FUZZY LOGIC AS A TOOL FOR ESTIMATING PRODUCTION POTENTIAL OF A SAND LAYER

by

Bambang Widarsono, Heru Atmoko, Fakhriyadi Saptono, Tunggal and Ridwan

I. INTRODUCTION

In production management, a prior knowledge over production potential of a candidate sand layer (geological complexity in Indonesia has led to existence of stratified reservoirs with a set of layers) to be opened is always desirable. The common practice performed during drilling and completion activities of a production well is through the use of well testing and fluid sampling. From the test, fluid dynamic data such as total liquid rate, water cut, and gas cut are produced. A similar set of data is also required for more mature fields for the purpose of monitoring through the running of routine production and/or swab tests.

Although the tests, especially flow tests during drilling and completion, are always regarded as the only source of proof about productive layer(s) production potential, an alternative means that can be used to provide estimates is always desired. The main reason is that flow tests are costly so that only layer(s) considered as the most potential are to be assigned for testing. Layer(s) that are considered less potential are left untested, eventhough in some cases they are also set on production during the well’s production phase.

The idea of establishing a method that can provide illustration over production potentials of all layer(s) always exists. Certainly, there are approaches to serve the purpose such as productivity index (PI) analogy and petrophysical through fractional flow measurement in a core laboratory. However, those approaches are often considered inadequate in accommodating various factors that may influence production potential.

To materialize the idea stated above, the pattern recognition approach was taken. This approach was taken in order to model the relationships between various factors in wellbore and production potential without being trapped by the certain complexity that occurs in any mathematical expressions trying to explain the relationships. For the purpose, fuzzy logic (a branch in Artificial Intelligence) has been used. The choice is actually based on its ability to accommodate both numeric and non-numeric data. Some non-numeric data such as lithology and pore system also have some degrees of influence on production potential.

With a tool that enables us to have production potential estimates of reservoir layers, from which layers with the most promising potential are taken to undergo flow tests. Furthermore, as flow test data has been acquired and used as feedback and calibration by the fuzzy model, production potential of layer(s) with less promising or ambiguous prospect can also be predicted.

II. METHODOLOGY AND CONCEPT OF FUZZY LOGIC

Unlike the conventional binary or boolean logic which is based on sets of ‘true’ and ‘false’, fuzzy logic is essentially a methodology that allows an object to belong to both ‘true’ and ‘false’ but with different ‘degrees of membership’ that range between 0 and 1. This underlines the uniqueness of fuzzy logic that can handle the concept of ‘partial truth’, which represents truth values that are neither ‘completely true’ nor ‘completely false’.

The most importance of fuzzy logic is that it simulates the way human thinks, i.e. combining quantitative, qualitative, and subjective information. It can handle overlapping, imprecise, approximate, and linguistic information such as ‘low’, ‘medium’ and ‘high’. This makes fuzzy logic suitable for the nature of data the geophysicists, geologists, and reservoir engineers have to evaluate (Wong et al, 2003).

The relationships between input and output variables can be described in forms of fuzzy IF-THEN rules. Given a set of input-output patterns, many similar fuzzy rules can be derived, in an overlapping manner, to cover the whole functional space. In the process, based on the available “historical” information (i.e. database) that is transformed into fuzzy variables, fuzzy inference rules
are constructed and can be easily validated or modified by engineers or geoscientists. When new geological and non-geological factors are obtained, specific decisions can be derived and uncertainties can be quantified in terms of the degree of reliability (i.e. degree of membership). The entire process is best simplified in the schematic diagram presented in Figure 1.

Practically, the ‘database’ part in the Figure 1 consists of key data that is obtained from some key layers chosen because of having individual flow test data. (when there is no such data available for a field, data from nearby fields can cautiously be used.) Using this key data, plus additional rock and fluid data for those layers, fuzzy rule bases are extracted from the data population of the key layers. One rule for each data pair. Hundreds or even thousands of data pairs mean hundreds or thousands fuzzy rules. The fuzzy rule bases are then reduced in size through eliminating similar fuzzy rules.

The fuzzy interpretation model (i.e. fuzzy rule bases) is then ready to be used to estimate production potential (i.e. hypothetic individual flow test data) for any candidate layers that do not have actual flow test data, either in the same or in other wells. The same type of data variables as the ones used in the modeling are used as input data.

