PRJ, Vol. 5 (January 1993); P. 1–5, 3 TABLES.

IMPROVEMENT OF THE ACCURACY OF STATISTICAL WELL LOG INTERPRETATION USING THE MATERIAL BALANCE EQUATION

G. Szigeti*

دقة الإستقراء الإحصائي لسرود الآبار باستخدام معادلة توازن المواد

ج. زيقيتي

عند تحديد مسامية كتلة الصخور داخل البئر النفطي تتم مقارنة القيم المقاسة بتلك التي تم حسابها من النموذج الإفتراضي للصخور. وتنتج عن هذه المقارنة معادلات اتفاقية. وعلى افتراض أن السجل المقاس متغير عشوائي بقيم متوقعة معطاة من النموذج النظري وبتباين معلوم، فانه بالإمكان تحديد التقدير الإحصائي والحجم الجزئي للمسام وللمكونات المعدنية وتباين تلك الكميات.

وباستخدام معادلة توازن المواد في مجموعة المعادلات الإتفاقية، كمعادلة اتفاقية تؤول إلى إنحراف معياري صفري، يقل التباين بالنسبة للمسامية والعناصر المعدنية الأخرى. كما أن استخدام معادلة توازن المواد يستوجب التعامل مع المحددة الدورية بدلاً عن المحددة المعتادة.

ABSTRACT

In the determination of the porosity of the rock bulk drilled through the oil well, we compare the measured physical quantities and the calculated quantities given by the idealized, theoretical rock model. The comparison gives stochastical equations. Supposing that the measured log is a random variable with expected value given by the theoretical model and with known variance, we can determine the statistical estimation of the volume fraction of the porosity and other mineral components and the variance of those quantities. If we involve the exact, theoretical material balance equation into the system of stochastical equations, as a stochastical equation tending to zero standard deviation, then the variance of the porosity and other mineral components will be decreased. The consideration of the material balance equation requires the use of the cyclical determinant instead of the regular determinant.

INTRODUCTION

An important task of the log survey is to determine the porosity of the rock bulk drilled through the oil well. For this reason the minerals of the given rock matrix are grouped on the basis of their physical properties which affect the given logs, and the task is to determine the volume fractions of each group. During this process all available logs which help to solve the problem are put in use.

The method of finding the volume fractions is to compare the measured physical quantities with the calculated quantities given by the theoretical description of idealized rock model containing the parameters.

Because of various reasons (see [1]) the measured quantity is a stochastical one; accordingly we have to consider it as a random variable. If only the first moments are known (i.e. the expected values) of these random variables, then as it is proved in [2], we can not get information on the rock volumes. Nevertheless knowing the second moments as well (by which the standard deviations of these variables also will be known) and assuming that, first, the measured quantities vary around the values of the theoretical model and second, rather large (infinitely large) number of observations, there is a possibility to restore the volume fraction parameters of the theoretical model. Enforcing the first assumption and taking into consideration that in a given depth point only a few stochastical observations are available, it is clear that the parameters to be determined are also stochastical quantities i.e. we can

* Petroleum Research Centre, P.O. Box 6431, Tripoli, G.S.P.L.AJ.

determine only the probability of the event, that the parameter belongs to a given section around the expected value.

During the process we consider the theoretical rock model which is defined by us, however not arbitrarily, as a certain model and not as a stochastic model. As a consequence, there is a basic difference between the stochastic vector equations of measured and theoretical values, and between the exact, purely theoretical material balance equation.

In the limited number of publications dealing with the complex process of interpretation of well log measurements (see e.g. [4], [5]) we could not find information about the fact, or about the manner of the above mentioned difference. According to our knowledge, part of log analysts proceed the following way: first they determine the volume fractions without using the material balance equation, and after that, they interpret the difference of the sum of components from the unit as a quantitative description of the quality of the interpretation process. We consider this procedure from a logical point of view, as a correct one, but not as the possible complete one, because it's inaccurate to ignore the only certain, exact equation in the determination of the volume fractions of mineral components.

FORMULATION OF THE PROBLEM AND SUMMARY OF THE STATISTICAL RESULTS

Let us suppose that we have m different log data at given depth point of the bore hole, and we divided the unit volume of the surrounding rock for $k+1$ mineral components as $x_0, x_1, \ldots, x_k$ so that

$$x_0 + x_1 + \cdots + x_k = 1 \tag{1}$$

(We begin to enumerate the x_i components from zero, because we suppose that the rock has k solid components and x_0 is the fluid component, which is in our case the main parameter, namely the porosity).

