SPE 37744

A Fast and Accurate Method for Parameter Estimation of Archie Saturation Equation

Noaman El-khatib, SPE, King Saud U.

Abstract

A modified regression analysis method is presented for the estimation of Archie equation parameters, a ,m and n. The method is a weighted least squares with weights chosen as to approximate the actual residuals of the independent variable which are nonlinear by the residuals of the logarithms of the dependent variable which are linear. The method thus combines the simplicity of linear regression on the logarithms with the physical meaningfulness at the nonlinear regression on the actual variables. The method is found to give almost identical results with the nonlinear regression method. Better results in terms of data fitting is obtained if the Archie parameter a is allowed to be estimated from regression analysis rather than fixing its value at 1. Other forms of regression analysis including sequential parameter estimation and straight line regression analysis are discussed.

Introduction

Porosity and fluid saturations are among the most important reservoir properties used in reserve estimates of oil and gas reservoirs. Because of the heterogeneity of most of the reservoirs, a continuous recording of these properties vs. depth is essential for accurate estimations. Since a complete coring and core analysis of the entire pay zone is impractical, well logging appears to be the most plausible mean to obtain such information. To achieve this task, different logs including resistivity, acoustic and radioactive logs are recorded in the well and well log interpretation is used to obtain the required properties.

Porosity and lithology of the formation is usually obtained from a combination of logs called the porosity logs. These include the neutron, sonic and density logs[1]. Fluid saturations are usually obtained from resistivity logs. Different resistivity logs with variable radii of investigation are recorded to obtains fluid saturations at different distances from the wellbore. A micro resistivity log is used to obtain the resistivity in the flushed zone behind the wellbore while a deep resistivity device (induction or laterolog) is used to obtain the resistivity in the virgin (uninvaded) zone.

Fluid saturations are estimated from resistivity measurements by the use of Archie equation[2]. This equation relates the resistivity of the formation to the porosity, water saturation and resistivity of the water saturating the formation

$$R = \frac{F R_w}{S_w^n} \quad \text{.................} \quad (1)$$

with

$$F = \frac{R_o}{R_w} = \frac{a}{\phi^m} \quad \text{..................} \quad (2)$$

Eq. [1] is usually written in the form

$$R = \frac{aR_w}{\phi^m S_w^n} \quad \text{....................} \quad (3)$$

Archie equation is used to determine the water saturation S_w provided the coefficients a, m and n are known and ϕ is obtained from independent porosity logs. Originally Archie used a value of a=1 in his equation but it was found that values other than one fits bettter the data for some formation. So ,Winsauer et al.[3] introduced thecoefficient a in Eq. (2) which may differ from unity. For the purpose of estimating S, Archie equation is written as

$$S = \left[\frac{a}{\phi^m}\frac{R_w}{R_t}\right]^{1/n} \quad \text{.....................} \quad (4)$$

Equation (4) can be used to estimate the saturation in the original zone or in the flushed zone. When used with resistivity devices with deep radius of influence, it provides a value of the original (connate) water saturation S. This is the saturation needed to estimate the original hydrocarbon in place. In this case the water resistivity used in Eq. (4) is the formation water resistivity R which may be obtained from the chemical analysis of formation water or from the SP log of the formation under consideration. On the other hand, when Eq. (4) is used with micro resistivity logs it gives the value of the aqueous phase saturation in the flushed zone which is composed mainly of mud filtrate. The resistivity of this zone is usually denoted by R_{xo} and Eq. (4) takes the form

$$S_{xo} = \left[\frac{a}{\phi^m} \frac{R_{mf}}{R_{xo}} \right]^{1/n} \qquad \text{.................. (5)}$$

The saturation in the flushed zone obtained from Eq. (5) can be used to estimate the residual oil saturation and the movable hydrocarbon saturation S_{hr} after a waterflooding process.

$$S_{or} = 1 - S_{xo} \qquad \text{................ (6)}$$

$$S = S_{xo} - S_{wi} \qquad \text{.................... (7)}$$

To apply Archie equation for estimation of fluid saturation, one must have values for the equation parameters a, m and n. A widely used value for the saturation exponent n is 2. Empirical values for a and m are also used in the literature depending on the type of formation. For compact formations the equation

$$F = 1/\phi^2 \qquad \text{.................. (8)}$$

is usually applied while for loose formations (sands), the so-called Humble formula

$$F = 0.62/\phi^{2.15} \qquad \text{..................... (9)}$$

is applicable. A modification of the Humble formula is usually used to eliminate the fraction in the cementation factor

$$F = 0.81/\phi^2 \qquad \text{................. (10)}$$

Perhaps the use of values of m=2, n=1 and a=1 was dictated by the simplicity of calculations before computers were used. When such values are used. Eq. (4) takes the form

$$S_w = \frac{1}{\phi} \sqrt{\frac{R_t}{R_w}} \qquad \text{.................. (11)}$$

Values of m and n, however, were found to vary significantly from the value of 2 depending on the type of the formation, depth, overburden pressure and formation temperature[4-7]. It is therefore desirable to have the value of the Archie equation parameters a, m and n that corresponds to the formation for which the saturation is to be estimated. This can be done by performing laboratory resistivity measurements on cores taken from the specific formation and estimating the values of Archie equation parameters. Measurements on a single core at different saturation levels provide a value of the saturation exponent n and the formation resistivity factor F. Values of F for cores of different porosities can be used to obtain the parameters a and m. Since values of n obtained from different cores will normally not be the same, these values must be averaged to obtain a representative value form for the entire. Since the averaging process is not unique, it is therefore preferable to use all the measurements on the different cores to determine simultaneously the values of a, m and n. This is usually done by performing regression analysis on the measured data.

The idea of parameter estimation by regression analysis is to find the values of model parameters from a set of observations. Most of parameter estimation methods are based on one form or the other of the method of least squares. This method estimates the values of the parameters which minimize an objective function representing the sum of squares of differences between observed and estimated values of the dependent variable at the observation points. These differences between observed and calculated values are known as the errors or the residuals. Modifications of the least squares method include the weighted least squares method where each observation point is assigned a nonnegative weight factor to be multiplied by the square of the residual before minimization.

The procedure of least squares estimation involves finding the derivatives of the objective function with respect to each of the model parameters and equating each of them to zero. The resulting set of equations which represent the conditions for the objective function to be minimum are known as the normal equations. Solution of the normal equations give the required estimation of the model parameters. Some forms of least squares method that do not involve estimation of derivatives can also used. Chen et al used the simplex method[8] and the Fuzzy regression analysis[9] to estimate the Archie equation parameters.

If the model is linear in the parameters to be estimated, the normal equations will be a system of linear algebraic equations. In this case the solution of the normal equations will be straight forward and does not need iteration or guessing an initial. If the model is nonlinear, the resulting normal equations will be nonlinear themselves and an iterative procedure must be used to solve for the model parameters. All iterative procedures require an initial guess of the model parameters. The number of iterations needed to reach a solution (convergence) will depend on the initial guess and its closeness to the actual solution. In many saturations convergence may not be achieved if the initial guess is too far from the true values of the parameters.

