are the original KH fuzzy interpolation technique [8], the modified α-cut fuzzy interpolation (MACI) technique [9] and the improved

0-7803-7078-3/01/$10.00 (C)2001 IEEE.

multidimensional fuzzy interpolation technique (IMUL) technique [10].

2. Fuzzy Rule Interpolation for Multidimensional Input Space

Fuzzy rule interpolation techniques provide a tool for specifying an output fuzzy set whenever at least one of the input spaces is sparse. Kóczy and Hirota [8] introduced the first interpolation approach known as (linear) KH interpolation. This is based on the Fundamental Equation of Rule Interpolation (refer to equation (1)). This method determines the conclusion by its α-cuts in such a way that the ratio of distances between the conclusion and the consequents should be identical with that among observation and the antecedents for all-important α-cuts (breakpoint levels). This is shown in the equation as follows (refer to Figure 1 for notations):

$$d(A^*,A_1):d(A^*,A_2)=d(B^*,B_1):d(B^*,B_2) \qquad (1)$$

Two conditions apply for the usage of the linear interpolation. First, there should exist an ordering on the input and output universes. This allows us to introduce a notion of distance between the fuzzy sets. Second, the input sets (antecedents, consequents, and the observation) should be convex and normal fuzzy sets.

The KH interpolation possesses several advantageous properties. Firstly, it behaves approximately linearly between the breakpoint levels. Secondly, its computational complexity is low, as it is sufficient to calculate the conclusion for the breakpoint level set. However, for some input situations it fails to result in a directly interpretable fuzzy set, because the slopes of the conclusion can collapse as shown in Figure 1. A modification of the original method has been proposed in the MACI [9] and IMUL [10] techniques, which can solve the problem of abnormal conclusions while maintaining its advantageous properties.

In the following sub sections, we will examine the three fuzzy interpolation techniques for use in multidimensional input spaces. For ease of computation and ease of interpretability, normally triangular or trapezoidal membership functions are used in well log analysis in petroleum engineering. The notations used in this paper are shown in Figure 2.

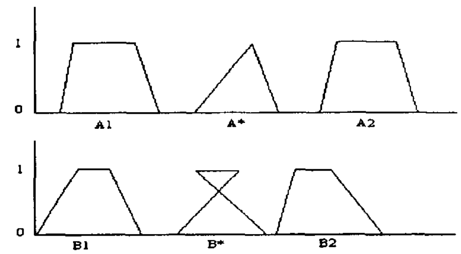

Figure 1: Fuzzy rule interpolation

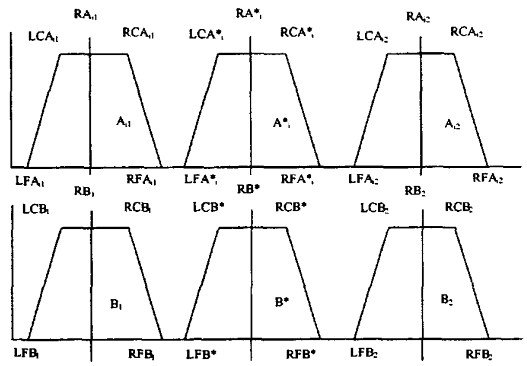

Figure 2: Notations used.

(A) KH Fuzzy Interpolation for Multidimensional Input Space

The KH fuzzy interpolation can be extended to multidimensional input space by using the Euclidean distance on all input spaces. For k input dimensions:

The right core for trapezoidal membership:

$$RCB^* = \frac{d_{1RC} * RCB_1 + d_{2RC} * RCB_2}{d_{1RC} + d_{2RC}} \qquad (2)$$

where

$$d_{1RC} = \sqrt{\sum_{i=1}^{k}(RCA_{i2} - RCA_i^*)^2}$$

and

$$d_{2RC} = \sqrt{\sum_{i=1}^{k}(RCA_i^* - RCA_{i1})^2}$$

The right flank:

$$RFB^* = \frac{d_{1RF} * RFB_1 + d_{2RF} * RFB_2}{d_{1RF} + d_{2RF}} \quad (3)$$

where

$$d_{1RF} = \sqrt{\sum_{i=1}^{k} (RFA_{i2} - RFA_i^*)^2}$$

and

$$d_{2RF} = \sqrt{\sum_{i=1}^{k} (RFA_i^* - RFA_{i1})^2}$$

The left flank and left core can be calculated similar to the above.