Let us suppose that each log can be described as a linear function of unknowns $x_0, x_1, \ldots, x_k$, and all the measured logs are random variables according to the one dimensional normal distribution, with known standard deviations $\sigma_1, \sigma_2, \ldots \sigma_m$ and with expected values which are the results of that linear function in a certain, unknown point $\hat{x}_0, \hat{x}_1, \ldots, \hat{x}_k$. Normalizing the coefficients of the first linear function and the first measured value by σ_1, the seconds by σ_2 and so on, we can write

$$\begin{aligned} a_{10} \cdot \hat{x}_0 + a_{11} \cdot \hat{x}_1 + \cdots + a_{1k} \cdot \hat{x}_k &= b_1 \\ a_{20} \cdot \hat{x}_0 + a_{21} \cdot \hat{x}_1 + \cdots + a_{2k} \cdot \hat{x}_k &= b_2 \\ &\vdots \\ a_{m0} \cdot \hat{x}_0 + a_{m1} \cdot \hat{x}_1 + \cdots + a_{mk} \cdot \hat{x}_k &= b_m \end{aligned} \tag{2}$$

where a_{ij} is the normalized jth coefficient of the ith linear tool response function, b_i is the normalized ith log, which is not necessarily the originally measured log, but a modified form of that, after different transformations like filtering, depth matching, etc.

According to the assumptions the right hand side of (2) is a random variable with normal distribution. The expected value of each b_i is the expression standing on the left hand side of the ith equation and the variance of each b_i is 1. In addition we shall suppose, that for any $i \neq j$ the b_i and b_j stochastic variables are independent of each other.

Denoting the set of matrices containing n rows and m columns by $\mathcal{M}_{n,m}$ and the set of vectors, containing k elements as one row by $\mathcal{V}_k$, the set of equations (2) can be rewritten in matrix form

$$A\underline{\hat{x}} = \underline{b} \tag{3}$$

where $A \in \mathcal{M}_{m,k+1}$, $\underline{b} \in \mathcal{V}_m$, $\underline{x} \in \mathcal{V}_{k+1}$ and $\underline{\hat{x}} = (\hat{x}_0, \hat{x}_1, \ldots, \hat{x}_k)$.

We consider that the statistical results, which can be summarized as follows, are well known to the reader (see e.g. [3], page 260). Keeping in force all the assumptions related to $\underline{b}$, among all the linear statistical estimations $\underline{\alpha}$ for the solution of the overdetermined linear system of equations (3), the smallest variance given by that $\underline{\alpha}$, which satisfies the Gauss normal equation

$$(A^T \cdot A)\underline{\alpha} = A^T \cdot \underline{b} \tag{4}$$

(A^T is the transpose of matrix A).

In this case $\underline{\alpha}$ is unbiased estimation of $\underline{\hat{x}}$, i.e. the expected value of $\underline{\alpha}$ is $\underline{\hat{x}}$.

In this way according to Cramer's rule, the ith component of $\underline{\alpha}$ can be calculated from

$$\alpha_i = \frac{\det[(A^T \cdot A)^i]}{\det[(A^T \cdot A)]} \tag{5}$$

where $A^T \cdot A \in \mathcal{M}_{k+1,k+1}$; $(A^T \cdot A)^i \in \mathcal{M}_{k+1,k+1}$ is the matrix in which the ith column of the matrix is changed to the right hand side vector of (4). The variance of α_i, denoted $D^2(\alpha_i)$ can be calculated from

$$D^2(\alpha_i) = \frac{\det[(A^T \cdot A)_{ii}]}{\det[(A^T \cdot A)]} \tag{6}$$

where $A^T \cdot A$ is the same as above, and $(A^T \cdot A)_{ii} \in \mathcal{M}_{k,k}$ is the matrix obtained from the matrix $A^T \cdot A$ by omitting it's ith column and it's ith row.

INVOLVEMENT OF THE MATERIAL BALANCE EQUATION INTO THE STOCHASTICAL SYSTEM

During the investigation of the system of equations (2) all the equations had a statistical character. However, equation (1), is not stochastical but a certain one. The simplest way to eliminate the difference between equation (1) and equation (2), is to handle the right hand side of (1) as a stochastical variable with unit expectancy and with a variance, which tends to zero.

