A PRACTICAL APPROACH TO STATISTICAL LOG ANALYSIS

by

WILLIAM K. MITCHELL and RICHARD J. NELSON
Sun Exploration & Production Company
Dallas, TX

ABSTRACT

The advantages of a statistical approach to log analysis have been extensively discussed. Statistical techniques use "inverse" methods to calculate formation parameters. These methods minimize the difference between actual tool responses and reconstructed theoretical tool responses. The use of statistical techniques has been limited, however, by the complexity of the mathematics and lengthy computer time required to minimize the non-linear equations.

When computing lithology and porosity the tool response equations are usually linear or can be closely approximated by linear equations. The "inverse" problem can be solved directly without extensive iteration by using linear response equations. The approach described computes standard deviations which are used to normalize the response equations. Additional linear equations are added to the system of equations to insure linear independence as well as deal with necessary constraints. The system is then solved by a Modified Gram-Schmidt algorithm. A Maximum Liklihood Estimate (M.L.E.) of the standard deviation of the model can be calculated and used to evaluate the quality of the solution. This method not only solves the "overdetermined" case, but provides considerable help in solving the more common "underdetermined" case in which there is not enough information to solve for all the minerals present. The computation is faster and more flexible than conventional non-linear methods.

Previous papers have not supplied the details of the algorithms used to solve their statistical models. The algorithm needed for the solution of the linear model is not overly complex. One of the objectives of this paper is to provide the algorithm in sufficient detail to enable its incorporation into existing log analysis programs.

This approach has become the standard method for computing lithology and porosity by Sun Exploration and Production Company world-wide. Extensive use has verified the effectiveness and flexibility of this approach.

INTRODUCTION

Statistical log analysis techniques calculate formation parameters by minimizing the difference between normalized tool responses and reconstructed theoretical tool responses. The objective equation which must be minimized to obtain the solution is:

$$\Delta = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\frac{(m_i - f_i)^2}{\tau_i^2 + \delta_i^2}} \qquad (1)$$

Where:
- n = the number of tool response equations
- m_i = the actual tool measurement of tool i
- f_i = the theoretical tool response of tool i (calculated from theoretical tool response equation)
- τ_i^2 = the variance of the error of the tool measurement of tool i
- δ_i^2 = the variance of the error of the theoretical tool response of tool i

The minimized solution of this equation is the set of petrophysical parameters that best fit all the theoretical response equations. The advantages of this approach include:

- the use of all log and other data according to the accuracy of the data, resulting in optimized answers,
- the use of data from other sources, e.g., geological, core, experimental logging tools,
- the ability to solve for very complex lithologies,
- the output of "quality" indicators which indicate how well the model fits the data,
- the flexibility to define unique models without additional software.

These advantages have been discussed in detail.[1,2,3] Despite general agreement that statistical techniques produce more accurate answers, these techniques have been slow to replace conventional approaches. The mathematical complexity of the minimization procedure and associated programming have slowed the adoption of statistical techniques. Use of these methods, even when available, has also been limited by the amount of computer time required. Statistical techniques are also difficult to use because of the subtle interrelationships of formation parameters (e.g., cementation factor affecting porosity, etc.) which occur when solving for all parameters simultaneously.

In theory the simultaneous solution for all parameters leads to the best answers. In reality, it often is more practical to solve for lithology and porosity independently of fluid saturations. Solving for fluid saturations requires the use of non-linear equations, but lithology and porosity can be computed with linear equations. The interrelationship of formation parameters is simplified by dealing with only linear equations, and the computation is much faster. The approach discussed in this paper solves for lithology and porosity using only linear response equations.

When solving for lithology and porosity most responses are linear. Simple algorithms may be used, if necessary, to transform non-linear data to that which is linear or nearly linear. The intrinsic error resulting from the use of linear approximations to classical tool response functions is usually small. The error associated with the mineral endpoint selection is, in most cases, much larger than that resulting from the linear approximation.[4] Linear approximations of tool responses to hydrocarbon are also adequate when uncertainties associated with invasion, hydrocarbon density, reservoir pressure, etc., are considered.

Methods used to solve these systems of equations, or "solvers" as they are sometimes called are not discussed in the literature, and are often considered confidential. The following describes in detail a method which uses linear response equations and will solve a user defined model of formation volumes.

GENERAL PRINCIPLES

The general principle used is the same as that described in earlier papers[1,2,3] However, only linear response equations are used and uncertainty values are defined for each mineral endpoint. The endpoint uncertainty values are used for the computation of standard deviations which are used to normalize the tool measurements and response equations.

The tool response equations are of the form:

$$f_i = \sum_{j=1}^{m} e_{ij} V_j \quad for\ i = 1,n \tag{2}$$

Where:
- V_j = fractional volume of mineral j
- m = the number of minerals being solved for
- f_i = theoretical tool response of tool i
- e_{ij} = mineral endpoint of tool i in mineral j

Although some of the fractional volumes may be gas-filled porosity or liquid-filled porosity rather than minerals, for simplicity all volumes will be described as "mineral" volumes for the rest of the paper.