In the process, re-evaluation can be performed when required. The evaluation work in this part may involve re-training of the fuzzy rule bases whenever required or simply addition of new fuzzy rules through external interference from the analyst based on his/her knowledge and experience. Hypothetical data or even data from nearby fields with similarity can also be integrated in the re-training. This also covers evaluation on prediction results for candidate wells in which external modification (i.e. intervention from outside the use of the software package) is a possibility when required.

III. ANALYSIS OF DATA REQUIRED

It is a dream of any production engineer to have accurate production potential data of all layers. Ideally, this data is acquisitioned through the use of direct measurement such as direct flow tests from time to time in a regular basis. This is, of course, often considered as too expensive and impractical, especially for fields with large number of production wells. This is not to mention cases in which the production condition does not suit to the condition required for running the equipment.

Considering the possibility of large number layers and wells, engineers tend to find a means of deriving the needed data through less direct means. Practices, as mentioned before, such as the use of productivity index analogy and hypothetical production flow schemes using fractional flow laboratory test data are often made.

In reality, the examples of practices mentioned above are often considered too simplistic for at least two reasons; Firstly, all two fluids (oil and water or gas and water) or three fluids (oil, water, and gas) contribute to the total well production in exactly the same proportion. This certainly defies some common occurrences in which there are possibilities that decline in production rate are caused by rise in water cut. Secondly, the properties mentioned above are not the only affecting factors on the individual layer’s production potential.

In this study attempts have been made to evaluate other factors that may play a significant influence in this matter. Apart from the three rock and fluid properties mentioned previously, rock physical properties such as porosity (\(\phi\)), water saturation (\(S_w\)), irreducible water saturation (\(S_{wir}\)), and shale contents/types. Other factors and data such as geological data (e.g. facies type, clay type, pore type/system, and lithology) may also have some causative pattern with individual layer’s production potential. Based on the evaluation, data collection was focused on the data mentioned above. This set of data was then prepared for fuzzy modeling, validation, and prediction.
IV. REQUIRED KEY DATA

To fulfill the fuzzy modeling, the following data is required from the assigned key layers (in case no data available for a field, data from nearby fields can cautiously be used):

**Fluid production data for each individual layer**
- Total liquid flow rates
- Oil production rates and water cut

**Total production data for key layers**
- Total liquid flow rates. The figures must be consistent when compared to the sum of individual layer’s flow rates.
- Total oil flow rate and water cut

**Supporting data relevant to the estimation of production potential**
- Petrophysics: permeability, porosity, water saturation, shale contents, irreducible water saturation.
- Sand geometry: sand thickness, perforation thickness, net sand thickness, sand-shale ratio.
- Geological information: facies, lithology, pore type, shale type
- Engineering data: reservoir pressure, bottom hole pressure (static and dynamic), liquid physical properties.

As a case study a Central Sumatran oil field, CS field, is used. The field was discovered in November 1964 through the exploratory well CS-01, which penetrated an oil column at a depth of 3100 - 4556 ft subsea. Commercial production was started in December 1968. Productive sands of the field are in the Sihapas Group (comprising the Bekasap, Bangko and Menggala formations) and Pematang Group. So far, 92 production wells have been drilled (as per December 2003), of which 79 are in the status of producing.