Many nonlinear problems of parameter estimation can be made linear by taking the logarithms of the dependent and/or independent variables as new variables. Although the solution of the normal equations becomes easier, the quantity to be minimized loses its physical significance.

It is the objective of this work to devise a method that has the simplicity of the linear regression yet retains the accuracy and the physically meaningfulness of the original nonlinear regression problem. This is achieved by introducing weights to the logarithmic residuals that make them approximate the actual residuals of the dependent variable.

Archie Parameter Estimation by Nonlinear Regression Analysis

If data obtained from core measurement are to be fitted to Eq. ( ), the objective function to be minimized is

$$G_1 = \sum\left[R_t - \frac{aR_w}{\phi^m S_w^n}\right]^2 \quad \dots\dots\dots\dots\dots\dots \quad (12)$$

Where the summation is taken over all the measurements and the subscripts are dropped for convenience. Equating derivatives of the objective functions w.r.t a, m and n to zero, the following normal equations are obtained

$$g_1 = \sum\left[R_t - \frac{aR_w}{\phi^m S_w^n}\right]\frac{R_w}{\phi^m S_w^n} = 0 \quad \dots\dots\dots\dots \quad (13)$$

$$g_2 = \sum\left[R_t - \frac{aR_w}{\phi^m S_w^n}\right]\frac{R_w}{\phi^m S_w^n}\ln\phi = 0 \quad \dots\dots\dots \quad (14)$$

$$g_3 = \sum\left[R_t - \frac{aR_w}{\phi^m S_w^n}\right]\frac{R_w}{\phi^m S_w^n}\ln S_w = 0 \quad \dots\dots \quad (15)$$

Equations (7) - (15) are three nonlinear equations in the three parameters a, m and n. An iterative solution must be used and it requires obtaining the derivatives of the three functions g_1 , g_2 and g_3 w.r.t. the three parameters a, m and n. The resulting matrix is called the Jacobian matrix. As stated before, convergence is not guaranteed and the estimation of the Jacobian matrix at each iteration is a time consuming task.

Archie Parameter Estimation by Linear Regression

The parameter estimation problem can be made linear by taking the logarithms of both sides of Eq. (4) which can be written as

ln(Rt/Rw) = ln a - m ln ϕ - n ln S .... (16)

In this case the objective function to be minimized is

$$G_2 = \sum\left[\ln\left(\frac{R_t}{R_w}\right) - \ln a + m\ \ln\phi + n\ \ln S_w\right]^2 \quad (17)$$

The normal equations are

$$\sum\left[\ln\left(\frac{R_t}{R_w}\right) - \ln a + m\ln\phi + n\ln S_w\right] = 0 \quad (18)$$

$$\sum\left[\ln\left(\frac{R_t}{R_w}\right) - \ln a + m\ln\phi + n\ln S_w\right]\ln\phi = 0 \quad (19)$$

$$\sum\left[\ln\left(\frac{R_t}{R_w}\right) - \ln a + m\ln\phi + n\ln S_w\right]\ln S_w = 0 \quad (20)$$

These three equations are linear in lna, m and ϕ. In matrix form, these equations can be written as

$$\underline{A}\,\underline{X} = \underline{b} \quad \dots\dots\dots\dots\dots\dots\dots\dots(21)$$

where

$$A = \begin{vmatrix} N & -\sum\ln\phi & -\sum\ln S_w \\ \sum\ln\phi & -\sum(\ln\phi)^2 & -\sum\ln\phi\ln S_w \\ \sum\ln S_w & -\sum\ln S_w\ln\phi & -\sum(\ln S_w)^2 \end{vmatrix} \quad (22)$$

$$X = \begin{vmatrix} \ln a \\ m \\ n \end{vmatrix}, \quad b = \begin{vmatrix} \sum\ln(R_t / R_W) \\ \sum\ln\phi\ln(R_t / R_W) \\ \sum\ln S_w Ln(R_t / R_W) \end{vmatrix} \quad \dots\dots \quad (23)$$

The solution for lna (and hence a), m and n can be easily obtained using Cramer's rule or the Gaus elimination method.

It was argued[10] that the quantity minimized in Eq.(17) does not have a physical meaning. Instead of minimizing the sum of the squares of the differences between the observed and calculated values of the dependent variable, one is working with the logarithms of the dependent variable which has no physical meaning. Furthermore, the errors in the dependent variable are assumed to be normally distributed about the expected value and has a zero mean. This is not true for y_1 = log y which implies that this transformation introduces a bias in the estimation of parameters. It is recognized that in the method of least squares, a deviation δ in the dependent variable is equally acceptable regardless of the value of y. A deviation of ± 0.1 in y would mean the same absolute error in the value of y independent of the specific value of y. On the other hand a deviation of ± 0.1 in logy would mean different absolute errors in the value of the dependent variable y depending on the value of y itself. For example for values of logy of 1, 2 and 3, a deviation of ± 0.1 means an absolute error of 2.589, 25.89 and 258.9 in the values of y respectively (y = 10, 100, 1000).

It is therefore desirable if the residuals of the dependent variable R_t can be expressed in terms of the residuals of the logarithms of the dependent variable. This would combine the accuracy and meaningfulness of the nonlinear regression with the simplicity and convenience of linear regression.

The Proposed Method

It is required to express the residuals of the dependent variable y in terms of the residuals of ln y. Let

$$y_{cal} = y_{obs} + \delta \quad \text{............} \quad (24)$$

$$y_{cal} / y_{ob} = 1 + \delta / y_{obs} \quad \text{..............} \quad (25)$$

Using Taylor's expansion we get,

$$\ln y_{cal} = \ln y_{obs} + \left(\frac{1}{y_{ob}}\right)\delta + \left(-\frac{1}{y_{obs}^2}\right)\delta^2 \quad \text{......} \quad (26)$$

Neglecting terms of δ^2 and higher we get

$$\left(\ln y_{cal} - \ln y_{ob}\right) \cong \frac{y_{cal} - y_{ob}}{y_{obs}} \quad \text{..........} \quad (27)$$

Or

$$y_{ob} - y_{cal} \cong y_{obs}\left(\ln y_{ob} - \ln y_{cal}\right) \quad \text{..........} \quad (28)$$

It is clear that the SSE for the dependent variable, Eq. (12) can be approximated by

$$SSE = \sum\left(y_{ob} - y_{cal}\right)^2 \cong \sum y_{ob}^2 (\ln y_{ob} - \ln y_{cal})^2 \quad \text{..................} \quad (29)$$

The foregoing mathematical derivation can simply be deduced from the fact that

$$\Delta \ln y \cong \left(\frac{d \ln y}{dy}\right)\Delta y = \left(\frac{1}{y}\right)\Delta y \quad \text{...........} \quad (30)$$

Equation (29) implies that a weight factor y^2 is assigned to each observation point when we minimize the sum of squares of residuals of the logarithms of the dependent variable.