Instead of specifying τ_i^2 and δ_i^2 for each log the total variance is defined as the variance σ_i associated with both the tool measurement and the tool response equation. The standard deviation (square root of variance) of the error associated with each tool reading is defined by:

$$\hat{\sigma}_i^2 = \sum_{j=1}^{m} \sigma_{ij}^2 V_j \quad for\ i = 1,n \tag{3}$$

Where:
- $\hat{\sigma}_i$ = the standard deviation associated with the error of tool i
- σ_{ij} = the standard deviation associated with the error of tool i in mineral j

Since the "answers" are required to compute these standard deviations some iteration is necessary. Fixed point iteration is used which converges rapidly.

The objective function (to be minimized) is now:

$$\Delta = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{m_i - f_i}{\hat{\sigma}_i}\right)^2} \tag{4}$$

or by defining $\hat{f}_i = f_i/\hat{\sigma}_i$ and $\hat{m}_i = m_i/\hat{\sigma}_i$

$$\Delta = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{m}_i - \hat{f}_i)^2} \tag{5}$$

Where $\hat{f}_i$ and $\hat{m}_i$ now have a standard deviation of one (1.0) for each i (or tool).

The system of equations

$$\hat{m}_i = \hat{f}_i \quad for \quad i = 1,n \tag{6}$$

is a system of linear equations each of which has the same standard deviation associated with it. It can be shown[6] that the least squares fit solution to eq. 6 is the minimum solution of eq. 5. By utilizing linear equations and defining $\hat{\sigma}_i$ using eq. 3, the problem of minimizing eq. 1 has been reduced to the simplier problem of finding the least squares fit to eq. 6.

To ensure a unique solution it is necessary to have at least as many independent equations as unknowns. An equation is said to be independent if it cannot be formed by any linear combination of the other equations. This is done by defining an equation for each unknown, and including these equations in the system of equations. These auxiliary equations will be of the form:

$$weight_j * V_j^{old} = weight_j * V \quad for \quad j = 1,m \tag{7}$$

Where $weight_j$ = scaling factor for auxiliary equation for mineral j
V_j^{old} = fractional mineral volume from previous iteration for mineral j

The above equations will always be independent and, therefore, the system of equations will have at least as many independent equations as unknowns. The scaling factors, $weight_j$, will also be used to deal with constraints.

The system of equations developed is linear, has been normalized, and now has at least as many independent equations as unknowns.

This system of equations satisfies the definition of a "general linear model". A complete discussion of the general linear model and its properties is given in the literature.[6,7] The least squares solution defined above is unique and will give the minimum value of Δ.[6,7] The Δ

defined above, is actually a "maximum liklihood estimator" (M.L.E.) of the standard deviation of the model and can be thought of as a measure of the fit of the data to the response equations. This value can be used, as discussed later, when evaluating multiple systems of equations to determine the system which best fits the data.

DEALING WITH CONSTRAINTS

Two constraints are needed to restrict the computed volumes to values that are physically possible. They are:

- Unity constraint - the sum of all volumes must be equal to one, and
- the non-negative constraint - no single volume can be allowed to be less than zero.

The first constraint could be satisfied by adding the following equation to the system:

$$1 = V_1 + V_2 + \ldots + V_m \tag{8}$$

If the system contains more equations than needed this equation will not be honored exactly. Therefore, a different method is needed to constrain the answer to volumes that sum exactly to one. Eq. 8 can be rewritten as

$$V_m = 1 - V_1 - V_2 - \ldots - V_{m-1} \tag{9}$$

The right side of this equation can be used in the system to replace V_m. This reduces the number of unknowns to $m-1$. Algebraically this is equivalent to simply subtracting the last endpoint from the tool response and each endpoint. The system of response equations, which includes eq. 6 and eq. 7, will be changed to

$$\hat{m}_i - e_{im} = \hat{f}_i - e_{im} \quad for \quad i = 1,n+m \tag{10}$$

The last mineral volume will then be given by eq. 9. This ensures that the mineral volumes will always sum to one. Reference 7 gives a general method for the incorporation of any linear constraints such as the unity constraint.

The auxiliary equations defined above are used to eliminate negative solutions. If a negative is obtained the negative mineral volume is set to zero, and the scaling factor is set to a large number, forcing the answer back to zero (or very near zero) in the next iteration. For example, if V_1, goes negative the following equation

$$0.01 * V_1^{old} = 0.01 * V_1 \tag{11}$$

would be changed to

$$0.0 = 100.0 * V_1 \tag{12}$$

and the value of V_1 will be very close to zero after the next iteration.

Other constraints may be defined and incorporated easily, but these are the only two that are necessary for the computation.

SOLUTION OF THE MATRIX

Eq. 10 above can be written in matrix form as

$$\underline{M} = E\,\underline{V} \tag{13}$$

where:

E = matrix of the endpoints of the first *m-1* volumes which have been normalized by their standard deviations, and have had the endpoint of the m'th mineral subtracted from them (as described earlier),

$\underline{V}$ = vector of the first *m-1* unknown volumes,

$\underline{M}$ = vector of data (log, core, etc.) which has been normalized by the standard deviations and have had the normalized endpoint of the m'th volume subtracted (as described earlier).

There are several methods for solving this equation.[6,7] The Modified Gram-Schmidt algorithm offers a good compromise of speed and reduction of round-off error. In this method matrix E is converted into a triangular matrix, and vector $\underline{V}$ is then determined by back substitution. A description of this algorithm is given in the appendix, as well as a short fortran subroutine which utilizes the Modified Gram-Schmidt algorithm to solve the matrix.