<table>
<thead>
<tr>
<th>Well</th>
<th>Formation</th>
<th>Reservoir</th>
<th>Data test</th>
<th>Perforated</th>
<th>H (Perf)</th>
<th>Production test</th>
<th>Swab test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Top</td>
<td>Bot</td>
<td>BFPD</td>
<td>BOPD</td>
</tr>
<tr>
<td>CS-43</td>
<td>Pematang</td>
<td>50xx SD</td>
<td>14-May-92</td>
<td>5008</td>
<td>5025</td>
<td>22</td>
<td>482</td>
</tr>
<tr>
<td>CS-44</td>
<td>Bangko</td>
<td>35xx SD</td>
<td>31-May-92</td>
<td>3738</td>
<td>3748</td>
<td>24</td>
<td>254</td>
</tr>
<tr>
<td>CS-54</td>
<td>Pematang</td>
<td>42xx SD</td>
<td>5-Oct-00</td>
<td>4264</td>
<td>4269</td>
<td>5</td>
<td>1004</td>
</tr>
<tr>
<td>CS-73</td>
<td>Menggala</td>
<td>39xx SD</td>
<td>17-Mar-02</td>
<td>4262</td>
<td>4270</td>
<td>8</td>
<td>1296</td>
</tr>
<tr>
<td>CS-74</td>
<td>Bangko</td>
<td>35xx SD</td>
<td>17-Apr-02</td>
<td>3682</td>
<td>3698</td>
<td>16</td>
<td>1144</td>
</tr>
<tr>
<td>CS-76</td>
<td>Pematang</td>
<td>42xx SD</td>
<td>28-Apr-02</td>
<td>4570</td>
<td>4582</td>
<td>12</td>
<td>1452</td>
</tr>
<tr>
<td>CS-76</td>
<td>Menggala</td>
<td>38xx SD</td>
<td>24-Aug-03</td>
<td>4072</td>
<td>4096</td>
<td>24</td>
<td>2352</td>
</tr>
<tr>
<td>CS-78</td>
<td>Pematang</td>
<td>43xx SD</td>
<td>22-May-02</td>
<td>4600</td>
<td>4606</td>
<td>6</td>
<td>1363</td>
</tr>
<tr>
<td>CS-80</td>
<td>Pematang</td>
<td>42xx SD</td>
<td>15-Jul-02</td>
<td>4614</td>
<td>4624</td>
<td>10</td>
<td>1164</td>
</tr>
<tr>
<td>CS-81</td>
<td>Pematang</td>
<td>44xx SD</td>
<td>15-Jul-02</td>
<td>4750</td>
<td>4758</td>
<td>8</td>
<td>696</td>
</tr>
<tr>
<td>CS-85</td>
<td>Bangko</td>
<td>36xx SD</td>
<td>22-Mar-03</td>
<td>3836</td>
<td>3842</td>
<td>6</td>
<td>579</td>
</tr>
<tr>
<td>CS-88</td>
<td>Menggala</td>
<td>39xx SD</td>
<td>27-Aug-03</td>
<td>4051</td>
<td>4058</td>
<td>7</td>
<td>1210</td>
</tr>
<tr>
<td>CS-90</td>
<td>Pematang</td>
<td>46xx SD</td>
<td>28-Oct-03</td>
<td>4638</td>
<td>4648</td>
<td>10</td>
<td>432</td>
</tr>
</tbody>
</table>

Table 1
Production (and swap) test data of the key layers
From the 92 wells that have production flow test data only 32 wells have tests on single layer (all production wells are of commingle completion scheme). Apart from the production tests data there is also swab test data that was taken as a source of comparison only. Combined with other data, only 13 data sets have a complete set of data. Table 1 presents the production test data for the 13 key layers that was to be taken as key data.

The petrophysical properties data used in this study is the results of log interpretation conducted by the field's operator. For the 13 data points presented in Table 1, the petrophysical data is presented in Table 2.

For fluid properties, Table 3 presents the supporting data of oil and water formation volume factors \(B_o\) and \(B_w\) respectively) and viscosities. Properties of water are taken as constant for all sands and wells.

For pressure data, the static bottom-hole pressure (SBHP) is taken as the reservoir pressure \(P_r\). A problem had risen when it appeared that there is no flowing bottom-hole pressure (FBHP or \(P_w\)) data for the entire field. It is known that the difference between the two pressures \(P_r - P_w\) is considered as a factor that also influence split factor. The solution to the problem is to estimate the needed \(P_w\) data based on available static fluid level. Not all wells have the data, however. Table 4 presents the data.
For geological data, analysis on data has concluded that there are two main depositional environments; tidal channel/bar and estuary. Data for the 13 data sets are listed in Table 5.

V. DEVELOPMENT OF FUZZY INTERPRETATION MODEL

In the fuzzy modeling, and following the standard steps taken in fuzzy logic, 3 general steps were taken:
1. Data fuzzification
2. Fuzzy inference
3. Defuzzification

Figure 2 presents the process.