It is to be noted that if the values of y^2 are large then assigning weight factors of y^2 may cause overflow in the calculations. The weight factors can be scaled by dividing by a number say y^2_{max} or y^2_{av}.

Solution Procedure

As seen from Eq. (29) each residual of the logarithm of the of the dependent variable should by weighed by a factor equal to the value of that variable squared. If the dependent variable to be minimized is the resistivity R_t then the objective function to be minimized is

$$G = \sum R_t^2 \left[\ln R_t - \ln\left(\frac{aR_w}{\phi^m S_w^n}\right)\right]^2 \quad \text{..........} \quad (31)$$

which may be written as

$$G = \sum R_t^2 \left[\ln\left(\frac{R_t}{R_w}\right) - \ln a + m \ln\phi + n \ln S_w\right]^2 \quad \text{........} \quad (31)$$

The normal equations can be written in a matrix form as in Eq.(21) with

$$A = \begin{vmatrix} \sum R_t^2 & -\sum R_t^2 \ln\phi & -\sum R_t^2 \ln S_w \\ \sum R_t^2 \ln\phi & -\sum (R_t \ln\phi)^2 & -\sum R_t^2 \ln S_w \ln\phi \\ \sum R_t^2 \ln S_w & -\sum R_t^2 \ln\phi \ln S_w & -\sum (R_t \ln S_w)^2 \end{vmatrix} \quad \text{................} \quad (32)$$

and

$$b = \begin{vmatrix} \sum R_t^2 \ln R_t \\ \sum R_t^2 \ln R_t \ln\phi \\ \sum R_t^2 \ln R_t \ln S_w \end{vmatrix} \quad \text{..............} \quad (33)$$

which can easily be solved for ln a, m and n. If one still desires to work with the dependent variable y, this method provides an initial guess for the solution procedure of the nonlinear regression analysis outlined earlier.

Variations in Parameter Estimation by Regression Analysis.

In parameter estimation by regression analysis, one tries to fit experimental data to a proposed model (equation). The equation may have theoretical basis or may be purely empirical. Since the model involves a number of variables, the equation can be written in different forms. The Archie equation, for example can be expressed in the following different forms.

$$R_t = \frac{aR_w}{\phi^m S_w^n} \quad \text{.................} \quad (34)$$

$$S_w = \sqrt[n]{\frac{aR_w}{\phi^m R_t}} \quad \text{.................} \quad (35)$$

$$\phi = \sqrt[m]{\frac{aR_w}{R_t S_w^n}} \quad \text{.................} \quad (36)$$

$$S_w \phi = \left[\frac{aR_w}{R_t}\right]^{1/n} \phi^{\frac{1-m}{n}} \quad \text{.................} \quad (37)$$

$$S_w \phi = \left[\frac{aR_w}{R_t}\right]^{1/m} S_w^{\frac{1-n}{m}} \quad \text{.................} \quad (38)$$

Although all these forms are equivalent, the values of model parameters obtained by fitting the data to each of these equations will not be the same. Maute et al.[10] argued that

Archie equation should be written in the form of Eq. (35) because it will be used to estimate the water saturation.They also proposed regression on the quantity ϕS_w if the objective is to estimate the hydrocarbon pore volume. They did not however point out the form of Archie equation to be used in this case. Furthermore, the hydrocarbon pore volume is related to $\phi(1-Sw)$ which is a more complex term than ϕS_w. The arguments of Maute et al however were based on the idea that one should make regression analysis on a quantity that has physical meaning rather than on the logarithm of a quantity. The objective of regression analysis is to try to find the best representation of the data relating a dependent variable to a set of independent variables. It is usually assumed that independent variables can be accurately determined while measurements of the dependent variable are subjected to statistical fluctuations. In most physical systems, independent variable are time and location while the dependent variable is a physically measurable quantity like temperature, pressure, velocity, ... etc. In our problem, a given core is saturated with water having a certain salinity (and hence resistivity) and the resistivity of the core is measured. It is logical to assume that the porosity is an independent property of the core together with the amount of water in the core S_w and the resistivity of the saturating fluid and that the resistivity of the core would be the dependent variable that is affected by the independent variables. Fitting the data to Archie equation in its original form, Eq. (3) would be logical. It seems that the argument of Maute et al of using Sw instead of log R_t is well founded but the same logic may not apply for using Sw instead of Rt. It was therefore decided to make regression analysis on R_t and S_w and compare the results.

Another variation of linear regression analysis was discussed by Tehrani[12] in his analysis of the material balance equation. This variation applies in linear regression and involves dividing both sides of the equation by one of the independent variables thus reducing the number of regression variables by one. This procedure is so tempting because one can sometimes reduce the problem to a straight line and thus making it amenable to graphical interpretation. In the Archie equation, if the parameter a is fixed to unity as was suggested by Maute et al or to a fixed specific value (0.62 or 0.81 for example), the equation in the logarithmic form can be written as

$$\ln (R_t/aR_w) = - m \ln\phi - n \ln S_w \quad .... \quad (39)$$

If Eq. (39) is divided by log ϕ or by log S_w we get the following two forms

$$\frac{\ln\left(\frac{R_t}{aR_w}\right)}{\ln\phi} = -m - n\frac{\ln(S_w)}{\ln\phi} \quad .... \quad (40)$$

and

$$\frac{\ln\frac{R_t}{aR_w}}{\ln S_w} = -m\frac{\ln\phi}{\ln S_w} - n \quad .... \quad (41)$$

Each of these equations can be plotted as a straight line. Tehrani has shown that such forms have low resolving power and may produce erroneous estimates.

Maute et al.[10] also argued that the parameter a in Archie equation is a weak-fitting parameter with no physical significance and they recommended it to be fixed to unity. Theoretically, a should be equal to unity since for pure water ($\phi=1$, $S_w=1$) the resistivity R_t should be equal to R_w and hence a equals unity. Since most of the porosities in actual reservoirs are far from unity and the parameters should be used to fit the field data, a value of a different than unity is usually obtained. As mentioned by Day[11], the parameter a accounts for many unseen factors and that its value should be determined from the data. Although the differences were insignificant, results from the CAPE model of Maute et al showed that results with the value of a different from unity showed lower standard deviation than for the cases with a set equal to unity. The reason of the insignificant differences may be due to the large number of data points they used in their analysis which must have wide scattering. As stated by Day, for data with wide scattering, fixing a to unity or determining if from regression analysis will not affect the prediction results of the model.

To investgate the effect of the different variations in the objective function a simulation example will be used.

Sequential Estimation of Archie Parameters

It is possible to make a sequential estimation of parameters in Archie equation by first estimating n and F for each core using the relation

$$(R_t/R_w) = F / S_w^n \quad .... \quad (42)$$

and then correlating the F obtained from the different cores with porosity to find the parameters a and m using the relation

$$F = a/\phi^m \quad .... \quad (43)$$

and then correlate the values of F obtained for t cores with different porosities to find the parameters a and m using Eq. (2).