PROGRAM FLOWCHART

A flowchart showing the fixed point iteration is shown in fig. 1.

For the first iteration all the mineral volumes are set to equal values (that sum to one). The non-negative constraint is imposed (for subsequent iterations) by checking the volumes for negative values. If a volume is found negative it is set to zero, and the weight of its auxiliary equation is increased (forcing that volume close to zero in the next iteration). The standard deviations, $\hat{\sigma}_i$, of the tool responses, m_i, are computed from the volumes, V_j, and the standard deviations, σ_{ij} . The endpoints, e_{ij}, and the log values, m_i, are normalized by the standard deviations, $\hat{\sigma}_i$, calculated for each tool response. The auxiliary equation for each volume is added to the system of equations. (These independent equations ensure that each iteration will always have a unique solution.) The unity constraint is imposed by subtracting (for each log) the last endpoint, e_{im}, from the log value, m_i, and the remaining endpoints, e_{ij}. The system of normalized, linear equations in matrix form is solved using the Modified Graham-Schmidt algorithm for mineral volumes , $V_1, \ldots, V_{m-1}$. The last volume, V_m, is then calculated from the unity equation (eq. 9). The mineral volumes are compared with the previous volumes,

V_j^{old}, to test for convergence. If the convergence criterion are not met the iteration continues. Once the convergence criterion is met the Δ value is calculated from eq. 4. This is, as the system has been defined, the maximum liklihood estimate (M.L.E.), and can be used to evaluate the fit of the model to the data.

USING Δ AS A QUALITY INDICATOR

The Δ as computed above, is the minimized value of the objective function (eq. 4) mentioned earlier. Under reasonable assumptions it can be shown that the answers obtained in this manner are the most probable[1].

In practical terms Δ can be thought of as a measure of the fit of the model to the data. In the following discussion "model" refers to the set of mineral endpoints and their associated standard deviations (e_{ij},σ_{ij}). Theoretically, with an infinite number of logs and correctly defined model the Δ value will be close to one. Since our systems have little or no degrees of freedom (slightly overdetermined or balanced systems) Δ should be less than one (again if the model has been properly defined). Δ values greater than one are usually the result of the following:

- bad hole resulting in unreliable data,
- incorrect selection of endpoint values,
- the presence of mineral volumes not accounted for in the model,
- data not properly depth correlated,
- bad data (improper calibration, tool malfunction, etc.).

Since the Δ is a measure of the fit of the model to the data, it is intuitive that it can be used as a comparison tool to evaluate different combinations of minerals (models). The combination (model) which best fits the data would then be assumed to be the most probable.

The above logic is utilized in several ways to help deal with "underdetermined" problems.

While a high Δ indicates the model does not fit the data, a low Δ does not verify that the model is correct (i.e., different models may fit the data equally well, but give different answers). This should be considered when evaluating the results of these computations.

Another quality check is the comparison of the measured log values, m_i, with the theoretical (reconstructed) tool response, f_i. Confidence intervals for the volumes, V_i, can also be used to evaluate the quality of each volume computed. Details of the computation of these confidence intervals are given in Appendix B.

UNDERDETERMINED PROBLEMS

In some cases enough logs are available to solve for the minerals present. Often, however, there are not enough logs to solve for all minerals in the formation. The number of equations or logs in a system, in relation to the number of unknowns (or mineral volumes), defines a

system of equations as either "underdetermined", "overdetermined, or "balanced". An "overdetermined" system of equations has more independent equations than unknowns. The statistical approach computes the solution which is the most probable. A system of equations is "balanced" when there are an equal number of equations and unknowns. In this case there is only one solution. An "underdetermined" system of equations has fewer equations than unknowns. In this case there will be an infinite number of solutions. Mathematically, the system of equations defined above will always have at least as many equations as unknowns because of the auxiliary equations, and therefore, a unique answer can always be found. (The answer in this case would be weighted toward making the volumes equal.) In order to compute a meaningful solution the system of equations must be either balanced or overdetermined. When there are not enough logs to define a balanced or overdetermined system a number of methods may be used.

Some methods commonly used include:

- Alternate Minerals
- Multiple Models
- Sequential Models
- Computed Data Input

Alternate Minerals - To use this method an alternate mineral is defined and entered in the initial matrix. The answer is computed as outlined earlier without using the alternate mineral. If one or more of the mineral volumes computed is negative (before forcing to zero) the Δ value is saved, and the system solved again using the alternate instead of the most "negative" mineral. If the solution again has a "negative" the solution with the lowest Δ will be retained. The alternate mineral approach can also be expanded by using more than one alternate and assigning a substitution priority for each alternate mineral.

Multiple Models - Two or more models can be defined and computed separately. The solution to be output corresponds to the model (or system of equations) which has the lowest Δ. This technique is useful in an interval containing several minerals of which two or more do not occur simultaneously.

Sequential Models - To use sequential models two (or more) models are defined. The first model is solved, and one (or more) of the volumes computed is used as data input into the next model (and so on if there are more than two models). An example of this is the use of "M", "N", and Gamma Ray, to compute dolomite, limestone, anhydrite and shale volumes, then using the anhydrite and shale values as inputs to the second model, which is then solved with the rest of the data.