(B) MACI Technique for Multidimensional Input Space

MACI works with the vector description of fuzzy sets. The fuzzy set A is represented by a vector $a = [a_{-m}, ..., a_0, ..., a_n]$ where $a_k (k \in [-m, n])$ are the characteristic points of A and a_0 is the reference point of A with membership degree one. This means that $a_L = [a_{-m}, ..., a_0]$, and $a_R = [a_0, ..., a_n]$ are the left flank and right flank of A, respectively. Similarly the basic technique of MACI is extended to multidimensional input spaces using Euclidean distance on all input space. In our case, the reference points of all membership functions can be calculated by taking the mid point of the membership function. For k input dimensions:

The reference point of the interpolated conclusion for trapezoidal membership function is:

$$RB^* = (1 - \lambda_{core}) RB_1 + \lambda_{core} RB_2 \quad (3)$$

where

$$\lambda_{core} = \frac{\sqrt{\sum_{i=1}^{k} (RA_i^* - RA_{i1})^2}}{\sqrt{\sum_{i=1}^{k} (RA_{i2} - RA_{i1})^2}}$$

With the reference point the left and right cores can be calculated. For the right core:

$$RCB^* = (1 - \lambda_{right}) RCB_1 + \lambda_{right} RCB_2 + (\lambda_{core} - \lambda_{right})(RB_2 - RB_1) \quad (4)$$

where

$$\lambda_{right} = \frac{\sqrt{\sum_{i=1}^{k} (RCA_i^* - RCA_{i1})^2}}{\sqrt{\sum_{i=1}^{k} (RCA_{i2} - RCA_{i1})^2}}$$

For the right flank:

$$RFB^* = (1 - \lambda_{rightflk}) RFB_1 + \lambda_{right} RFB_2 + (\lambda_{core} - \lambda_{right})(RB_2 - RB_1) + (\lambda_{right} - \lambda_{rightflk})(RCB_2 - RCB_1) \quad (5)$$

where

$$\lambda_{rightflk} = \frac{\sqrt{\sum_{i=1}^{k} (RFA_i^* - RFA_{i1})^2}}{\sqrt{\sum_{i=1}^{k} (RFA_{i2} - RFA_{i1})^2}}$$

MACI will yield only a singleton conclusion if and only if the consequents are singletons.

(C) IMUL Technique for Multidimensional Input Space

This method incorporates features of the MACI and the conservation of fuzziness technique [11]. It makes use of the vector description of the fuzzy sets by representing them as characteristic points, and the coordinate transformation features of the MACI. At the same time, it can take the fuzziness of the fuzzy sets in the input spaces at the conclusion as those presented in the conservation of fuzziness technique. The advantage of this fuzzy interpolation technique is not only that it takes the fuzziness of the sets at the input spaces, but also makes use of the information of the core at the consequents.

For k input dimensions, the reference characteristic point of the interpolated conclusion with the use of Euclidean distance is:

$$RB^* = (1-\lambda_{core})RB_1 + \lambda_{core}RB_2 \quad (6)$$

$$\text{where } \lambda_{core} = \frac{\sqrt{\sum_{i=1}^{k}(RA_i^* - RA_{i1})^2}}{\sqrt{\sum_{i=1}^{k}(RA_{i2} - RA_{i1})^2}}$$

By using the above reference point, the right cores of the conclusion are calculated as follows:

For left core:

$$\begin{aligned} &LCB^* = \\ &(1-\lambda_{left})LCB_1 + \lambda_{left}LCB_2 + \\ &(\lambda_{core} - \lambda_{left})(RB_2 - RB_1) \end{aligned} \quad (7)$$

where

$$\lambda_{left} = \frac{\sqrt{\sum_{i=1}^{k}(LCA_i^* - LCA_{i1})^2}}{\sqrt{\sum_{i=1}^{k}(LCA_{i2} - LCA_{i1})^2}}$$

After calculating the cores of the two sides, the two flanks can then be calculated. When calculating the left and right flanks of the conclusion, the relative fuzziness of the fuzzy sets in all the input spaces are taken into consideration as follows:

Based on A_{i1} and B_1

$$s_i = RFA_{i1} - RCA_{i1} \quad (8)$$

$$s' = RFB_1 - RCB_1 \quad (9)$$

$$r_i = LCA_i^* - LFA_i^* \quad (10)$$

$$r' = LCB^* - LFB^* \quad (11)$$

$$u_i = RA_i^* - RA_{i1} \quad (12)$$

$$u' = RB^* - RB_1 \quad (13)$$

In multidimensional input spaces,

$$s = \sqrt{\sum_{i=1}^{k}(s_i)^2} \quad (14)$$

$$r = \sqrt{\sum_{i=1}^{k}(r_i)^2} \quad (15)$$

$$u = \sqrt{\sum_{i=1}^{k}(u_i)^2} \quad (16)$$

For left flank:

$$LFB^* = LCB^* - r_k\left(1 + \left|\frac{s'}{u'} - \frac{s}{u}\right|\right) \quad (17)$$

IMUL will result in a singleton fuzzy set if and only if all the observations are crisp and the core of the consequent is only one point.