Data fuzzification

The first step in preparation of fuzzy modeling was to determine fuzzy sets into which all data variable to be used is grouped. In the process, the crisp key data that is available to the study is fuzzified or transformed from numeric to linguistic form. This transforms the Boolean logic that characterizes crisp data to fuzzy logic. To underline the difference between the two logics a classification of porosity population serves as an example. For instance, porosity data is grouped into:

LOW porosity < 18%
MEDIUM 18% ≤ porosity ≤ 30%
HIGH porosity > 30%

Unlike the Boolean approach (crisp approach) which uses the above classification set in a discontinuous way the fuzzy approach takes the class limits as merely a rough approximation for a continuous data such as porosity. In other words, crisp approach categorizes porosity of 17% as strictly LOW and 18% as strictly MEDIUM whereas fuzzy approach takes both values as belonging to both LOW and MEDIUM but with different degree of membership (see previous section for definition)
in regard to each class. Fuzzy sets allow some human expertise and decisions to be modeled more closely.

For the key data discussed above, each data variable was grouped into a set of fuzzy sets with their considered most appropriate membership functions. Figures 3 and 4 present examples of plots of the membership functions for water cut (WC) and shale contents \( V_{sh} \), respectively. The full membership functions of the two variables are presented in the Appendix. In essence, the functions convert the information from numerical level to linguistic level, as shown in Figure 2. Other numerical variables (porosity, permeability, transmissibility, water saturation, irreducible water saturation, oil rate, and delta pressure) also have their most suitable membership functions, whereas non-numerical variables (lithology, pore type, and depositional environment) are already in linguistic level (for instance, depositional environment: ESTUARINE and TIDAL CHANNEL).

As all data variables had been fuzzified, the next step was to apply fuzzy inference from which the fuzzy determination model would be derived.

**Fuzzy inference**

Following the approach proposed by Mamdani (Mamdani and Assilian, 1975; Kusumadewi, 2002), after the fuzzification of the *crisp* data, fuzzy rules were generated. These fuzzy rules cover the whole universe of discourse by taking all the possibilities into account. In general, a fuzzy rule takes a form of

\[
IF \ (x_1 \ is \ A_1) \ \& \ (x_2 \ is \ A_2) \ \& \ (x_3 \ is \ A_3) \ \& \ \ldots \ \& \ (x_n \ is \ A_n) \ \& \ THEN \ y \ is \ B
\]
with propositions following IF are called antecedent and propositions following THEN are called consequence. In the study the most suitable fuzzy operator is AND, hence making the general rule into

IF \( (x_1 \text{ is } A_1) \) AND \( (x_2 \text{ is } A_2) \) AND \( (x_n \text{ is } A_n) \) AND ...... AND \( (x_n \text{ is } A_n) \) THEN \( y \text{ is } B \).

In this fuzzy modeling, variables \( x \) and \( y \) are variables such as porosity, permeability, and water cut whereas \( A \) and \( B \) are fuzzy statements such as LOW, MEDIUM, SHALLY, etc. An example could be:

\[
\begin{align*}
\text{IF} & \ (K \text{ is HIGH}) \ \text{AND} \ (\theta \text{ is HIGH}) \ \text{AND} \\
& (S_w \text{ is LOW}) \ \text{AND} ...... \ \text{AND} \ (KH \text{ is LARGE}) \ \text{THEN} \ q_o \text{ is HIGH} \\
or \\
\text{IF} & \ (K \text{ is LOW}) \ \text{AND} \ (V_{wdl} \text{ is SHALLY}) \ \text{AND} \\
& (S_{wp} \text{ is MEDIUM}) \ \text{AND} ...... \ \text{AND} \ (KH \text{ is MEDIUM}) \ \text{THEN} \ q_o \text{ is LOW}
\end{align*}
\]

This rules establishment (fuzzy training) is the center of the fuzzy modeling. A fuzzy model consists of a set of fuzzy rules. For a problem that has two or three relevant variables, the number of fuzzy rules is small. However, for a problem with a large number of variables the number of fuzzy rules is likely to be too large to be established manually. An automatic fuzzy extraction method was used for the purpose.

Further in the inference process, the fuzzy rules undergo operations such as application of implication functions and rules composition in order to obtain correlation among the rules. Using the established fuzzy rules base, solution will be provided by the model when any sets of data crisp are input to it.

**Defuzzification**

So far, as the fuzzy interpretation model is used, output is still in the form of fuzzy linguistic data (for instance, \( q_o \text{ HIGH} \) and \( WC \text{ LOW} \)). This needs to be transformed back into crisp data (for instance, \( q_o = 615 \) BOPD and \( WC = 25\% \)). This is done through a process of defuzzification.