Computed Data Input - Raw data may be computed and input into the model in an alternate form. "M" and "N" values are typically crossplotted and used in log analysis. These values may be used in the model, the endpoints and standard deviations being read directly off the crossplot (discussed later). Another common application is the use of a shale volume (Vsh) computed from the minimum (or average, etc.) of several inputs. The computed shale volume

could be crossplotted with other data to define its standard deviation and then entered in the model.

While this is not a comprehensive list of techniques to handle underdetermined problems these are the ones found to be most useful. They may be employed in any combination allowing a very large amount of flexibility.

NON-LINEAR RESPONSE EQUATIONS

Several methods are employed when it is necessary to deal with non- linear responses. It is well known that SP, Gamma Ray, and resistivity can be used as shale indicators, however, they are often non-linear. These measurements can be transformed using Clavier, Steiber, reciprocal, or various S-curve or exponential relationships, and entered into the models as Vsh values. The minimum of several Vsh calculations can also be used, as well as a variety of averaging techniques. Porosity from the sonic log can be computed from the desired method (e.g., non-linear Hunt-Raymer transform) and used as sonic porosity in the model, instead of sonic transit time. Rather than using the non-linear Pe measurement the linear U can be used in the model. In dealing with the neutron log a linear approximation of the response can be made, or, if necessary, provisions can be made for the mineral endpoint to vary with porosity. In practice, therefore, the accuracy of the computation is not significantly reduced by using only linear equations.

ESTIMATION OF STANDARD DEVIATIONS

The standard deviation of each log reading in each mineral is a measure of the uncertainty of that log reading in that mineral (volume). Since the systems to be solved are only slightly overdetermined, at best, the standard deviation values do not need to be precise. Fig. 2 shows a table of reasonable standard deviation values for typical minerals and logging tools which have been found, by experience, to give satisfactory results. Crossplots can be used to define these standard deviations more precisely in some cases. In a section containing only one mineral approximately 70% of the data points will fall within ± one standard deviation of the endpoint of that mineral. In fig. 3 the bulk density and sonic transit time endpoints for salt can be identified as 2.08 gm/cc and 69.3 micro-sec/ft. The brackets enclosing the salt endpoint show the standard deviation to be .03 g/cc for the bulk density and 0.5 micro-sec/ft for the sonic transit time.

COMPLEX LITHOLOGY EXAMPLE

Fig. 4 illustrates a computation in complex lithology. The interval shown is known to contain limestone, dolomite, anhydrite, salt, shale and porosity. The available logs are sonic transit time, bulk density, neutron porosity, and Gamma Ray. There are six mineral volumes and four logs, therefore, a technique for dealing with underdetermined problems must be used. In this example the Multiple Models method was used. Model 1 mineral volumes include anhydrite, limestone, salt and shale, while Model 2 mineral volumes include limestone, dolomite, anhydrite, shale, and porosity (see fig. 5). Both models are solved at each foot and

the solution with the lowest Δ will be retained. Note that at the top of the example Model 1 best fits the data (lowest Δ). Below this interval Model 2 best fits the data, and correctly identifies the high porosity dolomite intervals. This example could have also been solved using the Alternate Minerals method, described earlier.

CONCLUSION

Using linear equations in a statistical algorithm to compute lithology and porosity retains the advantages of non-linear statistical approaches and eliminates many of the disadvantages. The linear approach presented runs much faster, is more flexible, and is easier to incorporate into existing software than non-linear methods. This approach was incorporated into Sun Exploration and Production Company log analysis software five years ago. Its subsequent use on thousands of wells has verified these advantages.

NOMENCLATURE

Δ = Solution of the objective function
n = Number of tool response equations in the system of equations
m = Number of unknown mineral volumes
f_i = Theoretical tool response of tool i
m_i = Actual tool measurement of tool i
τ_i = Standard deviation associated with the error of tool measurement i
δ_i = Standard deviation associated with the error of the theoretical response equation of tool i
V_j = Fractional volume of mineral j
e_{ij} = Endpoint of mineral j for tool i
σ_{ij} = Standard deviation associated with the error of tool i in mineral j
$\hat{\sigma}_i$ = Standard deviation associated with the error of tool i
$\hat{m}_i$ = Normalized value of tool measurement i
$\hat{f}_i$ = Normalized value of theoretical tool measurement i
V_j^{old} = Fractional mineral volume from previous iteration for mineral j
$weight_j$ = Scaling factor for auxiliary equation for mineral j
$\underline{M}$ = Measurement data vector
E = Endpoints matrix
$\underline{V}$ = Fractional volumes vector

ACKNOWLEDGEMENTS

We would like to thank Sun Exploration and Production Company for permission to publish this paper.