It is not true what Maute et al mentioned that only points with 100% water saturation can be used to estimate a and m. This followed from the belief that F must be determined from the relation

$$R_o = F R_w \quad .... \quad (44)$$

where R_o is the resistivity at 100% water saturation. The value of F however can be determined from measurements on the same core at different saturations according to Eq. (1). In fact, F can be determined from measurements that may not contain any point with Sw = 1.

After the value of F for each core is determined we may use the relation

$$I_R = \frac{R_t}{R_o} = \frac{R_t}{FR_w} = S_w^{-n} \qquad \text{.....................} \qquad (45)$$

to analyze the data for all cores to obtain average value of n.

Results and Discussion

A simulated example with twelve cores of porosities ranging from 0.05 to 0.6 is used in this study . The water saturation in each core is assumed to vary between 0.1 and 1 and the water resistivity used is 0.05 Ωm. The Archie equation is assumed to have the form

$$R_t = \frac{0.62 R_w}{\phi^{2.15} S_w^2} \qquad \text{.................} \qquad (46)$$

The resistivity R_t is calculated using the above relation. A random error was then introduced in R_r. The error is assumed to be normal with a zero mean and a standard deviation σ of 1 Ωm. Table 1 shows the calculated values of the core resistivities R_t .The exact values of R_t as well as those with generated random errors are used in the different parameter estimation methods. The following combinations were studied:

- Exact data and data with random errors.
- Regression on R_t and regression on S_w.
- Parameter a estimated from regression analysis and a=1 .
- Linear regression analysis, nonlinear regression analysis and weighted least sqares metod .

Table 2 shows the results for the different cases analysed.

First, the exact data without error was used for parameter estimation by linear regression (logarithmic), nonlinear regression and the proposed method (weighted least squares). In this case, the three methods produced the exact values of a, m and n. Since for data with no errors, the residuals are not exactly zeros due to roundoff errors, the weighted least squares methods which multiplies each residual by the value of R_t seems to have a relatively higher value for the standard deviation of R_t (6.8 x 10^{-3} as compared to 3.6 x 10^{-3} for the linear method and 8.1 x 10^{-3} for the nonlinear regression method)..

When the value of the parameter a is set equal to one, the linear regression method estimated values for m and n of 1.936 and 1.862 respectively with a standard deviation for R_r of 71.42 while the proposed method and the nonlinear regression methods estimated almost identical values for m (2.0377 and 2.0395) and n (1.935 and 1.9322) and standard deviation (5.555 and 5.549). Figure 1 shows the calculated values of R_t as compared to the input (exact) values for both cases of fixed a (a=1) and for a obtained by regression analysis. It is seen that for the case when the parameter a is estimated from regression analysis, the data almost lies on a straight line with unit slope while for the case of fixed a ,the data is shifted upward with the estimated values of R_t being larger than the input values. It can therefore be concluded that more accurate results will be obtained if the parameter a is allowed to be estimated from regression analysis.

When the saturation equation, Eq. (4) is used with the exact data, exact values of the parameters are also obtained by the three methods (linear, weighted, nonlinear) with the standard deviations for the saturation of 3 x 10^{-6}, 2.4 x 10^{-6} and 4.3 x 10^{-8} for the three methods respectively, and the standard deviations for the resistivity of 3.75 x 10^{-3}, 5.47 x 10^{-3} and 2.4 x 10^{-5} for three methods respectively. When the value of a is fixed to unity, the standard deviations for the three methods increase to a value of about 7.5 x 10^{-2} for the saturation and to larger values (66-74) for the resistivity. Fig. 2 shows the fitting of the saturation for the case where a is determined from regression analysis and Fig. 3 for the case of a = 1. The large departure of calculated values from the input values for the latter case is obvious. The same behavior was also observed when the data with generated random errors. .

It is also observed that the weighted and nonlinear regression methods give higher standard deviations for the saturation than the linear (logarithm) method when the resistivity equation is used for the regression analysis.In the linear regression analysis method the errors (residuals) of the logarithms are of the same order for the whole range of R_t as seen in Fig. 5. The actual errors in R_t are very small for low values of R_t and relatively large for high values of R_t as seen Fig. 4. This results in a high value for the standard deviation (root mean square error) of R_t. Since high values of R_t corresponds to low values of S_w and vise versa, the errors in S_w will be small.

On the other hand, in the nonlinear regression method and similarly the weighted least squares method, the errors on a normal scale are of the same value as can be seen from Fig. 7, while the errors are relatively high for the low values of R_t and relatively low for high values of R_t as can be seen from the log scale of Fig. 6. This results in a low value of the standard deviation of R_t. Because of the relatively large deviation in the low values of R_t which corresponds to high values of S_w, there will be large residuals in S_w which results in high standard deviation of the water saturation.

From Eq. (3) we can write

$$\frac{dR_t}{R_t} = 2\frac{dS_w}{S_w} \qquad \text{...............} \qquad (47)$$

So the relative error in the saturation when the saturation is high and R_t is low will be high in case of nonlinear regression and low in case of normal regression analysis.

When the saturation equation, Eq. (4) is used in regression analysis, the standard deviation for the saturation decreases from the linear to the weighted to the nonlinear regression method. The change in the standard deviation of the saturation between the three methods is not large. The reason for this is that, first the values of S_w are small ($0 < S_w < 1$) and these values usually lie within one logarithmic cycle (0.1 - 1.0) and so the residuals do not have wide variations. Also the standard deviations of R_t from the saturation equation are close but are large compared to the values obtained from the resistivity equation, Eq.(3) which minimizes the residuals of R_t.

The sequential procedure for parameter estimation was used for the data with random errors. First the data for each core was fitted according to the equation

$$\frac{R_t}{R_w}=\frac{F}{S_w^n} \qquad \text{......................} \qquad (48)$$

and the values of F and n were determined for each core. The resuls of the analysis are shown in Fig. 8. The arithmetic average of n for the 12 cores was found to be 1.991 as compared to the exact value of 2.0. The values of F obtained are correlated with the porosity ϕ according to Eq. (2).The reults of the analysis are given in Table 3. The cementation factor m was found to be 2.1138 as compared to the actual value of 2.15 and to the value of 2.154 obtained by simultaneous regression analysis. The value of a is determined as 0.689 as compared to the exact value of 0.62 and to the value from regression analysis of 0.611. When the value of a was fixed to unity the value of m obtained was 1.988. Figure 9 shows the results plotted on a log-log scale. The calculated values of F were then used to correlate the resistivity index with the water saturation according to Eq.(45).The value of n obtained is 2.00072 as compared to the actual value of 2 and to the value of 2.000827 from simultaneous regression analysis. The fitted data of I_R vs S_w is shown in Fig. 10.