There are several known methods proposed to perform the process. The most widely used one is centroid method (Kusumadewi, 2002)

\[
z = \frac{\int z \mu(z) \, dz}{\int \mu(z) \, dz} \quad \text{or} \quad z = \frac{\sum_{j=1}^{n} z_j \mu(z_j)}{\sum_{j=1}^{n} \mu(z_j)}
\]
VI. DISCUSSION

Application of the method in the field will certainly face different problems from one field to another. Problem faced in CS field is that most of the flow tests have been conducted in commingle manner, which causes the number of data sets is limited. In other fields, other problem may occur, such as possession of old log survey data only, absence of reliable fluid and pressure data, and inaccurate completion record.

Problems such as ones mentioned above may reduce the reliability of the fuzzy interpretation model. In this case, experience of the analyst provide a big help to maintain the model’s reliability through external inference on the fuzzy rule bases, a feature that make fuzzy logic an interesting pattern recognition tool. Nevertheless, reliability of any fuzzy models will increase as more actual observed flow test data has become available, so that re-modeling and model modification can be carried out to yield a more robust model.

VI. CONCLUSIONS

From the study, some conclusions have been drawn:

1. Fuzzy logic can be used satisfactorily to estimate production potential of untested reservoir layers.
2. Facility offered by fuzzy logic to enable analysts to incorporate their experience is a great advantage.
3. Data variables used in the fuzzy modeling may be reduced in order to make a simpler model and larger number of fuzzy data training sets. However, this move has a risk of reducing the model’s reliability. Balance between data availability and model reliability has to be exercised.

REFERENCES

APPENDIX

The following examples of membership functions are for water cut (triangular model) and shale contents (sigmoid). Other data variables used in the fuzzy modeling may different membership functions, such as trapezoidal and bell-shape, dependent on the variables' data structure and pattern in general. The limiting values in the functions bellow are adopted after some trials. Note the overlapping nature of the functions.

Water cut (WC, fraction)

Membership function: Triangular and Linear

Data mapping resulted in the division of LOW, MEDIUM, and HIGH.

\[ \mu_{\text{LOW}}(WC) = \begin{cases} 1; & \text{WC} \leq 0.1 \\ (0.5 - WC)/0.4; & 0.1 < \text{WC} \leq 0.5 \\ 0; & \text{WC} > 0.5 \end{cases} \]

\[ \mu_{\text{MEDIUM}}(WC) = \begin{cases} 0; & \text{WC} \leq 0.1 \text{ or } \text{WC} \geq 0.9 \\ (WC - 0.1)/0.4; 0.1 \leq \text{WC} \leq 0.5 \\ (0.9 - WC)/0.4; & 0.5 < \text{WC} < 0.9 \end{cases} \]

\[ \mu_{\text{HIGH}}(WC) = \begin{cases} 1; & \text{WC} \leq 0.5 \\ (WC - 0.5)/0.4; & 0.5 < \text{WC} < 0.9 \\ 1; & \text{WC} \geq 0.9 \end{cases} \]

Shale contents (V_{sh}, fraction)

Membership functions: Sigmoid (S-curve), descending and ascending

Data mapping resulted in the division of CLEAN and SHALLY.

For descending S-curve (CLEAN), inflection point = 0.11

For ascending S-curve (SHALLY), inflection point = 0.18

\[ \mu_{\text{CLEAN}}(V_{sh}) = \begin{cases} 1; & V_{sh} \leq 0.05 \\ 1 - 2[(V_{sh} - 0.05)/0.15]^2; & 0.05 \leq V_{sh} \leq 0.1225 \\ 2[(0.11 - V_{sh})/0.15]^2; & 0.1225 \leq V_{sh} \leq 0.2 \\ 0; & V_{sh} \geq 0.2 \end{cases} \]

\[ \mu_{\text{SHALLY}}(V_{sh}) = \begin{cases} 0; & V_{sh} \leq 0.05 \\ 2[(V_{sh} - 0.18)/0.12]^2; & 0.05 \leq V_{sh} \leq 0.24 \\ 1 - 2[(0.3 - V_{sh})/0.12]^2; & 0.24 \leq V_{sh} \leq 0.3 \\ 1; & V_{sh} \geq 0.3 \end{cases} \]