REFERENCES

1. Mayer, C., and Sibbit, A.: "Global, A New Approach to Computer-Processed Log Interpretation", 55th Annual Meeting of Society of Petroleum Engineers of AIME, Dallas, Sept., 21-24, 1980.
2. Alberty, M., and Hashmy, K.: "Application of ULTRA to Log Analysis", SPWLA 25th Annual Logging Symposium, June 10-13, 1984.
3. Mezzatesta, A., Rodriquez, E., and Frost, E.: "OPTIMA - A Statis tical Approach to Log Analysis", International Symposium, Petroleum Exploration in Carbonate Areas, Nanjing, China. Nov. 12-19, 1986.
4. Quirein, J., Kimminau, S., LaVigne, J., Singer, J., Wendel, F.: "A Coherent Framework for Developing and Applying Multiple Formation Evaluation Models", SPWLA 27th Annual Logging Symposium, June 9-13, 1986.
5. Gysen, M., Mayer, C., Hashmy, K.: "A New Approach to Log Analysis Involving Simultaneous Optimization of Unknowns and Zoned Parameters", CWLS 11th Formation Evaluation Symposium, Calgary, Sept. 8-11, 1987.
6. DeGroot, M. H.: *Probability and Statistics*, Addison-Wesley Publishing Co., Inc., 1975.
7. Dahlquist, G., Bjorck, A., Anderson, N.: *Numerical Methods*, Prentice Hall, Inc., 1974.

APPENDIX A

The modified Gram-Schmidt method solves the matrix equation (eq. 13 from text)

$$\underline{M} = E\,\underline{V}$$

by decomposing E into the product of two matrices,

$$E = QR \qquad (A1)$$

Where Q is a matrix with orthoganal columns and R is an upper triangular matrix, and

$$Q^TQ = D \qquad (A2)$$

where D is diagonal matrix with positive diagonal elements.

Then $$R\,\underline{V} = \underline{Y} \qquad (A3)$$

where $$\underline{Y} = D^{-1}\,\underline{Q^T M} \qquad (A4)$$

Therefore, if Q and R are known $\underline{Y}$ can be determined and eq. 3 can be solved by back substitution. The modified Gram-Schmidt algorithm for computing R and Y along with the solution of eq. 3 by back substitution is given in Fortran below:

```
      SUBROUTINE MINMAT(NCE,NRE,E,M,V)
      REAL E(20,20),M(20),V(20)
      REAL R(20,20),D(20),Y(20)
      REAL B(20),W(20,20)
C
C     E     -     MATRIX OF NORMALIZED ENDPOINTS
C     M     -     VECTOR OF NORMALIZED LOG VALUES
C     NCE   -     NUMBER OF COLUMNS IN E
C     NRE   -     NUMBER OF ROWS IN MATRIX E
C     V     -     VECTOR CONTAINING SOLUTION
C
C     R     -     UPPER TRIANGULAR MATRIX WITH NRE COLUMNS AND NRE ROWS
C     D     -     DIAGONAL MATRIX WITH NRE COLUMNS AND NRE ROWS
C     Y     -     VECTOR OF LENGTH NRE
C
C     B     -     WORK VECTOR OF LENGTH NRE
C     W     -     WORK MATRIX WITH NCE COLUMNS AND NRE ROWS
C
C
C     COMPUTE R AND Y FOR R*V=Y BY THE MODIFIED GRAM-SCHMIDT ALGORITHM
C
      DO 500 I=1,20
      DO 500 J=1,20
      W(I,J) = 0.0
500   B(I) = 0.0
```

```
      DO 501 I=1,NRE
      B(I)  = M(I)
      DO 501 J=1,NCE
501   W(I,J) = E(I,J)
      DO 110 K=1,NCE
      R(K,K) = 1.0
      D(K)  = 0.0
      DO 120 J=1,NRE
      D(K) = D(K) + W(J,K) ** 2.0
120   IF (D(K) .EQ. 0.0) D(K) = .0000001
      DO 130 J=K+1,NCE
      R(K,J) = 0.0
      DO 140 L=1,NRE
140   R(K,J) = R(K,J) + W(L,K) * W(L,J)
      R(K,J) = R(K,J) / D(K)
      DO 130 L=1,NRE
130   W(L,J) = W(L,J) - R(K,J) * W(L,K)
      Y(K) = 0.0
      DO 160 L=1,NRE
160   Y(K) = Y(K) + W(L,K) * B(L)
      IF (D(K) .EQ. 0.0)D(K)=.00001
      Y(K) = Y(K) / D(K)
      DO 110 L=1,NRE
110   B(L) = B(L) - Y(K) * W(L,K)
C
C           SOLVE THE EQUATION R*V=Y FOR THE V ARRAY
C
      V(NCE) = Y(NCE)
      IF (NCE.EQ.1) GO TO 101
      DO 200 I=NCE-1,1
      SUM = 0.0
      DO 210 J=I+1,NCE
210   SUM = SUM + R(I,J) * V(J)
200   V(I) = Y(I) - SUM
101   CONTINUE
      RETURN
      END
```

APPENDIX B

The system of equations (eq. 13 from text)

$$\underline{M} = E\,\underline{V}$$

can be solved by the so-called "normal" equations

$$E^T E\,\underline{V} = E^T \underline{M} \tag{B1}$$

The solution to these equations are

$$\underline{V} = (E^T E)^{-1} E\,\underline{M}. \tag{B2}$$

However, when E is ill-conditional, serious round off errors can result from computing $\underline{V}$ by this technique.

The variance, γ_j^2, of the estimate of V_j (for $j = 1,m$) is given by the diagonal elements of $(E^T E)^{-1}$. E is the matrix of endpoints which have been normalized by their standard deviations (in this case the last endpoint has not been subtracted from each endpoint as in eq. 13).

