



**HAL**  
open science

## Application of Bayesian in determining productive zones by well log data in oil wells

Pedram Masoudi, Behzad Tokhmechi, Majid Ansari, Mohammad Jafari,  
Shahram Zamanzadeh

► **To cite this version:**

Pedram Masoudi, Behzad Tokhmechi, Majid Ansari, Mohammad Jafari, Shahram Zamanzadeh. Application of Bayesian in determining productive zones by well log data in oil wells. *Journal of Petroleum Science and Engineering*, 2012, 94-95, pp.47-54. 10.1016/j.petrol.2012.06.028 . insu-01382398

**HAL Id: insu-01382398**

**<https://insu.hal.science/insu-01382398>**

Submitted on 17 Oct 2016

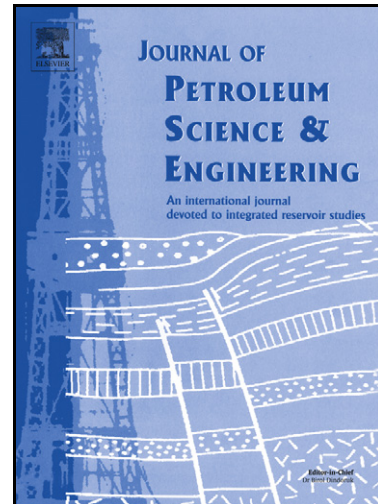
**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Author's Accepted Manuscript

Application of Bayesian in determining productive zones by well log data in oil wells

Pedram Masoudi, Behzad Tokhmechi, Majid Ansari Jafari, Mohammad Zamanzadeh, Shahram Sherhati



[www.elsevier.com/locate/petrol](http://www.elsevier.com/locate/petrol)

PII: S0920-4105(12)00170-2  
DOI: <http://dx.doi.org/10.1016/j.petrol.2012.06.028>  
Reference: PETROL2280

To appear in: *Journal of Petroleum Science and Engineering*

Received date: 30 June 2011  
Accepted date: 21 June 2012

Cite this article as: Pedram Masoudi, Behzad Tokhmechi, Majid Ansari Jafari, Mohammad Zamanzadeh and Shahram Sherhati, Application of Bayesian in determining productive zones by well log data in oil wells, *Journal of Petroleum Science and Engineering*, <http://dx.doi.org/10.1016/j.petrol.2012.06.028>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## Application of Bayesian in Determining Productive Zones by Well Log Data in Oil Wells

Pedram Masoudi<sup>1</sup>, Behzad Tokhmechi<sup>1</sup>, Majid Ansari Jafari<sup>1</sup>, Mohammad Zamanzadeh<sup>2</sup>, Shahram Sherkati<sup>3</sup>

<sup>1</sup>*School of Mining, Petroleum and Geophysics Engineering, Shahrood University of Technology, Shahrood, Iran*

<sup>2</sup>*Faculty of Geography, University of Tehran, Tehran, Iran*

<sup>3</sup>*Exploration Directorate of National Iranian Oil Company (NIOC), Tehran, Iran*

Correspondence: PedramMasoudi, masoudip@alumni.ut.ac.ir

### Abstract:

Exploration specialists conventionally utilize a cut-off-based method to find productive zones inside the oil wells. Using conventional method, pay zones are separated crisply from non-pay zones by applying cut-off values on some petrophysical features.

In this paper, a Bayesian technique is developed to find productive zones (net pays), and Bayesian Network is used to select the most appropriate input features for this newly developed method. So, two Bayesian methods were developed: the first one with conventional pay determination inputs (shale percent, porosity and water saturation), the other with two inputs, selected by Bayesian Network (porosity and water saturation). Two developed Bayesian methods are applied on well log dataset of two wells: one well is dedicated for training and testing Bayesian methods, the other for checking generalization ability of the proposed methods. Outputs of two presented methods were compared with the results of conventional cut-off-based method and production test results (i.e. a direct procedure to check validation of proposed methods).

Results show that the most prominent advantage of developed Bayesian method is determination of net pays fuzzily with no need to identify cut-offs, in addition to higher precision of classification: nearly 30% improvement in precision of determining net pays of first well (training well), and about 50% improvement in precision of determining productive zones through the generalizing well.

**Key words:** net pay; productive zone; Bayesian Network; petrophysics; well test

## 1. Introduction

After drilling an oil well, exploration specialists should determine productive zones through the well by means of well tests and logs. To commence production, explored productive zones should be perforated to let the hydrocarbon flow from the reservoir to the well bore. Conventionally, by applying cut-off values on well logs, productive zones are determined crisply.

Probably, the oldest and simplest methods, used to distinguish net pay zones from non-pays, were direct techniques determined while drilling. Two direct techniques, namely gas-metric and flour-metric, have been referred in the literature (Connell et al., 1986; Millikan, 1925). In these two techniques, fluid samples withdrawn from the wells are tested by specific processes, and net pay horizons are identified from the analysis of the tests results. The other group of methods to determine productive zones is indirect techniques that utilize wire line well log data for pay determination.

Pioneer petrophysicists used combination of gamma and resistivity logs (Snyder, 1971) and sonic-shear-wave and resistivity logs (Flower, 1983) to determine net pay (productive) zones. Later, formation pressure tester was introduced as a quick-look indicator of net pays (Cooke-Yarborough, 1984). Another investigation was fulfilled on detecting low contrast pays in a gas reservoir on an integrated dataset, which consists of image-log interpretations, conventional logs and core capillary pressures (Deakin and Manan, 1998). Based on dominated geological features, Worthington classified low-resistivity pays to six classes. In addition, he proposed a single-universal algorithm for detecting all classes of low-resistivity pay zones (Worthington, 2000). In a pioneer work, geochemical analysis of side-wall cores were incorporated to detect net pays (Mathur et al., 2001).