Accordingly let ξ_e be a stochastical variable with Gaussian normal distribution, with the expectation value 1, and with the variance σ^2,

$$M(\xi_e)=1; \quad D^2(\xi_e)=\sigma^2 \tag{7}$$

Denoting $1/\sigma$ by s and so $s=1/\sigma$ tends to infinity, then the stochastical variable $s\cdot\xi_e$ shows the behavior

$$M(s\cdot\xi_e)=s; \quad D^2(s\cdot\xi_e)=1 \tag{8}$$

Handling equation (1) in this way, now the system of equations (1) and (2) for $x_0, x_1, \ldots, x_k$ looks like

$$\begin{aligned} &s\cdot x_0+s\cdot x_1+\cdots+s\cdot x_k=b_0\\ &a_{10}\cdot x_0+a_{11}\cdot x_1+\cdots+a_{1k}\cdot x_k=b_1\\ &\quad\vdots\\ &a_{m0}\cdot x_0+a_{m1}\cdot x_1+\cdots+a_{mk}\cdot x_k=b_m \end{aligned} \tag{9}$$

where

$$b_0=s\cdot\xi_e \tag{10}$$

We shall apply the same statistical results for the system of equations (9), which have been applied for the system (3). Accordingly, taking the special structure of (9) in consideration, in analogy with (4), the unbiased statistical estimation $\underline{\beta}$ of $x_0, x_1 \ldots x_k$ of (9) fulfills the equation

$$(A^T\cdot A+s^2\cdot E)\cdot\underline{\beta}=A^T\cdot\underline{b}+s^2\cdot\underline{e} \tag{11}$$

where $A\in\mathcal{M}_{m,k+1}$ is the matrix of equation (2), $\underline{b}\in\mathcal{V}_m$ is the vector on the right hand side of the same equation. $E\in\mathcal{M}_{k+1,k+1}$ is the matrix each element of which is unit, while s is given by (10). In the case when s tends to infinity, the estimation $\underline{\beta}$ of (9) and (10) gives the statistical estimation of the system consisted of (1) and (2).

So, in full analogy with (5) and (6) we get

$$\beta_i=\lim_{s\to\infty}\frac{\det[(A^T\cdot A+s^2\cdot E)^i]}{\det[(A^T\cdot A+s^2\cdot E)]} \tag{12}$$

and

$$D^2(\beta_i)=\lim_{s\to\infty}\frac{\det[(A^T\cdot A+s^2\cdot E)_{ii}]}{\det[(A^T\cdot A+s^2\cdot E)]} \tag{13}$$

(Note that according to (11) at the calculation of the numerator of (12) the $(A^T\cdot A+s^2\cdot E)^i$ matrix preserves its structure, having the s^2 second term after replacing the ith column by the right hand side of (11), because the right hand side has the same second term).

The limit process in (12) and (13) can be derived in a simple form. On the basis of algebraic type results given in the Appendix, the derivation gives us the following final form for the statistical estimation $\underline{\beta}$:

$$\beta_i=\frac{\text{cycl det}[(A^T\cdot A)^i]}{\text{cycl det}[(A^T\cdot A)]} \tag{14}$$

and

$$D^2(\beta_i)=\frac{\text{cycl det}[(A^T\cdot A)_{ii}]}{\text{cycl det}[(A^T\cdot A)]} \tag{15}$$

where all the matrices in square brackets are exactly the same as in formulas (5) and (6), but instead of the regular determinant, β_i and $D^2(\beta_i)$ can be expressed by the "cycl det" operator, which is defined as follows:

For an arbitrary $C\in\mathcal{M}_{n,n}$ matrix cycl det is the sum of n determinants:

$$\text{cycl det } C=\sum_{i=1}^{n}\det[C_i] \tag{16}$$

where $C_i\in\mathcal{M}_{n,n}$ is developed from C by replacing its ith column by $\underline{e}\in\mathcal{V}_n, \underline{e}^T=(1, 1\ldots, 1)$.

Comparing (14), (15) with (5), (6) we can say, that the effects of consideration of the material balance equation in statistical estimation of volume fraction parameters, is that instead of the regular determinant calculation, we have to calculate the cyclical determinant of the same matrixes.

EXAMPLES

First we illustrate how the two different approaches work on a simplified "artificial" case. The rock matrix consists of the quartz component only. Our task is to determine the porosity filled up by interstitial water, using the acoustic, density and hydrogen index logs, and to determine the standard deviation (the square root of the variance) of this quantity, which displays a Gauss normal distribution

as a result of the assumption that the measured logs show Gauss normal distribution as well.