A new variable U_j can be defined by

$$U_j = \left(\frac{n-m}{n}\right)^{1/2}\left[\frac{v_j - v_j^*}{\Delta\gamma_j}\right] \quad for\ j = 1, m \tag{B3}$$

The variable U_j will have a t distribution with n-m degrees of freedom.

If $\propto$ is the choosen confidence level then the value of u_j such that

$$P(-u < U_j < u_j) = \propto \quad for\ j = 1,m \tag{B4}$$

can be determined from a table of the t distribution. A table of the distribution can be found in reference 6.

Then the confidence interval about V_j is $(V_j - d_j,\ V_j + d_j)$, where

$$d_j = u_j \Delta\gamma_j \left(\frac{n}{n-m}\right)^{1/2} \quad for\ j = 1,m \tag{B5}$$

ABOUT THE AUTHORS

Richard J. Nelson is a Log Analyst with Sun Exploration & Production Company in Dallas, TX. He attended The Johns Hopkins University and received a Bachelor of Engineering Science in 1977. He worked for Schlumberger Well Services eight years before joining Sun Exploration and Production Company in 1985. He is a Registered Professional Engineer in the State of Texas, and a member of SPE and SPWLA.

William K. Mitchell is a Log Analyst with Sun Exploration & Production Company in Dallas, TX. He holds a BS in Physics from Indiana University and a MS in Mathematics from University of Texas at Dallas. He worked for Schlumberger Well Services six years before joining Sun Exploration and Production Company in 1974. He is a member of SPE and SPWLA.

S

Program Flow Chart

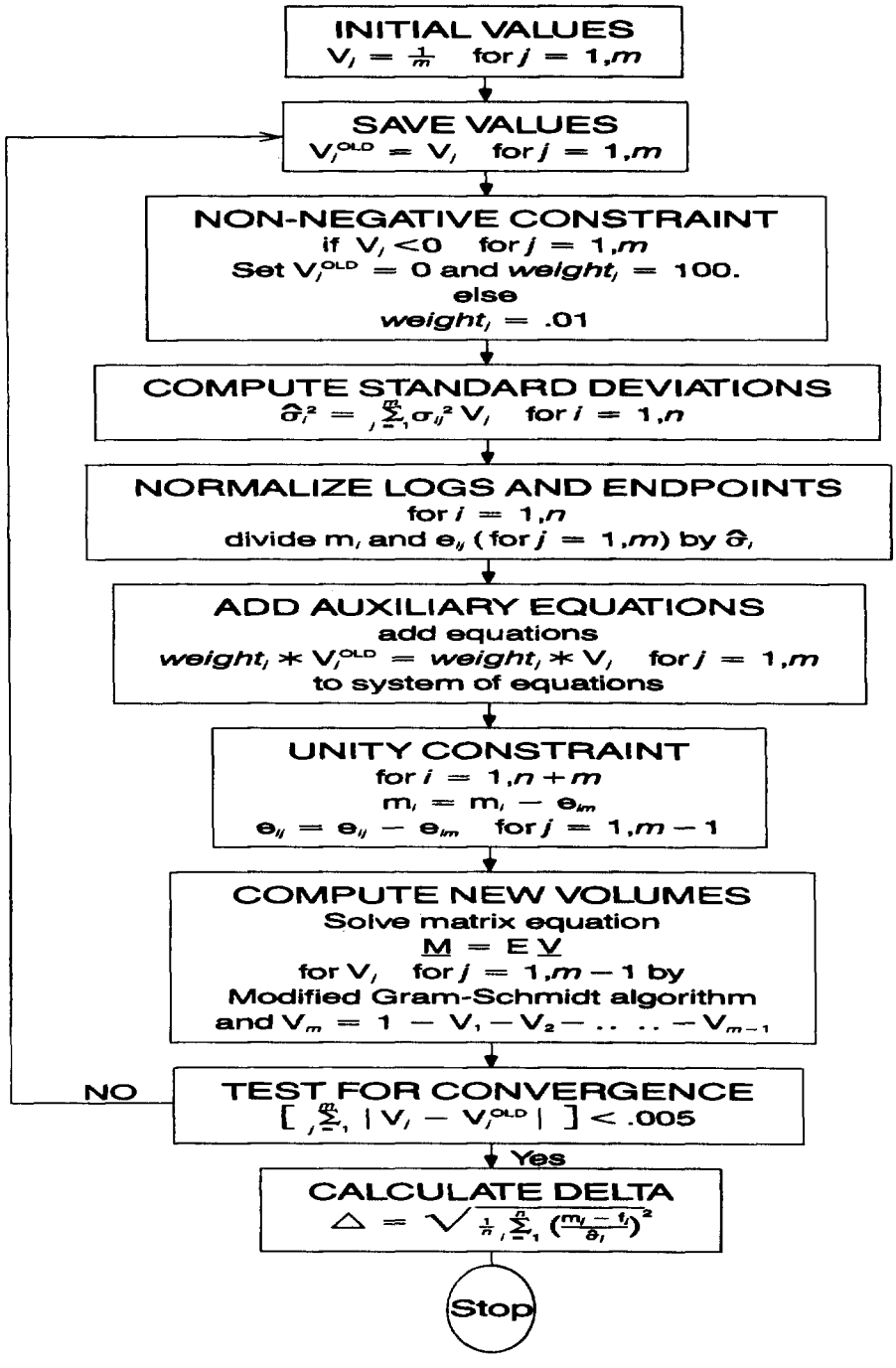


Fig. 1 Flow Chart showing fixed point iteration.