Cut-off of petrophysical parameters, e.g. porosity, shale volume and water saturation, is the most popular mean to determine net pays up to now. Worthington and Cosentino provided a comprehensive study on the role of cut-offs in determining net pays. They collected and summarized different combinations of cut-offs of shale volume, porosity, permeability, water saturation, resistivity and moveable hydrocarbon index (MHI), which have been used in 31 previous investigations from 1980 to 2002 (Worthington and Cosentino, 2005). A statistics-based procedure was introduced to determine cut-offs in a way to minimize error of calculating net to gross ratio (Jensen and Menke, 2006). Worthington also fulfilled another valuable research that determines cut-offs dynamically in accordance with depletion strategy (Worthington, 2008). A pioneer work of pay zone detection on seismic sections was later published (Singleton, 2008).

In a recently published paper, definition of different nets, especially net pay, is provided and application of net pay in petroleum industry is discussed (Worthington, 2010). Recently in another article a data-driven approach to identify cut-offs is proposed (Worthington, 2011). Masoudi et al. have proposed two new free of cut-off procedures to find productive zones fuzzily. One of them is based on flow equation (Masoudi et al., 2011), the other is based on Sugeno integral, which is a fuzzy operator (Masoudi et al., 2012).

The novelty of the current work is to develop a fusion-based methodology to identify net pays fuzzily without determining cut-offs and give priority to each depth. The results of this newly developed method are compared with conventional net pay determination and well test results.

### *1.1. Definition of Net Pay and Cut-off*

Although there is no universal definition of net pay, Worthington (2010) introduced an acceptable definition on which this investigation is based on. In his classification, total evaluation interval of an oil well is called gross rock, potential reservoir is named net sand, and net reservoir is defined as a subset of net sand that has supracritical amount of porosity and permeability. Finally, net pay includes some intervals of net reservoir, which contain supracritical amounts of recoverable hydrocarbons, and can produce oil or gas (Worthington, 2010).

Above mentioned nets, net sand, net reservoir and net pay, are distinguished from each other easily by applying cut-off values on petrophysical features. In other words, cut-offs are some criteria on petrophysical parameters to determine nets from each other. It means that intervals, which contain shale percentages of less than cut-off of shale, are considered as net sand. The range of cut-off of shale is between 30% to 50%. Those intervals of net sand that contain porosity of higher than cut-off of porosity (between six to eight percent for sandstones and between four to five percent for carbonates) are classified as net reservoir that are capable of containing economical amounts of hydrocarbon in the rock pores. Finally, net pay is some part of net reservoir in which water saturation is less than cut-off of water saturation (the range of cut-off of water saturation is between five to six percent) (Worthington and Cosentino, 2005).

## **2.Dataset**

This investigation is applied on well log data of interval of Sarvak Formation through two wells, located on an oil field in Abadan Plain, SW Iran. Raw data of this work are conventional well logs, porosity and permeability values (measured from cores) and well tests of the wells. Table 1 presents a summary of utilized dataset.

Table 1

Dataset, which was available from two wells of the oil field. “*npv*” stands for “net pay value”. *npv*=1 means that the interval of testing is not oil producing, *npv*=2 means that the tested interval produces oil less than

$$1500 \frac{\text{bbl oil}}{\text{day}} \left[ \cong 0.001104 \frac{\text{m}^3}{\text{sec}} \right], \text{ and } npv=3 \text{ means that the}$$

corresponding interval produces oil more than

$$1500 \frac{\text{bbl oil}}{\text{day}} \left[ \cong 0.001104 \frac{\text{m}^3}{\text{sec}} \right].$$

		Well 1	Well 2
No. of Well Test Intervals	<i>npv</i> =1	2	1
	<i>npv</i> =2	1	1
	<i>npv</i> =3	2	1
Petrophysical Well Logs	Caliper (CALI)	✗	✓
	Gamma Ray (GR)	✓	✓
	Corrected Gamma Ray (CGR)	✓	✓
	Sonic Log (DT)	✓	✓
	Neutron Porosity (NPHI)	✓	✓
	Bulk Density (RHOB)	✓	✓
	Density Correction (DRHO)	✓	✓
	Deep Laterolog Resistivity (LLD)	✓	✓
	Shallow Laterolog Resistivity (LLS)	✓	✓
	Microspherically Focused Log (MSFL)	✓	✓
	Photoelectric Effect Log (PEF)	✗	✓
Core Tests	Porosity	✓	✓
	Permeability	✓	✓

Sarvak Formation is a shallow carbonate reservoir rock with the age of Upper Albian to Upper Turonian, which is conformably laid on Kazhdumi Formation. The unconformity is interpreted as an erosional unconformity occurred after the formation deposition (Ghabeishavi et al., 2010). In the studied field, Laffan Formation is overlaid on Sarvak Formation above this sharp unconformity.

### 3.Methodologies: Cut-off and Bayesian

#### 3.1. Conventional Method: Based on Cut-off

As mentioned earlier, conventionally net pays are determined by applying cut-off values on some petrophysical or well log data (Deakin and Manan, 1998; Svec and Grigg, 2000; Worthington, 2010; Worthington and Cosentino, 2005). In this work, to determine net pay zones by conventional method, the following process is utilized:

1. Calculating shale volume ( $V_{sh}$ ) by normalizing CGR log. For this purpose, simply CGR values of each well are rescaled within zero and one by equation (1) (Darling, 2005); and water saturation is calculated from Archie relation, which is shown on equation (2) (Johnson and Pile, 2002).

$$V_{sh} = \frac{CGR - CGR_{min}}{CGR_{max} - CGR_{min}} \quad (1)$$

$$S_w = \sqrt{\frac{FR_w}{R_t}} = \sqrt{\frac{R_o}{R_t}} \quad (2)$$

2. Effective porosity is estimated by an Artificial Neural Network (MLP-structured) on input logs: CGR, DT, LLD, LLS, MSFL, NPHI and RHOB. The concept of Artificial Neural Network (ANN) and its application in petroleum industry is comprehensively explained in earlier works (Jalali Lichaei and Nabi Bidhendi, 2006; Mohaghegh et al., 2000).