3. Petroleum Engineering Case Study

In this case study, data from two wells in the same region are used. The input well logs used in this case study are gamma ray (GR), deep induction resistivity (ILD) and sonic travel time (DT). They are used to predict the petrophysical property, porosity (PHI). Core data from one well was used to establish a prediction model based on the fuzzy extraction technique used in [5]. The model was then used to predict the porosity in the second well. All the variables were normalised between the values of 0 and 100. The first well has a total of 71 core data and was used to establish the fuzzy rules. The second well had 51 core data and was used as the testing well to test the prediction accuracy. A few membership functions (3,5,7,9) have been tested, and 9 membership functions appeared to give the best prediction results. For ease of computation, only triangular membership functions were used. The total number of rules extracted from the training well was 63.

After all the fuzzy rules have been set up, the inputs from the second well ware used to infer the predicted PHI. In this well, two of the input vectors found no fuzzy rules to fire, i.e. fell into the gap. However, this

0-7803-7078-3/01/$10.00 (C)2001 IEEE.

number of input vectors that find no fuzzy rules to fire is considered very small. If more than half the input vectors in well 2 find no fuzzy rules to fire, then this fuzzy model built is considered useless. After the two input instances which did not have any fuzzy rule to fire have been picked up, the nearest fuzzy rules in the established fuzzy rule base need to be selected. From the observation and Euclidean distance measured on each input variable, the nearest fuzzy rules for the two input instances are determined.

The parameters used to interpolate the first input instance are as follows (refer to Figure 3):

A11: 25, 38, 38, 50	B1: 38, 50, 50, 63
A12: 0, 13, 13, 25	B2: 25, 38, 38, 50
A1*: 27, 27, 27, 27	B* (KH): 34, 43, 43, 51
A21: 50, 63, 63, 75	B* (MACI): 30, 42, 42, 54
A22: 13, 25, 25, 38	B* (IMUL): 42, 42, 42, 42
A2*: 37, 37, 37, 37	
A31: 63, 75, 75, 88	
A32: 25, 38, 38, 50	
A3*: 51, 51, 51, 51	

The parameters used to interpolate the second input instance are as follows (refer to Figure 4):

A11: 25, 38, 38, 50	B1: 25, 38, 38, 50
A12: 25, 38, 38, 50	B2: 38, 50, 50, 63
A1*: 36, 36, 36, 36	B*(KH): 32, 45, 45, 57
A21: 38, 50, 50, 63	B*(MACI): 33, 45, 45, 58
A22: 13, 25, 25, 38	B*(IMUL): 45, 45, 45, 45
A2*: 32, 32, 32, 32	
A31: 38, 50, 50, 63	
A32: 63, 75, 75, 88	
A3*: 60, 60, 60, 60	

The prediction accuracy for this case study is calculated using correlation factor as follows, the results are shown in Table 1.

$$\ell_{x,y} = \frac{\operatorname{cov}(X,Y)}{\sigma_x . \sigma_y} \qquad (18)$$

where $-1 \leq \ell_{x,y} \leq 1$, and

$$\operatorname{cov}(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_x)(y_i - \mu_y)$$

Table 1: Correlation Factors

Test	Correlation factor
Without fuzzy interpolation	0.654
With KH	0.859
With MACI	0.860
With IMUL	0.860

From the results, all three methods generate compatible results, however IMUL has the advantage of obtaining the porosity prediction directly from the interpolated results. As for KH and MACI, the interpolated fuzzy sets are added to the original fuzzy system in order to obtain a crisp porosity prediction.