Eqation (40) has also been tested. the value of a is assumed to be 0.62 and 1, then the terms [$\ln(R_t/a\,R_w) / \ln\phi$] and [$\ln S_w / \ln\phi$] are calculated and fitted to a straight line according to Eq.(40). For the case of a = .62, the values of m and n obtained were 2.1469 and 2.0034 respectively. When the value of a was fixed at unity, the values obtained for the parameter m and n were 1.7715 and 1.88 respectively. The plot of the data is shown in Fig. 11 for a=.62 and in Fig.12 for a =1. It is also clear that large deviations occur when the parameter a is fixed at a value of 1.

Conclusions

From the results of this study the following conclusions can be made.

1- The weighted least squares method presented in this study with the weight factor equals to the square of the dependent variable gives close results to the nonlinear regression analysis method.
2- When the data is free of measurement errors, any of the three regression methods can be used. In this case the simple linear regression method based on logarithmic transformation is recommended.
3- When the resistivity residuals are minimized, large variations in S_w at high values of R_t will result for the nonlinear least squares method and the proposed weighted least squares method. On the other hand when the saturation residuals are minimized no such large changes occur..
4- Sequential estimation of the Archie parameters may be used with confidence. Other forms of analysis that amend themselves to graphical interpretation may be used if the value of the parameter a is known.
5- More accurate results are obtained if the parameter a is allowed to be estimated by regression analysis rather than fixing it at the a value of one.

Nomenclature

a	= Constant in Archie equation
F	= formation resistivity factor
I_R	= resistivity index
m	= cementation exponent
b	= saturation exponent
R	= resistivity, Ωm
R_{mf}	= resistivity of mud filtrate, Ωm
R_o	= resistivity of rock with 100% water saturation, Ωm
R_t	= true resistivity of formation, Ωm
R_w	= resistivity of formation water, Ωm
R_{xo}	= resistivity of flushed zone, Ωm
S_{hr}	= movable hydrocarbon saturation
S_o	= oil saturation
S_{or}	= residual oil saturation
S_w	= water saturation
S_{wi}	= irreducible water saturation
S_{xo}	= water saturation in the flushed zone
ϕ	= porosity

Subscript

f	= filtrate
h	= hydrocarbon
i	= initial
m	= mud
o	= oil
t	= true
w	= water

Acknowledgment

The author wishes to thank the Petroleum Engineering Department of King Saud University for the support provided during this study.

References

1. *Schlumberger Log Interpretation - Principles/Applications*, Schlumberger Educational Services, Houston, Texas, 3rd printing, May 1991.

2. Archie, G.E., "Electrical Resistivity Log as and Aid in Determining Some Reservoir Characteristics," *Trans., AIME* (1942) 146, 54-62.
3. Winsauer, W.O. et al., "Resistivity of Brine-saturated Sands in Relation to Pore Geometry," *AAPG Bulletin* (1952) 36, No. 2, 253-77.
4. Dokla, Mahmoud El-Awady, "Effect of Temperature on Formation Resistivity of Some Saudi Reservoir Rocks," paper SPE 9617 presented at the 1981 SPE Middle East Technical Conference, Bahrain, 9-12 March 1981, 387.
5. Sanyal, S.K., Marsden, S.S., Jr., and Ramey, H.J., Jr., "The Effect of Temperature on Electrical Resistivity of Porous Media," paper I presented at SPWLA 13th Annual Logging Symp., Tulsa, Okla., May 7-10, 1972.
6. Soendenna, E., Bratteli, F., Kolltveit, K. and Normann, H.P., "A Comparison Between Capillary Pressure Data and Saturation Exponent Obtained at Ambient Conditions and at Reservoir Conditions," paper SPE 19592 presented at the 64th Annual Tech. Conf. and Exhib. Of SPE, San Aantonia, TX., Oct. 8-11, 1989.
7. Glanville, C.R., "Laboratory Study Indicates Significant Effect of Pressure on Resistivity of Reservoir Rock," JPT (April 1959.
8. Chen, D.S., Fang, J.H. Kortright, M.E. and Chen, H.C., "Novel Approaches to the Determination of Archie Parameters I:Simplex Method," *SPE Advanced Technology Series*, Vol. 3, No. 1, 39.
9. Chen, D.S., Fang, J.H., Kortright, M.E., and Chen,H.C., "Novel Approaches to the determination of Archie Parameters II: Fuzzy Regression Analysis," *SPE Advanced Technology Series*, Vol. 3, NO. 1, 44.
10. Maute, R.E., Lyle, W.D. and Sprunt, E.S., "Improved Data Analysis Method Determines Archie Parameters from Core Data," *JPT*, (Jan. 1992), 103-7.
11. Day, Peter I., "Discussion of Improved Data-Analysis Determines Archie Parameters from Core Data, *JPT* (July 1992) 846.
12. Tehrani, D.H., "An Analysis of Volumetric Balance Equation for Calculation of Oil in Place and Water Influx," *JPT* (Sept. 1985) 1664-70.

TABLE 1- GENERATED VALUES OF CORE RESISTIVITIY

No	ϕ	Water Saturation S_W									
		.1	.2	.3	.4	.5	.6	.7	.8	.9	1
1	.05	1943.463	485.8658	215.9403	121.4664	77.73853	53.98508	39.66251	30.36661	23.99337	19.43463
		1945.747	486.4067	215.7275	122.4105	76.07392	53.62247	39.30131	30.608	23.53677	18.2138
2	.1	437.8867	109.4717	48.65408	27.36792	17.51547	12.16352	8.936464	6.84198	5.406009	4.378867
		436.8488	109.2901	49.56764	27.58348	18.73243	11.08129	8.855636	7.526178	5.327984	3.94098
3	.15	183.1325	45.78313	20.34806	11.44578	7.325302	5.087015	3.737399	2.861446	2.260896	1.831326
		181.9423	46.52401	20.41438	12.59036	6.592772	5.595716	3.63763	2.575301	2.486985	1.892562
4	.2	98.66139	24.66535	10.96238	6.166337	3.946456	2.740594	2.013498	1.541584	1.218042	.986614
		96.72214	25.7275	10.72533	6.782971	4.341102	3.014654	1.812148	1.695743	1.096238	.8879526
5	.25	61.06477	15.26619	6.784974	3.816548	2.442591	1.696244	1.24622	.954137	.7538861	.6106477
		62.50983	14.42378	6.106477	3.434893	2.198332	1.711185	1.121598	.8587233	.6784975	.6717125
6	.3	41.26207	10.31552	4.584674	2.578879	1.650483	1.146169	.8420831	.6447198	.5094083	.4126207
		39.49026	10.4748	4.151421	2.836767	1.485435	1.260785	.7578748	.5802479	.4584675	.3892662
7	.35	29.62207	7.405519	3.291341	1.85138	1.184883	.8228353	.6045321	.4628449	.3657047	.2962207
		30.48672	6.664967	3.545874	2.036518	1.066395	.9051189	.5440789	.4165604	.3291342	.2665987
8	.4	22.22966	5.557414	2.469962	1.389354	.8891863	.6174904	.4536665	.3473384	.2744402	.2222966
		22.67662	6.113155	2.222965	1.528289	.9781049	.5557414	.4082998	.3820722	.2469962	.2445262
9	.45	17.25658	4.314146	1.917398	1.078537	.6902634	.4793496	.3521752	.2696341	.2130443	.1725658
		17.37961	4.745561	2.109138	1.122153	.621237	.5272845	.3169577	.2965975	.2343487	.1612391
10	.5	13.75866	3.439666	1.52874	.8599164	.5503465	.382185	.280789	.2149791	.16986	.1375866
		13.85836	3.251882	1.681614	.945908	.6053811	.4204035	.308868	.236477	.186846	.1513453
11	.55	11.20939	2.802347	1.245488	.7005868	.4483756	.3113719	.228763	.1751467	.1383875	.1120939
		10.08845	2.522112	1.364263	.6305281	.4932131	.2802347	.2058867	.157632	.1245488	.1054287
12	.6	9.296865	2.324216	1.032985	.581054	.3718746	.2582462	.1897319	.1452635	.1147761	..099686
		9.748693	2.091794	1.136283	.6391594	.409062	.2324216	.1707588	.1597899	.1262537	.1022655