3. Determining net pays by applying cut-offs of porosity, shale volume and water saturation on estimated petrophysical parameters, i.e. those intervals, with porosity values of higher than cut-off of porosity, in addition to shale percentage of lower than cut-off of shale volume and water saturation of lower than cut-off of water saturation, are considered as net pay zones, whereas other intervals are defined as not producing zones or non pays.

#### 3.2. Fusion-Based Method: Bayesian Classifier

Bayesian technique has been used as a method to fuse data (Challa and Koks, 2004; Hall and Llinas, 2001). Fusing data reduces error of identification (Hall and Llinas, 2001), and consequently lessens uncertainty and risk. Bayes theorem offers a method to calculate conditional probability of a happening, i.e. probability of occurrence of a phenomenon by considering the effect of other evidences (Niedermayer, 2008). The mathematical formula of Bayes is:

$$P(A|B) = \frac{P(A) \times P(B|A)}{P(B)} \quad (3)$$

where  $P(A)$  and  $P(B)$  are probabilities of A and B respectively.  $P(B|A)$  is the probability of B, considering A, known as “priors knowledge”, and reversely  $P(A|B)$  is the probability of A, considering B, known as “posterior probability”. The algorithm, used in this investigation is introduced as follows:

1. One of the wells is considered as training well, that Bayesian technique should be trained on, the other is considered as generalization well to check generalization ability of Bayesian technique through it.

Whole data of training well is divided into three classes according to production rate: if a specific depth interval does not produce hydrocarbon, net pay value is considered 1, if the corresponding depth produces hydrocarbon

less than  $1500 \frac{\text{bbl oil}}{\text{day}} \left[ \cong 0.001104 \frac{\text{m}^3}{\text{sec}} \right]$ , then net pay value is 2, and finally if that

specific depth produces hydrocarbon more than  $1500 \frac{\text{bbl oil}}{\text{day}} \left[ \cong 0.001104 \frac{\text{m}^3}{\text{sec}} \right]$ ;

then, net pay value is 3. The intervals are shown on Figure 7 and 8 for training and generalization wells respectively.

2. Randomly, dataset of each class is divided into two parts: 70% for training Bayesian classifier, and 30% for testing the efficiency and validity of trained classifier.

3. In each production class (i.e. npv=1, 2 or 3), probability distribution function (PDF) of intake features are determined by training data. Features are generated curves, derived by the means of petrophysical relations or ANN from raw logs as were mentioned previously. E.g. PDF of porosity is determined three times by randomly selected training data: one time within intervals of npv=1, another time in intervals with npv=2, and the last time in the intervals, which npv=3. The same PDF determinations are done for other features: water saturation and shale volume.

4. Probability of belongings of each test data (30% of whole dataset) to each possible net pay value is calculated by the equation 4 (Duda et al., 2000).

$$P(npv) = \sum_{i=1}^n P(npv|d_i) \times P(d_i) \quad (4)$$

where n is number of input features,  $P(d_i)$  is probability of occurrence of i-th feature, which can be calculated simply by counting number of occurrences, divided by all samples, “npv” stands for “net pay value” and is described previously,  $P(npv|d_i)$  is named priory knowledge and can be calculated easily



by counting number of occurrences of a specific npv in all specific  $d_i$  conditions. This equation is called discrimination function and is a non-parametric type of Bayesian classifier.

5. The most probable net pay value is selected for each test data. E.g. if  $P(1)$  is bigger than  $P(2)$  and  $P(3)$ , then npv is defined 1 for that specific interval, hence the interval only produces water.

6. For comparing these net pay determination methods, Confusion Matrix (CM) and Classification Correctness Rate (CCR) are calculated.

In the present research, net pay value is determined two times by Bayesian classifier (two different procedures). The difference between these procedures is their input features. The inputs of first procedure are those of conventional cut-off based method (porosity, shale percentage and water saturation). But the inputs of second procedures are determined by Bayesian Network (porosity and water saturation).

### 3.3. Confusion Matrix and CCR

Confusion matrix evaluates preciseness of classifiers. This matrix shows that in which fraction, data of actual classes are classified among all possible classes (Theodoridis and Koutroumbas, 2003). CCR is a parameter which is calculated from confusion matrix, and is an index for evaluating the classifier. It is calculated by dividing summation of accepted cells (mostly on the trace of confusion matrix) by number of classes. In this work, CCR is calculated in two states: crisp (CCRC) and fuzzy (CCRf). Crisp state is when the expected output only accepts two values (i.e. binary: 0 or 1) but fuzzy state is when expected output can accept more than two states that is three here (i.e. npv=0, 1 or 2). For better understanding of difference of crisp and fuzzy states, look at Figure 7 and 8.

### 3.4. Bayesian Network

Bayesian Network (BN) is a directed acyclic graph. Each node of this graph represents a variable, and directed edges, outgoing of a node, show children of the parent node (note that occurrence of children nodes is dependent on the occurrence of parent nodes). The BN is useful in managing variables when an event is dependent on the probability of other events (Niedermayer, 2008).

Methods and algorithms for constructing Bayesian Networks are classified into two classes: constraint-based and score-based methods (Lauría, 2008). In this paper, a score-based method called K2 algorithm, is utilized to construct BN. For running this algorithm, the inputs should be ordered in a

way that shows dependency of parameters. The algorithm K2 is presented in appendix A. For a better understanding about BN and K2 algorithm, the respected reader is referred to (Doguc and Ramirez-Marquez, 2009).

#### 4. Results of Cut-off and Bayesian Methodologies

Both cut-off-based and fusion-based (Bayesian technique) methods are applied on datasets of two available wells. For the proposed method (fusion-based), one of the wells is used as training, the other as checking generalization ability. Conventional method serves sharp results whereas fusion-based method provides fuzzy output. The Classification Correctness Rate (CCR) is calculated for conventional output crisply (CCR<sub>c</sub>), and for fusion-based method is calculated both crisply (CCR<sub>c</sub>) and fuzzily (CCR<sub>f</sub>).

