Journal of Petroleum Science and Engineering 112 (2013) 310–321

Contents lists available at ScienceDirect

Journal of Petroleum Science and Engineering

journal homepage: www.elsevier.com/locate/petrol

Fast fuzzy modeling method to estimate missing logs in hydrocarbon reservoirs

Fouad Bahrpeyma [a,b], Bahman Golchin [c,d,*], Constantin Cranganu [e]

[a] Department of Computer Engineering, Faculty of Electrical and Computer Engineering, Shahid Beheshti University, Tehran, Iran
[b] Department of Information Technology, Information Technology Research Center (ITRC), Tehran, Iran
[c] Department of Geology, Faculty of Earth Sciences, Shahid Beheshti University, Tehran, Iran
[d] Drilling and Measurement (D&M), SeaLand Engineering and Well Services, Tehran, Iran
[e] Department of Earth and Environmental Sciences, Brooklyn College of the City University of New York, Brooklyn, NY, USA

ARTICLE INFO

Article history:
Received 12 April 2012
Accepted 3 November 2013
Available online 26 November 2013

Keywords:
fast fuzzy modeling method (FFMM)
Ink Drop Spread (IDS) operator
center of gravity (COG) operator
back-propagation artificial neural network (BP-ANN)
Takagi–Sugeno fuzzy inference system (TS-FIS)

ABSTRACT

In order to deal with huge amounts of computational complexities, conventional modeling systems always had to choose a tradeoff between accuracy and rapidity and usually one prevails over the other. Thus there is a need for a solution which provides acceptable accuracy and rapidity at the same time. In this research we propose a new fast fuzzy modeling method (FFMM) using Ink Drop Spread (IDS) and Center of Gravity (COG) operators. We applied this method to estimate missing logs of sonic and density. In the petroleum industry, characterization of pore–fluid pressures and rock lithology, along with estimation of porosity, permeability, fluid saturation and other physical properties are crucially important for successful exploration and exploitation. For many reasons, such as incomplete logging, inappropriate data storage and measurement errors, log suites are either incomplete or unreliable. By applying the proposed method, we estimated sonic and density logs. Correlation coefficients and MSEs for DT and RHOB logs were equal to 0.92, 21.07 and 0.85, 0.006 respectively. These results show that, despite the algorithm's very fast computation speed, its performance is comparable with that of methods like artificial neural network (ANN) and conventional Fuzzy Logic (FL); but the latter requires large amounts of storage and computing time, while the resources needed for the proposed method are limited to fixed values with respect to the number of data points.

1. Introduction

System identification involves using mathematical tools and algorithms to build dynamical models describing the behavior of real-world systems from measured data (Söderström and Stoica, 2001). There are always two conflicting requirements in the modeling process: the model capability to faithfully represent the real system (accuracy) and its ability to express the behavior of the real system in an understandable way (interpretability). Obtaining high degrees of accuracy and rapidity is usually contradictory and, in practice, one of the two properties prevails over the other (Cordón, 2011). Fuzzy systems have demonstrated their superb ability as system modeling tools (Bardossy and Duckstein, 1995; Driankov et al., 1996; Kuncheva, 2000). Conventional fuzzy modeling approaches like Takagi–Sugeno and Sugeno–Yasukawa try to use different types of methods to give them the abilities to deal with the accuracy and the computational complexity of the system at the same time, and almost a tradeoff between accuracy and time performance should take place. Inspired by the ability of the human brain to identify, model and estimate quickly and almost accurately, it seems that using the same method like what the brain uses can produce the same results. Bagheri Shouraki and Honda (1997a) introduced a new method, "Establishment and Saving Fuzzy Membership Functions", that enables the use of the same method that the brain is using when solving problems. Then Bagheri Shouraki and Honda (1999) introduced a new recursive approach called "Active Learning Method (ALM)" using Ink Drop Spread (IDS) operator to recursively find the best results. The proposed algorithm uses the same method of devoting membership degrees to the elements of the system domain with respect to the variations between the system elements and general behavior of each input–output plane called the Narrow Path (NP). That represents a simple use of ALM. Taheri Shahriyani et al. (2006) used ALM in marine ecology. Considering the robustness of fuzzy modeling where the level of uncertainty is high, the approach presented in this paper exhibits rapidity, robustness, and examined

* Correspondence to: Faculty of Earth Sciences, Shahid Beheshti University, Daneshjoo Avenue, Velelnjak, Tehran 1983963113, Iran. Tel.: +98 93 6008 7968; fax: +98 21 2243 1690.
E-mail addresses: Bahmangolchin@hotmail.com, B.golchin@sealandws.com (B. Golchin).

0920-4105/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.petrol.2013.11.019

accuracy at the same time. We applied this fast fuzzy modeling method (FFMM) for estimation of problems in hydrocarbon reservoir parameter domain with an example of estimating the missing well logs.

Because petrophysical and geophysical logs are such important tools for reservoir evaluation, engineers and petroleum geoscientists use well logs to obtain reservoir parameters such as porosity, permeability, volume of shale, oil and water saturation, rock lithology, fluid contacts and productive zones. Because of borehole conditions, instrument failure, poor logging condition, loss of data due to inappropriate storage and incomplete logging, most wells do not have a complete set of data in wells (Rezaee et al., 2008; Cranganu and Bautu, 2010; Cranganu and Breaban, 2013). Therefore, we propose a FFMM to accurately synthesize missing log data, such as density (RHOB) and sonic (DT) logs. The model performance was compared and evaluated by results obtained using artificial neural networks (BP-ANN) and conventional fuzzy logic (TS-FIS) methods.