TABLE 2 - RESULTS OF NONLINEAR REGRESSUON ANALYSIS

L = Linear Regr.		Resistivity Equation						Saturation Equation					
W = Weighted R.		Exact Data			Data with Error			Exact Data			Data with Error		
NL = Nonlinear R		L	W	NL	L	W	NL	L	W	NL	L	W	NL
	a	0.620	0.620	0.620	0.618	0.611	0.611	0.620	0.62	0.62	0.615	0.612	0.607
	m	2.149	2.149	2.15	2.143	2.155	2.155	2.149	2.149	2.15	2.143	2.137	2.140
	n	1.999	1.999	2.00	2.009	2.001	2.000	1.999	2.000	2.0	2.016	2.060	2.058
	$\sigma(R_t)$	3.6E-3	6.8E-3	8.1E-6	1.132	0.504	0.504	3.7E-3	5.5E-3	2.4E-5	1.560	16.54	15.81
	$\sigma(S_w)$	2.9E-6	1.1E-4	7.5E-8	2.7E-2	2.81E-2	2.81E-2	3E-6	2.4E-6	4.3E-8	2.7E-2	2.71E-2	2.7E-2
a = 1	m	1.936	2.038	2.039	1.928	2.039	2.041	1.920	1.880	1.856	1.909	1.861	1.832
	n	1.862	1.935	1.932	1.870	1.934	1.931	1.896	1.954	1.957	1.911	2.011	2.009
	$\sigma(R_t)$	71.42	5.554	5.549	72.31	5.743	5.737	67.69	66.3	74.34	67.93	57.78	69.21
	$\sigma(S_w)$	8.1E-2	0.114	0.115	8.7E-2	0.124	0.125	7.8E-2	7.4E-2	7.4E-2	8.4E-2	8E-2	7.9E-2

TABLE 3 - RESULTS OF SEQUENTIAL REGRES. ANALYSIS					
Core No.	**Porosity**	**F**	**n**	F_{cal}	F_{cal}
1	.05	387.8394	2.001482	387.8023	386.0085
2	.1	89.03197	1.991764	89.59229	97.30042
3	.15	38.72439	1.973027	38.02206	43.45353
4	.2	22.28286	1.938824	20.69812	24.52633
5	.25	9.896675	2.101324	12.91443	15.73854
6	.3	8.809331	1.952897	8.784073	10.95325
7	.35	5.261445	2.063242	6.341308	8.062042
8	.4	5.027718	1.955807	4.781796	6.182304
9	.45	4.218547	1.916433	3.727873	4.891627
10	.5	2.796255	1.995463	2.983565	3.967182
11	.55	2.349175	1.933552	2.439139	3.282379
12	.6	1.667724	2.066919	2.029346	2.760964
Results of regression analysis			av. n = 1.991	a = .689 m= 2.114	a = 1 m=1.988