The Confusion Matrices and CCR values of net pay determination by these two methods (inputs are the same) in training well are shown on Figure 1. As it is revealed in the table, crisp CCR of Bayesian method shows an improvement of 9%, in comparison to crisp CCR value of conventional method. It is important to note that even fuzzy CCR of Bayesian method is higher than crisp CCR value of conventional method.

Decided (Conventional)	True Output (well test)			Decided (Bayesian)	True Output (well test)		
		1	2		3		1
0	$\begin{bmatrix} 0.73 & 0.08 & 0.04 \end{bmatrix}$			1	$\begin{bmatrix} 0.99 & 0.08 & 0.02 \end{bmatrix}$		
1	$\begin{bmatrix} 0.27 & 0.92 & 0.96 \end{bmatrix}$			2	$\begin{bmatrix} 0.01 & 0.87 & 0.11 \end{bmatrix}$		
	$CCR_c = \frac{0.73 + 0.92 + 0.96}{3} = 0.87$				$CCR_c = \frac{0.99 + 0.87 + 0.11 + 0.05 + 0.02}{3}$		
	$CCR_f = NAN$				$CCR_f = \frac{0.99 + 0.87 + 0.87}{3} = 0.91$		

A) Confusion Matrix and CCR of conventional (cut-off-based) method

B) Confusion Matrix and CCR of fusion-based method

Figure 1: Confusion Matrices and the way CCRs are calculated for both results of conventional and fusion-based methods (input parameters: shale volume, porosity and water saturation) in classifying whole well test intervals of training well. A) Conventional method. B) Fusion-based method

The trained Bayesian classifier (trained in the first well) is applied on the other well data (generalization well) to check generalization ability of the proposed method. The results of generalization are shown on Figure 2.

Higher precision of newly developed method is obvious when comparing crisp CCR values of these two methods: CCR<sub>c</sub> of fusion-based method (94%) is 34% higher than CCR<sub>c</sub> of conventional method (60%). Whereas CCR<sub>f</sub> of proposed method is lower than CCR<sub>c</sub> of conventional method, but it is not a drawback for fusion-based method, because its corresponding CCR<sub>c</sub> is higher than CCR<sub>c</sub> of cut-off-based method.

Decided (Conventional)	True Output (well test)			Decided (Bayesian)	True Output (well test)				
		1	2		3		1	2	3
	0	$\begin{bmatrix} 0.01 & 0.00 & 0.20 \end{bmatrix}$			1	$\begin{bmatrix} 1.00 & 0.00 & 0.17 \end{bmatrix}$			
1	$\begin{bmatrix} 0.99 & 1.00 & 0.80 \end{bmatrix}$			2	$\begin{bmatrix} 0.00 & 0.01 & 0.65 \end{bmatrix}$				
$CCR_c = \frac{0.01+1.00+0.80}{3} = 0.80$				$CCR_c = \frac{1.00+0.01+0.99+0.65+0.00+0.01+0.18}{3}$					
$CCR_f = NAN$				$CCR_f = \frac{1.00+0.01+0.18}{3} = 0.40$					

A) Confusion Matrix and CCR of conventional (cut-off-based) method

B) Confusion Matrix and CCR of fusion-based method

Figure 2: Confusion Matrices and the way CCRs are calculated for both results of conventional and fusion-based methods (input parameters: shale volume, porosity and water saturation) in classifying whole well test intervals of generalization well. A) Conventional method. B) Fusion-based method

#### 4.1. Input Selection by Bayesian Network

For selecting inputs of fusion-based method, one can simply choose conventional net pay determination inputs, like those utilized previously (shale percent, porosity and water saturation). The other choice is to run a specific input selection method. Here, Bayesian Network is used to select input features. To do so, K2 algorithm is utilized to generate the network. The available features to check the dependency of net pay value on them are: shale percent, porosity, permeability and water saturation. The order of input features, fed to the algorithm is:

shale percent- porosity- permeability- water saturation- net pay

The reason for selecting the mentioned order is that: Net pay is fixed at the end of this order to check its dependency on the other features. Porosity and permeability are located after shale percent, because porosity and permeability features are mainly controlled by lithological and structural phenomena. On the other hand, Shale volume is a parameter that lithology is based on. So, lithology is controlled by shale percentage, and porosity and permeability are controlled by lithology. It shows dependency of porosity and permeability on shale volume.

Porosity and permeability have interacting effect on each other (see Figure 3). Permeability is located after porosity, because permeability is a more sensitive parameter in comparison with porosity. To support this idea, crossplot of porosity- permeability is shown on Figure 3. In this figure, linear trend line with the equation of  $y=1.8x-9.2$  shows that the angle between this line and horizontal line is approximately 60 degrees (i.e.  $\tan^{-1} 1.8 \cong 61^\circ$ ), which is higher than 45 degree. So, any move on the porosity axis, results in a larger move on the axis of permeability due to sketched trend line. On the other hand:  $\frac{dy}{dx} = 1.8$ , which is higher than one. It shows higher speed of increase in y axis in comparison with x axis. Furthermore, permeability is a more sensitive parameter than porosity.

Mentioned math-based conclusion can be supported by this geological fact: In the same pore size, different pore arrangement will lead to different permeability values. In other words, different permeability values can occur for the same porosity, especially in carbonates where various types of pore arrangement exist, which are controlled by complex geometry of fractures. This rule is less true in the same porosity value, based on semi-vertical trend line on the Figure 3. Based on conclusions, it is considered that permeability is sensitive to porosity or porosity is a controlling factor of permeability.