Table of Typical Standard Deviations

	Bulk Density	Neutron Porosity	Sonic Δt	Volumetric Cross Section-U	GR API units
Porosity (Liquid)	.01	3.0	5.0	.30	5
Porosity (Gas)	.01	1.0	5.0	.30	5
Sandstone	.02	1.0	3.0	.30	5
Limestone	.01	0.5	3.0	.50	5
Shale	.03	3.0	3.0	.50	5
Dolomite	.02	2.0	3.0	.30	5
Anhydrite	.02	1.0	1.0	1.50	5
Gypsum	.02	1.0	3.0	.30	5
Salt	.03	1.0	1.0	.30	5
Coal	.03	1.0	3.0	.30	5

Fig. 2 Table of standard deviations for commonly encountered minerals and tools. The data is an empirical combination of tool theory and experience.

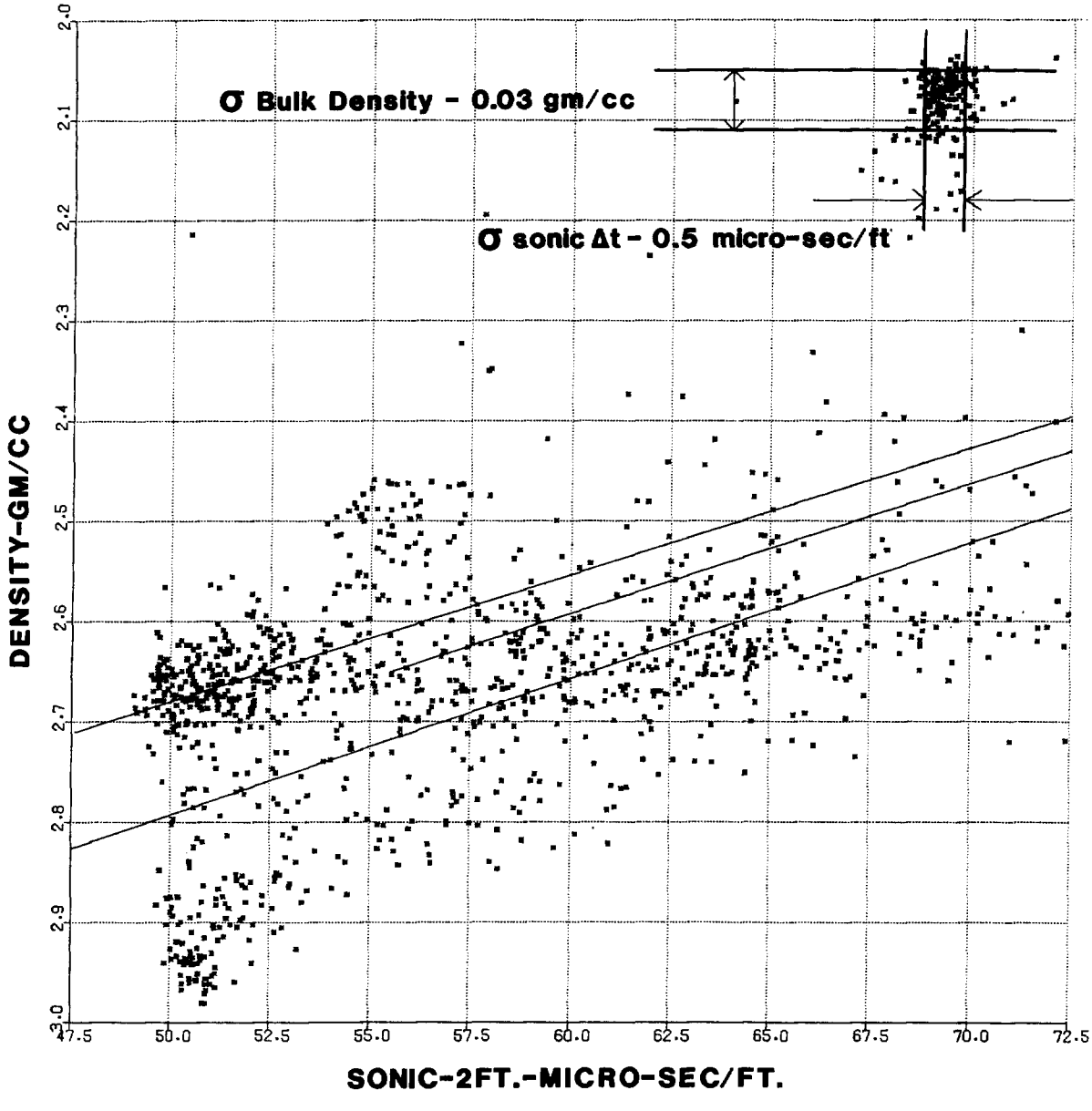
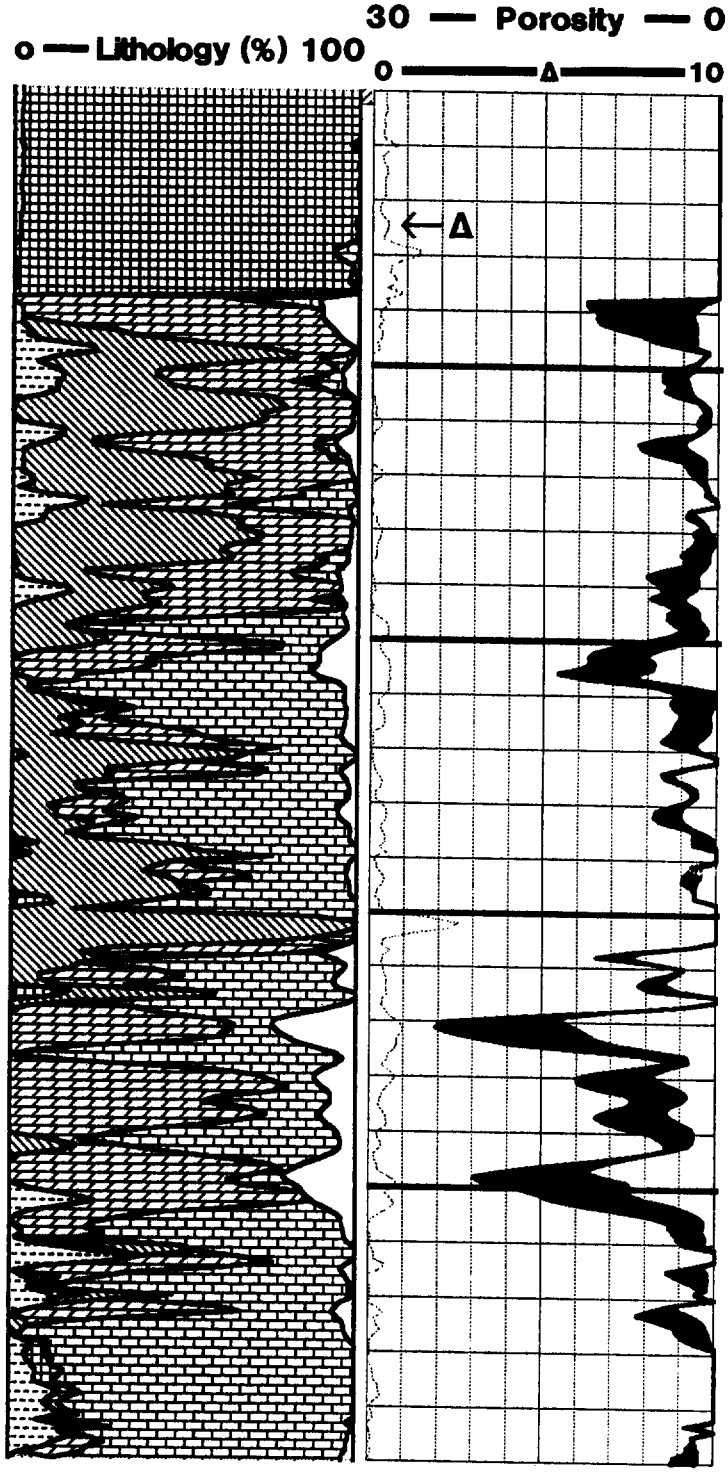