Table 1a shows the slowly changing, or the so called zone parameters of each rock component for each log and the standard deviations of the measurements.

The calculated results for the standard deviation of the porosity are given in Table 1b using the three different possible combinations from the three different logs, without the use of (formula (6)) and with the use of (formula 15)) of the material balance equation. It can be noticed that the standard deviation is decreasing in all cases as a result of the presence of the material balance equation. The largest decrease is shown by the (HI, Δt) pair, while the most accurate way of solving the porosity determination, using two out of the three logs above, is offered by the (HI, ρ_b) combination.

The second example is a more complex one from the practice of geophysical interpretationl. Here the bore hole intersects fractured granite rocks while the fractures are partly filled with zeolite as a result of hydrothermal processes. We have neutron density (ρ_b), neutron porosity (HI), acoustic (Δt), photoelectric index (P_e) and two different spectral gamma ray (Th and K) logs. According to these conditions quartz (Qz), feldspar (FP), mica and heavy minerals (M & Hm), zeolit (Zeo) and total porosity (ϕ_t) filled up by interstitial water are the main mineral groups in the volumetric rock model. The upper part of Table 2 contains the zone parameters and the standard deviations of the log measurements. The lower part of Table 2 shows the standard deviations of the volume fractions for each mineral group without and with the material balance equation taken into account. Comparison of the two last rows of Table 2 shows again that the uncertainty of the interpretation, belonging to a given level of probability (in our case, i.e. in the case of Gaussian distribution and standard deviation this probability level is 0.6827) decreases by involving the material balance equation into the estimation. This decrease is less in the case of Table 2, than it was in the case of Table 1, which is related to the fact that the dimension of the equation system, in the second case, is higher.

Table 1a. Slowly Changing Parameters of the Two Mineral Components Model

Tool	Mineral group		Standard deviation of the measurement
	Water	Quartz	
acoustic, Δt [μsec/foot]	185.0	52.0	5.0
density, ρ_b [g/cm^3]	1.08	2.65	0.05
hydrogen index, HI [dimensionless]	0.93	−0.02	0.01

Table 1b. Standard Deviation of Porosity Determination in Percentage Without and With the Material Balance Equation Being Taken into Consideration

Used logs	Standard deviation (%)	
	without	with
(Δt, ρ_b)	2.47	2.43
(HI, ρ_b)	1.92	1.00
(HI, Δt)	9.62	1.14

Table 2. Slowly Changing Zone Parameters of the Five Mineral Components Model and the Standard Deviation of the Volume Fraction in Percentage Without and With the Material Balance Equation Being Taken into Consideration

Tool		Mineral Group					Standard deviation of the measurement
		t	Q_z	FP	M & Hm	Zeo	
ρ_b [g/cm^3]		1.08	2.65	2.55	3.01	2.11	0.05
HI [dim. less]		0.93	−0.02	−0.03	0.05	0.01	0.01
Δt μsec/foot]		185.0	52.0	55.0	41.0	60.0	5.0
Pe [barn/electron]		1.13	1.85	2.86	4.2	1.78	0.5
Th [ppm]		2.6	1.8	17.0	40.0	50.0	1.0
K [dim. less %]		0.68	0.68	7.0	4.0	7.0	1.0
Standard deviation %	without	1.03	6.09	8.32	5.68	6.98	
	with	1.00	5.51	8.29	3.79	4.86	

SUMMARY

In calculating the volume fractions of mineral components by means of complex well log interpretation, the accuracy of the determination of these fractions can be increased by involving the material balance equation into the interpretation system. The calculation of the standard deviations of the volume percentages becomes only slightly more difficult, as a result of taking the material balance equation into consideration. Whereby the regular determinant is replaced by the cyclical determinant operation. Practical examples show, that in some cases the increase in accuracy is significant.

ACKNOWLEDGEMENT

I would like to thank Mr. Mustafa Idris, the General Manager of the Petroleum Research Centre for Providing the opportunity to continue this research.

The continuous stimulating consultation with Prof. Zoltan Barlai which helped considerably in the process of this study is highly acknowledged.

REFERENCES

[1] Theys P., 1988, Log Quality Control and Error Analysis: A Prerequisite to Accurate Formation Evaluation, Eleventh European Formation Evaluation Symposium.