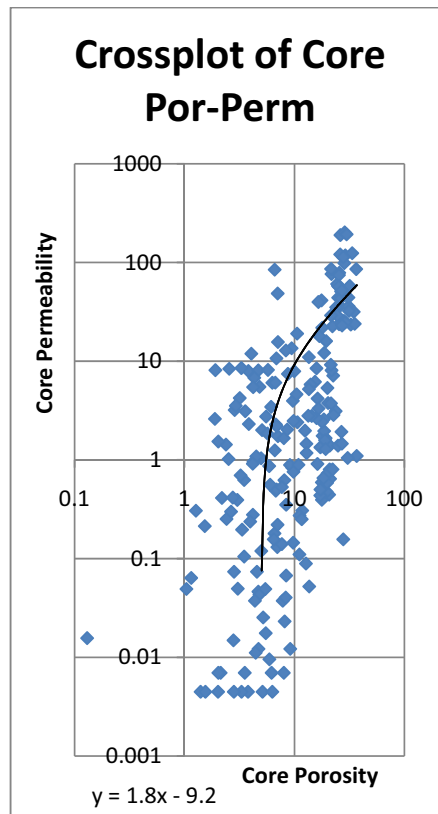


Figure 3: Crossplot of porosity-permeability with linear regression line. Equation of this line shows that the angle, created by this line and horizontal line is:  $\tan^{-1} 1.8 \cong 61^\circ$ . So, because of the angle being higher than 45 degree, any increase in porosity value, results in higher increase in permeability value.

Finally, water saturation is considered dependent on shale percent, porosity and permeability because usually this parameter is fixed after the others (on time series, water saturation fixes after hydrocarbon migration, whereas porosity and permeability are mostly constant before and after hydrocarbon migration). Figure 4 shows Bayesian Network, constructed by K2 algorithm.

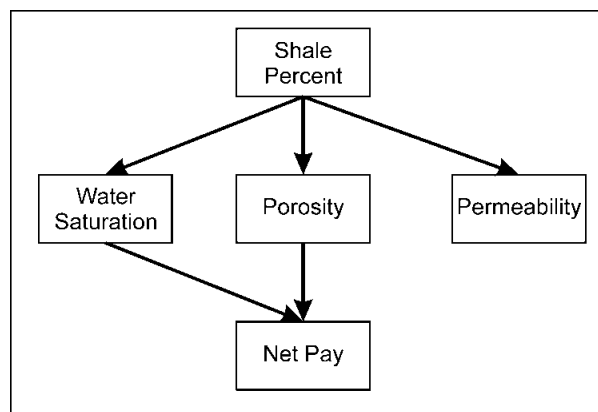


Figure 4: Bayesian Network constructed by K2 algorithm that shows dependency relation between petrophysical parameters and net pay.

As it is visible in the generated Bayesian Network, the net pay feature is directly dependent on porosity and water saturation parameters. It is also indirectly dependent on shale percent value. But permeability value is not an affecting feature on net pay value in presence of the other three parameters. Absence of permeability has two positive effects: One of them is to remove error of estimation of permeability; the other is to keep procedure of net pay determination as simple as possible in absence of permeability in calculations.

By considering results of BN, all three parameters of shale percent, porosity and water saturation (those features that are being used in conventional method) are effective features on net pay value but for continuing this investigation, only direct features are selected for net pay determination to compare its accuracy with previous fusion-based net pay determination procedure, which used three input features. The confusion matrices and CCR values of net pay determination by two input features (porosity and water saturation) are presented below. Figure 5-A and B show confusion matrices and CCR values of net pay determination in training and generalization wells, respectively.

Decided (Bayesian)		True Output (well test)			Decided (Bayesian)		True Output (well test)		
		1	2	3			1	2	3
	1	[ 1.00 0.13 0.02 ]				1	[ 1.00 0.00 0.09 ]		
	2	[ 0.00 0.78 0.10 ]				2	[ 0.00 0.01 0.76 ]		
3	[ 0.00 0.09 0.88 ]			3	[ 0.00 0.99 0.15 ]				
		$CCR_c = \frac{1.00 + 0.78 + 0.10 + 0.09 + 0}{3}$					$CCR_c = \frac{1.00 + 0.01 + 0.99 + 0.76 + 0}{3}$		
		$CCR_f = \frac{1.00 + 0.78 + 0.88}{3} = 0.89$					$CCR_f = \frac{1.00 + 0.01 + 0.15}{3} = 0.39$		

A) Confusion Matrix and CCR of net pay determination in training well

B) Confusion Matrix and CCR of net pay determination in generalization well

Figure 5: Confusion Matrices and the way CCRs are calculated in net pay determination by fusion-based method with two input parameters (porosity and water saturation) in classifying whole well test intervals of A) training well, and B)

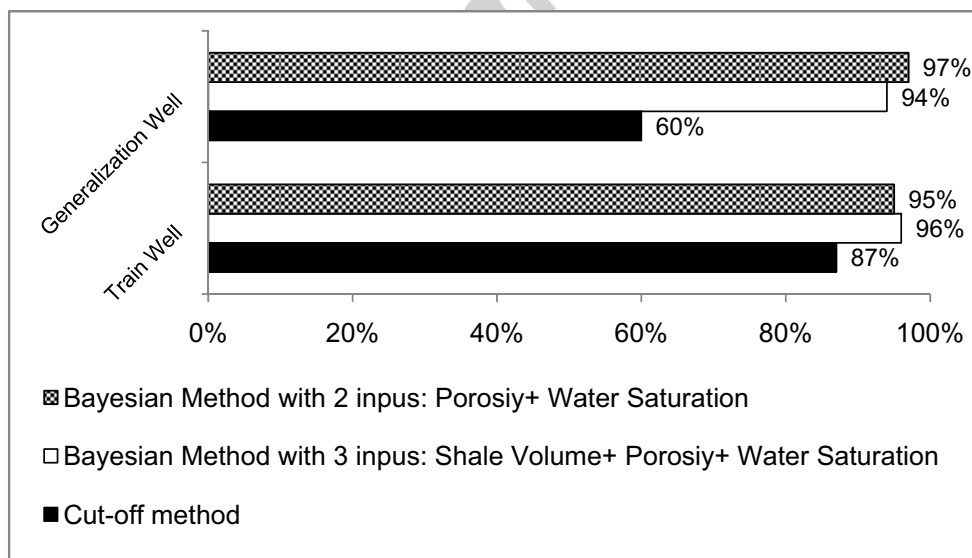
generalization well.