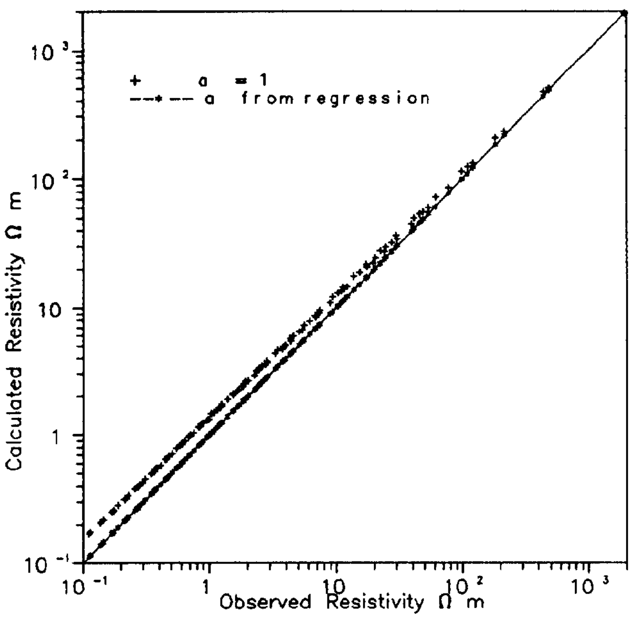

Fig. 1-Results of Regression Analysis , Exact Data

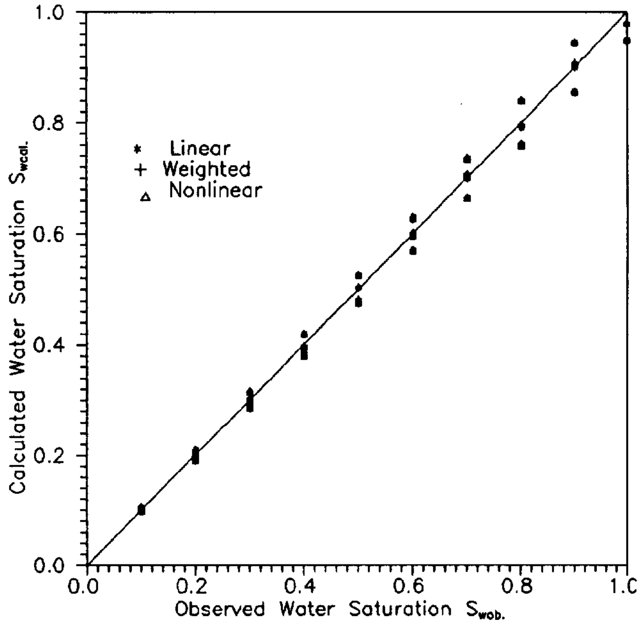

Fig. 2-Results of Regression Analysis for S_w

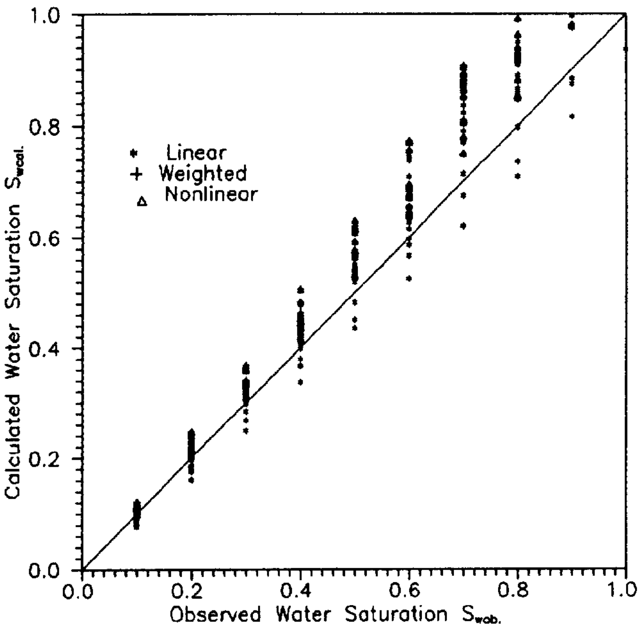

Fig. 3-Results of Regression Analysis for S_w (a=1)