[2] Halfin L.A., 1958, Informacionnaja teorija interpretacii geofzicsszkih iszledovanij [The Information Theory of the Interpretation of Geophysical Investigation], (in russian): D.A.N. CCCP T 122 No. 6 [Reports of the Academy of Sciences of USSR, Volume 122, No. 6].

[3] Vincze I., 1968, Matematikai statisztika ipari alkalmazasokkal [Mathematical Statistic with Industrial Applications], (in hungarian), Muszaki Konyvkiado, Budapest: Technical Publishing Company).

[4] Rodriquez E., A. Mezzatesta, D. Tetzlaff, 1989, Determination of Statistical Confidence Intervals for Petrophysical Formation Properties: SPWLA Laser Symposium, London.

[5] Mayer C., A. Sibbit, 1980, Global, a New Approach to Computer processed Log Interpretation, 55th Annual Conference of Society of Petroleum Engineers, Dallas.

APPENDIX

Here we dicuss the limit process used in the derivation of formulas (14) and (15).

Let $C \in \mathcal{M}_{n,n}$ be an arbitrary matrix and $E \in \mathcal{M}_{n,n}$ the matrix, each element of which is the unit. Then for any number p

$$\mathrm{dt}(C+pE)=p\cdot\mathrm{cycl\ det}\ C+\det C \tag{a1}$$

here we use the notations used in the paper. The definition of the $\mathcal{M}_{n,n}$ can be found before expression (3) and the definition of the operator "cycl det" is given by formula (16).

For providing eq. (a1) let us denote the column vectors of C by $\underline{c}_1, \underline{c}_2, \ldots \underline{c}_n$ and the columns of E by $\underline{e}$. Then we have

$$\det(C+pE)=\det(\underline{c}_1+p\cdot\underline{e}_1\underline{c}_2+p\cdot\underline{e}\ldots,\underline{c}_n+p\cdot\underline{e}) \tag{a2}$$

where on the right hand side there is just another notation for the matrix under discussion.

For two arbitrary vectors $\underline{a}, \underline{b} \in \mathcal{V}_n$ and for any index i

$$\begin{aligned}&\det(\underline{c}_1,\underline{c}_2,\ldots,\underline{c}_{i-1},\underline{a}+\underline{b},\underline{c}_{i+1}\ldots,\underline{c}_n)\\&\quad=\det(\underline{c}_1,\underline{c}_2,\ldots,\underline{a},\ldots,\underline{c}_n)\\&\qquad+\det(\underline{c}_1,\underline{c}_2\ldots\underline{b},\ldots\underline{c}_n)\end{aligned} \tag{a3}$$

and for any indexes $i, j,; i \neq j$

$$\begin{aligned}&\det(\underline{c}_1,\underline{c}_2,\ldots,\underline{c}_{i-1},p\underline{e},\underline{c}_{i+1},\ldots,\\&\quad\underline{c}_{j-1},p\underline{e},\underline{c}_{j+1},\ldots,\underline{c}_n)=0\end{aligned} \tag{a4}$$

The two last identities express obvious behavior of the determinant operation, they can be proved easily starting from any definition of the determinant.

Repeatedly applying (a3) to the right hand side of (a2) we get the sum of 2^n number of determinants, but on the basis of (a4) only $n+1$ out of them differ from zero:

$$\begin{aligned}&\det(C+pE)=p\cdot\{\det(\underline{e},\underline{c}_2,\ldots,\underline{c}_n)\\&\quad+\det(\underline{c}_1,\underline{e},\underline{c}_3,\ldots,\underline{c}_n)+\cdots+\det(\underline{c}_1,\underline{c}_2,\ldots,\underline{e})\}\\&\quad+\det(\underline{c}_1,\underline{c}_2,\ldots,\underline{c}_n)\end{aligned} \tag{a5}$$

This shows the validity of (a1).

Now let $C \in \mathcal{M}_{n,n}, B \in \mathcal{M}_{k,k}$ and $E_n \in \mathcal{M}_{n,n}, E_k \in \mathcal{M}_{k,k}$ be the matrices each element of which is unit and suppose cycl det $B \neq 0$.

Then on the basis of (a1) the equation

$$\lim_{p\to\infty}\frac{\det(C+p\cdot E_K)}{\det(B+p\cdot E_n)}=\frac{\mathrm{cycl\ det}\ C}{\mathrm{cycl\ det}\ B} \tag{a6}$$

is a simple limit process for the ratio of linear, related to p, functions.