## 5. Discussion: the Best Approach

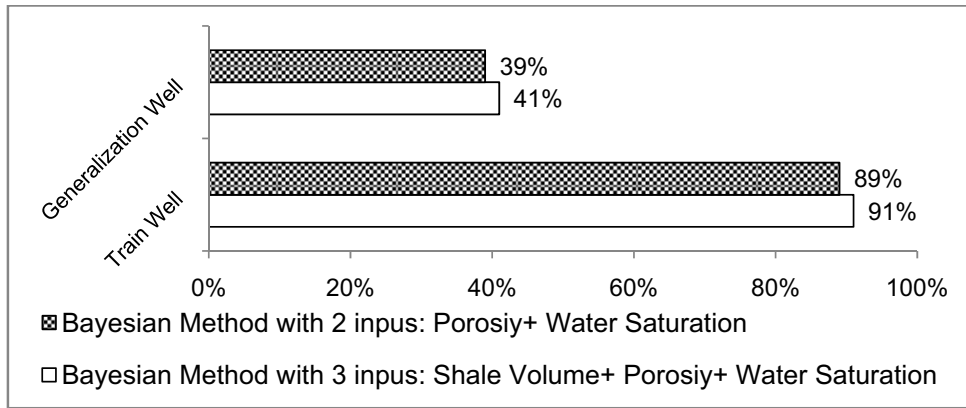
Figure 6 compares all three introduced net pay determination procedures. On Figure 6-A, crisp CCR values are compared. It is obvious that conventional cut-off-based method (black bars) has lower accuracy in comparison with two proposed fusion-based methods (based on lower CCRc) in both training and generalization wells. Hence; the most important advantage of the proposed methodology is its higher precision in comparison with conventional cut-off-based method.

There are two indices (CCRc and CCRf) for comparing two proposed methods with each other. Considering CCRc values (Figure 6\_A), it is impossible to give priority to one of the Bayesian methods. Because three-input-Bayesian procedure shows higher precision in training well while two-input-Bayesian procedure shows higher precision in generalization well. Hence; in crisp determination, these two procedures have the same precision.

Although three-input-Bayesian procedure shows higher CCRf, difference of CCRf values is negligible (only 2%). Furthermore; it is logical to use two-input-Bayesian procedure not to include error of estimation of shale percent, and keep the model as simple as possible.



A) Crisp CCR values (CCRc) of three mentioned methods



B) Fuzzy CCR values (CCR<sub>f</sub>) of two fusion-based method.

Figure 6: Comparing CCR values of mentioned methods in determining net pays.

A) Comparison of crisp CCRs B) Comparison of fuzzy CCRs

To have a more realistic sense and better understanding of differences between results of these methodologies, determined net pay values are sketched and compared to well tests in both training and generalization wells (Figure 7 and 8). Though two fusion-based methods and well test results are relatively similar to each other, there are considerable differences between conventional method and well tests.

#### Comparing Outputs in well Test Intervals of Training Well

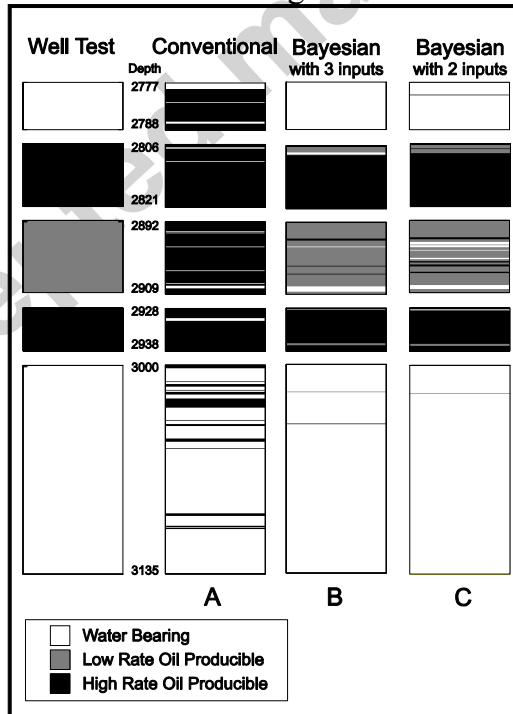




Figure 7: Results of three net pay determination procedures are compared with well test in training well. The track (A) is results of conventional cut-off-based method; track (B) shows results of Bayesian method with three input features (porosity, water saturation and shale content); and right track (C) reveals results of Bayesian method with two input features (porosity and water saturation).

According to Figure 8, conventional method has classified nearly all the depths as productive zone, which is fully incompatible with well test results. But Bayesian method has determined productive zones from non-productive zones precisely. Outputs of both Bayesian methods are very close to each other, so it is difficult to give priority to each of them. A drawback of these fusion-based methods is in distinguishing high rate and low rate productive zones that are classified reversely in generalization well.

Comparing Outputs in well Test Intervals of Generalization Well

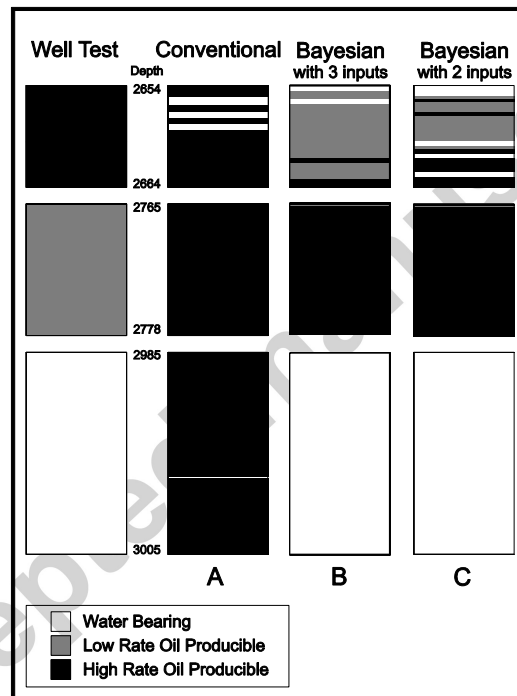


Figure 8: results of three net pay determination procedures are compared with well test results (in generalization well). The track (A) is results of conventional cut-off-based method; track (B) shows results of Bayesian method with three input features (porosity, water saturation and shale content); and right track (C) reveals results of Bayesian method with two input features (porosity and water saturation).