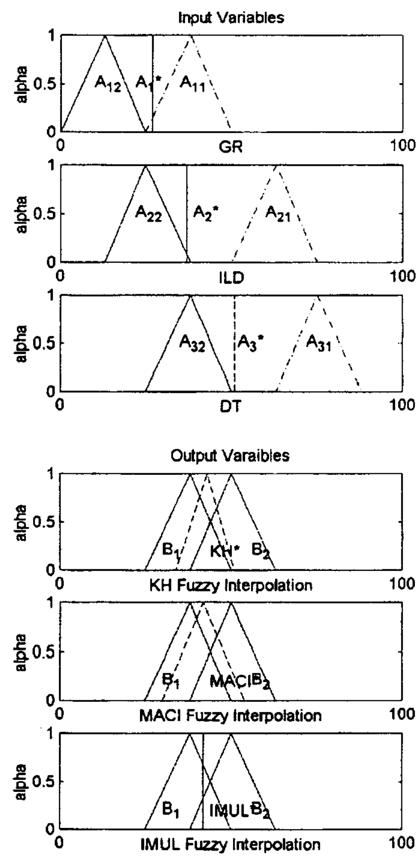

Figure 3: The first input instance

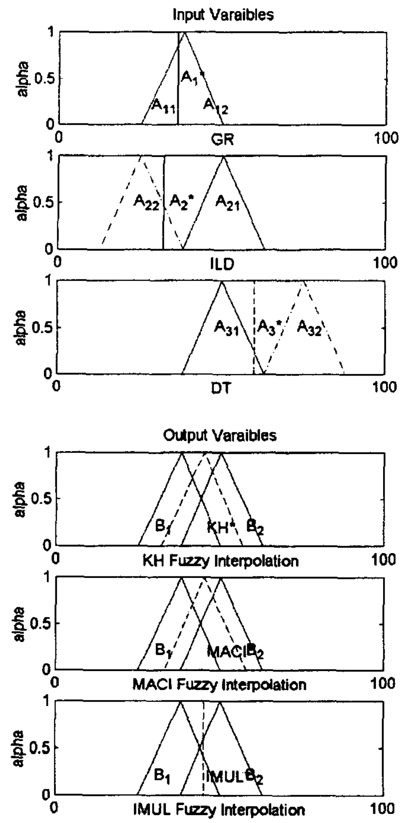

Figure 4: The second input instance

4. Conclusions

In this paper, the problem of sparse rule base and insufficient core data may cause undesirable prediction outcomes in petroleum engineering. This is mainly due to input instances that could not find any rule in the fuzzy rule base. To provide a solution to this problem, fuzzy interpolation techniques have been applied. The method can be used to interpolate the gaps between the rules. This ensures that the set of sparse fuzzy rules generated by the fuzzy rule extraction technique will be usable in a practical system. This is significant, as this will allow the use of fuzzy systems as an alternative for petrophysical properties prediction in petroleum engineering, at the same time without increasing the number of fuzzy rules that allows more human control.

5. References

[1] M. Rider, *The Geological Interpretation of Well Logs*, Second Edition, Whittles Publishing, 1996.

[2] E.R. Crain, *The Log Analysis Handbook Volume 1: Quantitative Log Analysis Methods*, Penn Well Publishing Company, 1986.

[3] G.B. Asquith and C.R. Gibson, *Basic Well Log Analysis for Geologists*, The American Association of Petroleum Geologists, 1982.

[4] L.A., Zadeh "Fuzzy Sets," *Information and Control*, vol. 8, 1965, pp. 338-353

[5] C.C. Fung, K.W. Wong, and P.M. Wong, " A Self-generating Fuzzy Rules Inference Systems for Petrophysical Properties Prediction," *Proceedings of IEEE International Conference on Intelligent Processing Systems*, 1997, pp. 205 – 208.

[6] H. Kuo, T.D. Gedeon, and P.M. Wong, "A Clustering Assisted Method for Fuzzy Rule Extraction and Pattern Classification," *Proceedings of 6th International Conference on Neural Information Processing*, vol. 2, 1999, pp. 679-684.

[7] 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.

[8] L.T. Kóczy, and K. Hirota, "Approximate reasoning by linear rule interpolation and general approximation," *Int. J. Approx. Reason*, Vol. 9, 1993, pp.197-225.

[9] D. Tikk, and P. Baranyi "Comprehensive Analysis of a New Fuzzy Rule Interpolation Method," *IEEE Transactions on Fuzzy Systems*, vol. 8, no. 3, 2000, pp. 281-296.

[10] K.W. Wong, T.D. Gedeon, and D. Tikk,, "An Improved Multidimensional Alpha-cut Based Fuzzy Interpolation Technique," *Proceedings of International Conference on Artificial Intelligence in Science and Technology 2000*, Hobart, Australia, December 2000, pp. 33-38.

[11] T.D. Gedeon, and L.T. Kóczy, "Conservation of fuzziness in rule interpolation," *Intelligent Technologies, International Symposium on New Trends in Control of Large Scale Systems*, vol. 1, Herlany, 1996, pp. 13-19.