Fig.3 **Crossplot showing how standard deviation can be estimated for salt endpoint. (Data from complex Lithology example.)**

Complex Lithology Example



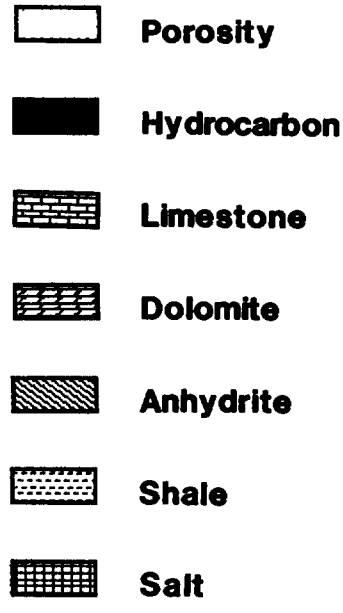


Fig. 4 Complex lithology example using data shown in fig. 3 and models in fig. 5.

Complex Lithology Example - Multiple Models

MODEL 1

	Limestone	Anhydrite	Salt	Shale
Bulk Density (gm/cc)	2.71	2.94	2.08	2.62
Standard Deviation	0.01	0.02	0.03	0.03
Neutron Porosity (%)	0.0	-0.6	-1.2	24.0
Standard Deviation	0.5	1.0	1.0	3.0
Sonic Δt (micro-sec/ft.)	47.6	50.5	69.3	85.0
Standard Deviation	3.0	2.0	0.5	3.0
Gamma Ray (API)	34	11	24	91
Standard Deviation	16	3	5	16

MODEL 2

	Porosity	Limestone	Dolomite	Anhydrite	Shale
Bulk Density (gm/cc)	1.10	2.71	2.87	2.94	2.62
Standard Deviation	0.02	0.01	0.02	0.02	0.03
Neutron Porosity (%)	100	0.0	6.0	-0.6	24.0
Standard Deviation	3.0	0.5	2.0	1.0	3.0
Sonic Δt (micro-sec/ft.)	185	47.6	44.5	50.5	85.0
Standard Deviation	5	3.0	3.0	2.0	3.0
Gamma Ray (API)	34	34	34	11	91
Standard Deviation	16	16	16	3	16

Fig. 5 Models used to solve complex lithology example in fig.4.