Above figures show that performance of conventional method is weak below the oil water contact (OWC), i.e. lowermost well test, whereas proposed method performs so precisely below OWC. Definitely by removing lower OWC from investigation, preciseness of conventional method will increase. However, because there is no need to find oil water contact to determine net pays, it is still a great advantage of the proposed method in comparison with the conventional method. It is worthy of note that finding

OWC is usually difficult and sometimes impossible, specifically in fractured carbonate reservoirs. Besides, usually there is no sharp boundary for OWC, because it is a transition zone.

A fundamental drawback of proposed method is its higher resolution of well logs in comparison with well test intervals. As it is shown on both Figure 7 and 8, each well test interval contains interpreted thinner layers with different productivity rates. Although in each well test there are thinner layers with different productivity rates, the same rate is considered for all of them to train Bayesian classifier. Despite this fundamental disadvantage, newly developed method is still better than conventional method because of final outputs. Based on outputs, Bayesian method represents a more precise classification in comparison with conventional method. So, newly developed method is recommended for net pay determination. To diminish effect of this resolution inconsistency, it is suggested to increase number of well tests and decrease thickness of each test interval.

The other disadvantage of proposed method is that it may not result in satisfactory outputs in heterogeneous environments in which samples and observations do not cover all different situations. This pitfall rises from essence of data driven techniques that Bayesian is one of them.

## 6. Conclusion

In this paper, a new methodology for net pay determination is developed, and its results are compared to a well-known conventional cut-off-based method.

Bayesian technique can be used as a fusing method to integrate petrophysical data/information to determine net pays more precisely. Eight to 37% improvement in preciseness of classifying net pays in comparison with conventional cut-off-based method.

By means of Bayesian theory, it is possible to determine net pays fuzzily with appropriate preciseness. CCR of training well is around 90% and around 40% for CCR of generalization well.

Bayesian Network is a suitable tool for feature selection (reduction in dimension of input features) for net pay determination by Bayesian technique. Bayesian Network introduces porosity, water saturation and shale volume as the most suitable features to determine productive zones.

A great advantage of net pay determination by proposed method is that there is no need to find oil water contact (OWC) to find net pays. Sometimes it is difficult to find this boundary mostly because there is not a sharp line or boundary. Actually OWC is a transition zone.

Also, there is no need to identify cut-offs for proposed method. It is another great advantage of Bayesian method in complex reservoirs. Cut-off determination is usually difficult, and sometimes impossible in some reservoir rocks.

Although each well test interval consists of different conditions, Bayesian method considers productivity of whole interval of each well test the same. It is the main drawback of Bayesian method. But based on the outputs, newly developed method is more precise than conventional methodology. For solving this problem and coming to more reliable results, it is suggested to use thin-interval well tests and as much available well tests as there are.

The other pitfall of Bayesian method is due to essence of data driven methods. That is insufficient training of Bayesian classifier, when there are different conditions, and the available data do not cover all variations, consequently will result in inappropriate outputs.

### **Acknowledgment**

The authors wish to acknowledge Exploration Directorate of National Iranian Oil Company (NIOC) for providing data and information of this investigation. Also great thank to Mr. Amir Mollajan and other four anonymous reviewers for their valuable comments that have improved the quality of the paper.

### **Appendix A**

Algorithm K2(T,u), from (Doguc and Ramirez-Marquez, 2009):

Input: T is a matrix with n columns (i.e. features or nodes) and m rows (i.e. observations), and u is the maximum possible number of parents for each node.

Output: An acyclic directed graph, called BN.

1. For column  $i=1$  to n in dataset T:
  - Create node  $X_i$  and add it to BN.
  - Consider  $\pi_i = \emptyset$  for node  $X_i$ .

Calculate scoring function,  $f(i, \pi_i)$ , using empty set:

$$f(i, \pi_i) = \prod_{j=1}^{q_i} \frac{(d_i - 1)!}{(\alpha_{ij} + d_i - 1)!} \prod_{k=1}^{d_i} \alpha_{ijk}! \quad (\text{A.1})$$

where  $q_i$  is number of subsets of the parent set for node  $X_i$ ,  $d_i$  is size of the input domain, which here is six (each continuous input value is converted to six discrete values: 0, 1, 2, 3, 4 and 5).  $\alpha_{ijk}$  is number of rows that both parent set  $j$  and node  $X_i$  have the same value of  $k$ : 0, 1, 2, 3, 4 or 5, and finally  $\alpha_{ij}$  is summation of all  $\alpha_{ijk}$  values, i.e.  $\alpha_{ij} = \sum_{k=1}^{d_i} \alpha_{ijk}$ .

- Until the size of  $\pi_i$  is equal or less than  $u$ :
  - Consider  $X_z$  as a node prior to the node  $X_i$ .

Calculate scoring function another time, using  $X_z$ :  $f(i, \pi_i \cup \{X_z\})$

If  $f(i, \pi_i \cup \{X_z\}) > f(i, \pi_i)$  then add  $X_z$  to  $\pi_i$ .