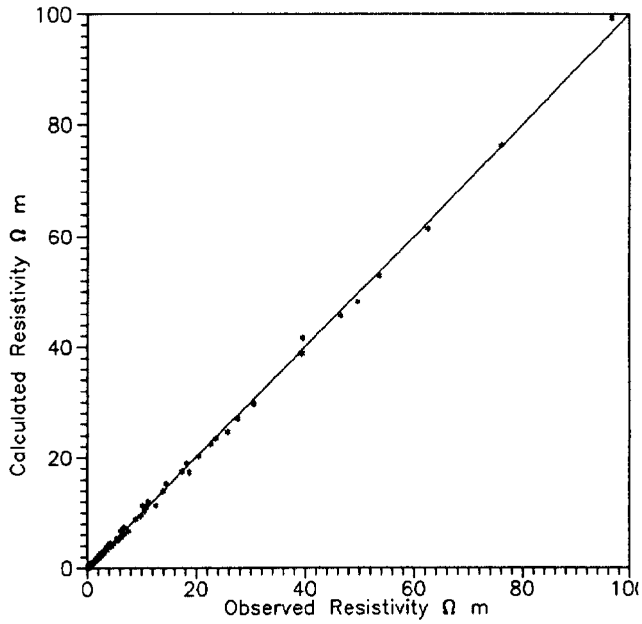

Fig. 4-Results of Linear Regression Analysis, Normal Scal

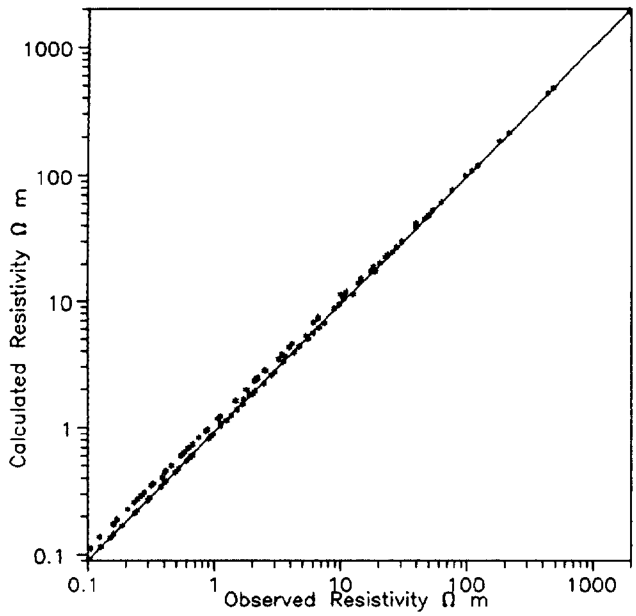

Fig. 5–Results of Linear Regression Analysis , Log Scale

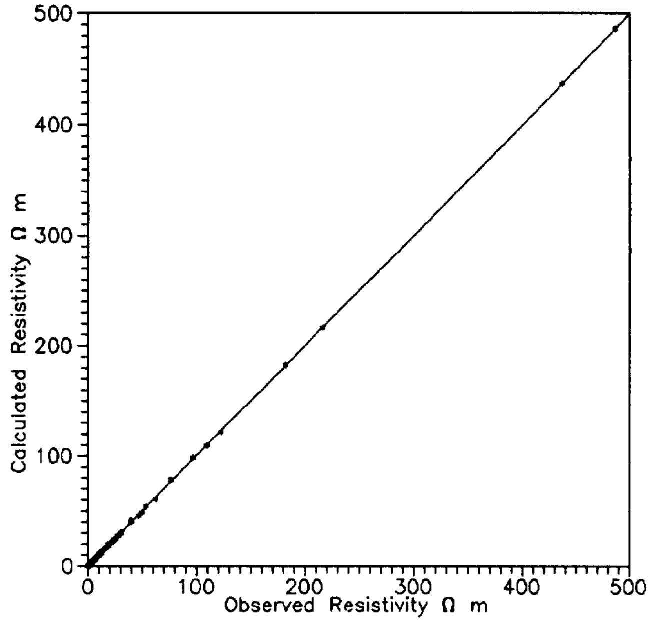

Fig. 6–Results of Weighted Regression Analysis, Normal Sca

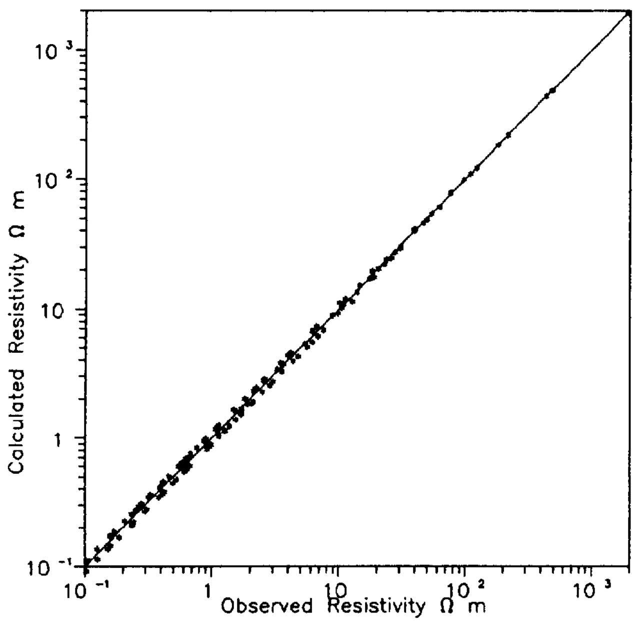

Fig. 7–Results of Weighted Regression Analysis, Log Scale

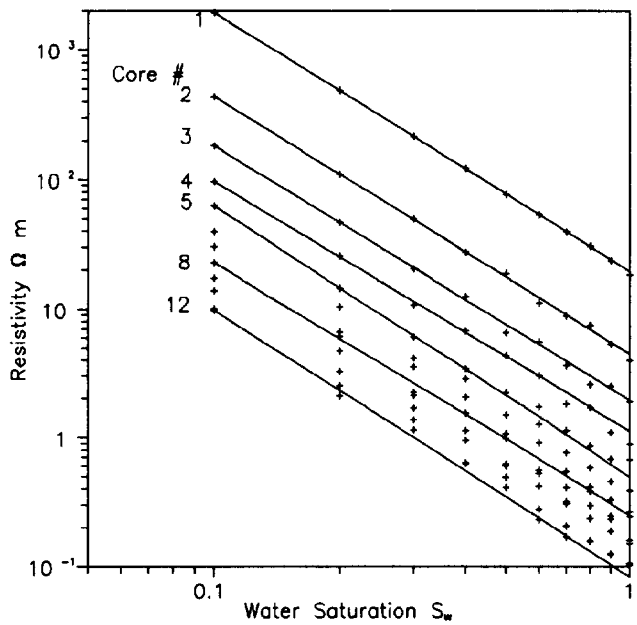

Fig. 8–Results of Regression Analysis for F and n

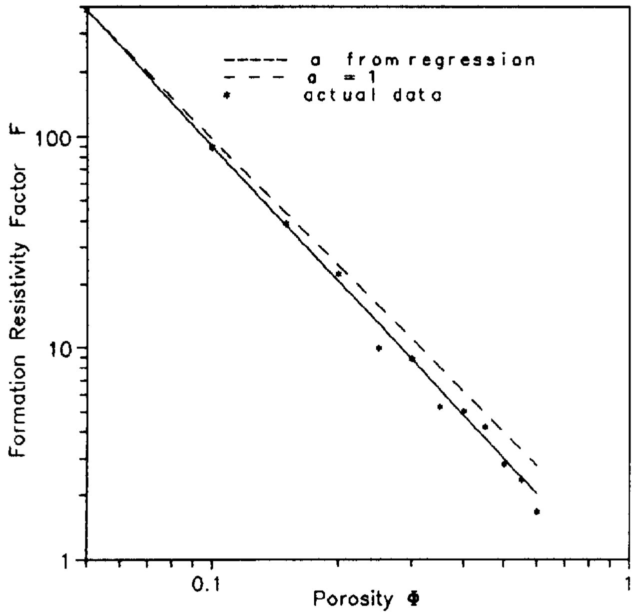

Fig. 9–Results of Regression Analysis for F and n

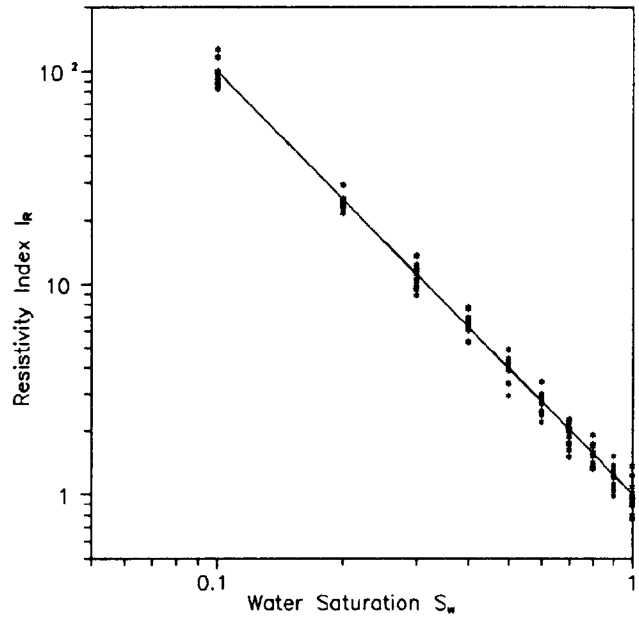

Fig. 10–Results of Sequential Regression Analysis for n

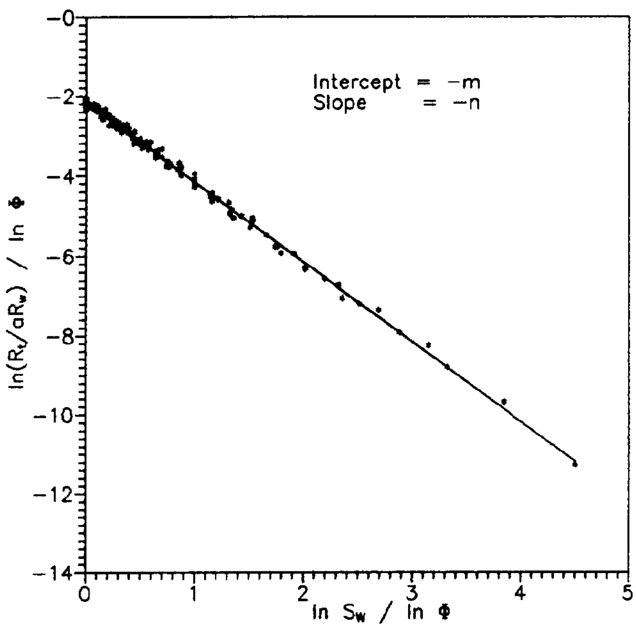

Fig. 11–Straight Line Regression Analysis , a = .62

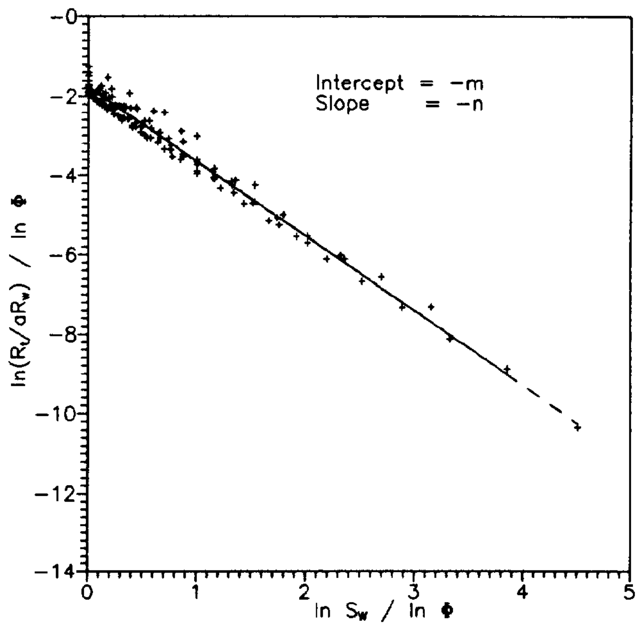

Fig. 12–Straight Line Regression Analysis , a = 1