The dataset in this study came from a well of Southern Iran. The dataset consists of density (RHOB), photoelectric (PEF), sonic (DT) and neutron (NPHI) logs. In this research, we divided data in two parts: 4025 (89%) data points for model construction and 500 (11%) data points for testing the models.

2. Methodology

Considering the ability of the human brain to estimate and model systems so quickly, it seems that the mind uses a specific non-mathematical approach with no computational complexity. On the other hand, the research shows that the human brain tries to remember the general behaviors of the system instead of storing the exact values of system elements (Bagheri Shouraki and Honda (1997a)). Breaking a Multiple Input, Single Output (M.I.S.O.) system into some Single Input, Single Output (S.I.S.O.) subsystems provides for estimating the general behavior of system output with respect to each input. Furthermore, taking the system apart causes it to forget about the internal relationship between input variables and instead uses a proper interpolation mechanism to obtain the final results.

Then, by applying a similar approach in the field of identification and modeling, we may obtain the same results, sparing us from dealing with huge amounts of computational complexities. As a caveat one has to check the relevance of each input to the output first, so as not to add extra attenuations and spreads to the model. This goal could be achieved by using the following algorithm.

First, the M.I.S.O. system should be broken into some S.I.S.O. sub-systems to obtain the general behavior of each X_i–Y plane. Each S.I.S.O. sub-system is expressed as a data plane (called IDS plane) resulting from the projection of the gathered data on each input–output (I–O) plane. The main problem that would be taken care of is the recognition of the most relevant inputs (the inputs which have greater effect on determining the value of the chosen parameter) to the output so as not to lower the performance of the system.

Recognizing the most relevant inputs to the desired output is obtained by calculating the spread of each X_i–Y plane to see if the spread is more than a predetermined value; the input is not optimistically relevant and could be removed from the process. In fact the spread of each X_i–Y plane is a result of involving other inputs in the process of recognition, but it seems that each X_i–Y plane should follow a general behavior that shows the relevance of the input to the output.

It is obvious that a criterion should be chosen to measure the degree of the spread. Bagheri Shouraki and Honda (1999a) used the sum of the variations of data points around the main behavior of 2-D data planes as the spread which is obtained by applying Ink Drop Spread (IDS) and center of gravity (COG) operators on data planes. This behavior is the narrowest continuous path on the data plane which is commonly called the Narrow Path (NP). IDS operator is used as a thickening operator to find a continuous path through the data plane, and COG is then used as a thinning operator trying to find the narrowest path which shows the general behavior of the selected plane. The level of confidence for each input variable is proportional to the reciprocal of variance of the data around NP.

If x_i^1 is in X_1–Y IDS plane: y_i is NP_1,

If x_i^2 is in X_2–Y IDS plane: y_i is NP_2,

If x_i^3 is in X_3–Y IDS plane: y_i is NP_3, where i shows ith data point.

It seems that there is a problem with the number of answers. There are multiple answers for each input point. There should be a

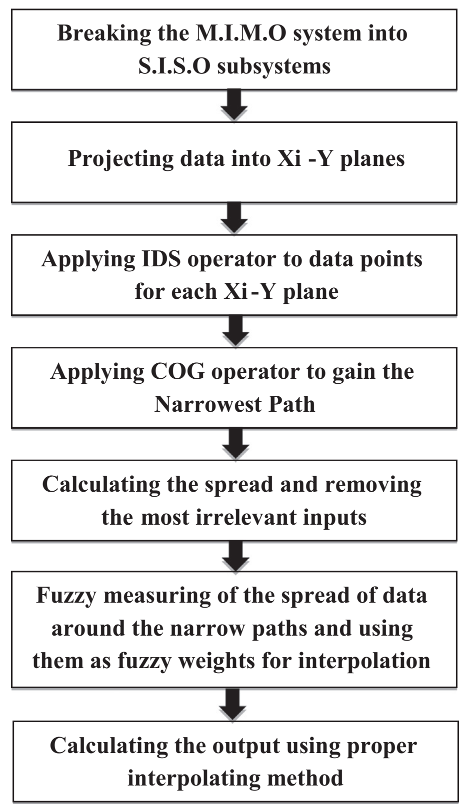

Fig. 1. The fast fuzzy modeling method (FFMM) flow chart.

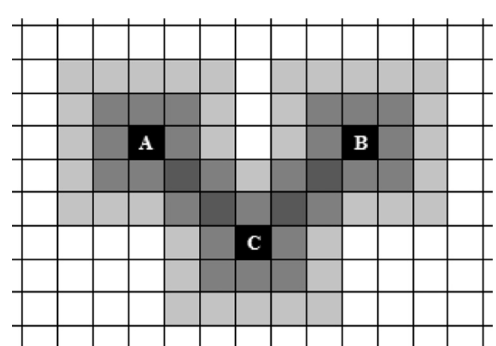

Fig. 2. Extracting the midpoints relationships between observed data samples in a S.I.S.O. system.