2. Return BN.

## References

- Challa, S. and Koks, D., 2004. Bayesian and Dempster-Shafer fusion. *Sadhana*, 29(2): 145-174.
- Connell, D.L., Coates, J.A. and Frost, D.A., 1986. Development of a Fluorimetric Method for Detection of Pay Zones During Drilling With Invert Muds. *SPE Formation Evaluation*, 1(6): 595-602.
- Cooke-Yarborough, P., 1984. Reservoir Analysis By Wireline Formation Tester: Pressures, Permeabilities, Gradients And Net Pay. *The Log Analyst*, 15(6): 36-46.
- Darling, T., 2005. Quicklook log interpretation, Well Logging and Formation Evaluation. Gulf Professional Publishing, Burlington, pp. 29-48.
- Deakin, M. and Manan, W., 1998. The Integration of Petrophysical Data for the Evaluation of Low Contrast Pay, SPE Asia Pacific Conference on Integrated Modelling for Asset Management. 1998 Copyright 1998, Society of Petroleum Engineers, Inc., Kuala Lumpur, Malaysia, pp. 327-339.
- Doguc, O. and Ramirez-Marquez, J.E., 2009. A generic method for estimating system reliability using Bayesian networks. *Reliability Engineering & System Safety*, 94(2): 542-550.
- Duda, R.O., Hart, P.E. and Stork, D.G., 2000. Pattern Classification. Wiley India Pvt. Ltd.
- Flower, J.G., 1983. Use of Sonic-Shear-Wave/Resistivity Overlay as a Quick-Look Method for Identifying Potential Pay Zones in the Ohio (Devonian) Shale. *SPE Journal of Petroleum Technology*, 35(3): 638-642.
- Ghabeishavi, A., Vaziri-Moghaddam, H., Taheri, A. and Taati, F., 2010. Microfacies and depositional environment of the Cenomanian of the Bangestan anticline, SW Iran. *Journal of Asian Earth Sciences*, 37(3): 275-285.
- Hall, D.L. and Llinas, J., 2001. Handbook of multisensor data fusion. CRC Press LLC, Boca Raton.
- Jalali Lichaei, M. and Nabi Bidhendi, M., 2006. Comparison between Multiple Linear Regression and Artificial Neural Networks for Porosity and Permeability Estimation. *Geosciences Scientific Quarterly Journal*, 61: 140-149.

- Jensen, J.L. and Menke, J.Y., 2006. Some Statistical Issues in Selecting Porosity Cutoffs for Estimating Net Pay. *PetroPhysics*, 47(4): 315–320.
- Johnson, D.E. and Pile, K.E., 2002. Formation Parameters. In: M.M. Patterson (Editor), *Well Logging in Nontechnical Language*. Penn Well Publishing Company, Tulsa, pp. 25-44.
- Lauría, E., 2008. An Information-Geometric Approach to Learning Bayesian Network Topologies from Data. In: D. Holmes and L. Jain (Editors), *Innovations in Bayesian Networks. Studies in Computational Intelligence*. Springer Berlin / Heidelberg, pp. 187-217.
- Masoudi, P., Tokhmechi, B., Jafari, M.A. and Moshiri, B., 2012. Application of Fuzzy Classifier Fusion in Determining Productive Zones in Oil Wells. *Energy Exploration and Exploitation*, 30(3): 403-415.
- Masoudi, P., Tokhmechi, B., Zahedi, A. and Jafari, M.A., 2011. Developing a Method for Identification of Net Zones Using Log Data and Diffusivity Equation. *Journal of Mining and Environment*, 2(1): 53-60.
- Mathur, N., Raju, S.V. and Kulkarni, T.G., 2001. Improved Identification of Pay Zones through Integration of Geochemical and Log Data: A Case Study from Upper Assam Basin, India. *AAPG Bulletin*, 85(2): 309-323.
- Millikan, C.V., 1925. Use of Gas Meters for Determination of Pay Strata in Oil Sands. 183-195.
- Mohaghegh, S.D., Goddard, C., Popa, A., Ameri, S. and Bhuiyan, M., 2000. Reservoir Characterization Through Synthetic Logs, SPE Eastern Regional Meeting. Copyright 2000, Society of Petroleum Engineers Inc., Morgantown, West Virginia, pp. 1-10.
- Niedermayer, D., 2008. An introduction to Bayesian networks and their contemporary applications. In: D.E. Holmes and L.C. Jain (Editors), *Innovations in Bayesian Networks, Theory and Applications. Studies in computational intelligence*. Springer, Berlin, pp. 117-130.
- Singleton, S., 2008. The use of seismic attenuation to aid simultaneous impedance inversion in geophysical reservoir characterization. *The Leading Edge*, 27(3): 398-407.
- Snyder, R.H., 1971. A Review of the Concepts and Methodology of Determining "Net Pay", Fall Meeting of the Society of Petroleum Engineers of AIME. Society of Petroleum Engineers, New Orleans, Louisiana, pp. 1-12.
- Svec, R.K. and Grigg, R.B., 2000. Reservoir Characterization and Laboratory Studies Assessing Improve Oil Recovery Methods for the Teague-Blinebry Field, SPE Permian Basin Oil and Gas Recovery Conference. 2000,. Society of Petroleum Engineers Inc., Midland, Texas, pp. 1-12.
- Theodoridis, S. and Koutroumbas, K., 2003. *Statistical Pattern Recognition*. Elsevier, USA.
- Worthington, P.F., 2000. Recognition and evaluation of low-resistivity pay. *Petroleum Geoscience*, 6(1): 77-92.
- Worthington, P.F., 2008. The Application of Cutoffs in Integrated Reservoir Studies. *SPE Reservoir Evaluation & Engineering*, 11(6): pp. 968-975.
- Worthington, P.F., 2010. Net Pay-What Is It? What Does It Do? How Do We Quantify It? How Do We Use It? *SPE Reservoir Evaluation & Engineering*, 13(5): pp. 812-822.
- Worthington, P.F., 2011. A Data-Driven Approach to Identify Bet-Pay Cutoffs. *World Oil*, 232: 49-56.
- Worthington, P.F. and Cosentino, L., 2005. The Role of Cut-offs in Integrated Reservoir Studies. *SPE Reservoir Evaluation & Engineering*, 8(4): pp. 276-290.

