A Rigorous Mathematical Approach for Petrophysical Properties Estimation

Dario Viberti
Department of DITAG, Politecnico di Torino,
Corso Duca Degli Abruzzi 24, 10129 Torino, Italy

Abstract: Problem statement: Over the last few decades, the oil industry has shown a growing interest in the new risk analysis methodologies aimed at the evaluation of the uncertainties associated with reservoir exploitation. In particular, the effort is made to take all possible sources of uncertainties into account so that not only the strengths but also the potential weaknesses of each possible technical and economic exploitation strategy are highlighted. Approach: The main parameters used to calculate strategic information, such as the Hydrocarbon Originally In Place (HOIP) and to define proper field development plans, are porosity, fluid saturations and Net To Gross (NTG). These quantities are typically obtained through the log interpretation process, which is an inverse problem where the main petrophysical characteristics are calculated as the acceptable minimum of a cost function. The cost function describes the discrepancy between measured and simulated logs with the latter being reproduced on the basis of an assumed formation model. Results: The results of the calculation process can be affected by several uncertainties related to the physics and calibration of the measuring tools, the identification of the proper formation model and the quantification of the formation model coefficients. An effective and robust methodology able to provide a reliable evaluation of petrophysical properties and the assessment of the associated uncertainties is presented and discussed in this study. The log interpretation process was approached as a linearly constrained optimization problem, solved by coupling a Lagrange-Newton method with a primal active set algorithm. Conclusion: The evaluation of the uncertainties was obtained by coupling the optimization algorithm with the Monte Carlo approach. The results obtained by the application of the methodology to a real case, where the interpretation was complicated by a poor characterization of the reservoir fluids, are also presented the study.

Key words: Log interpretation, petrophysical properties, uncertainties, inversion problem, Hydrocarbon Originally In Place (HOIP), Net To Gross (NTG), Lagrange-Newton

INTRODUCTION

Even if the evaluation of the uncertainties in the well log interpretation process dates back to some years ago (Ventre, 1994; Bowers and Fitz, 2000; Hall et al., 2001; Verga et al., 2001; Rocca, 2009; Verga and Rocca, 2010), the application of the risk analysis concepts to reservoir characterization is very recent because only modern computer science allowed rapid processing of extremely large amounts of data.

The calculation of the Hydrocarbon Originally In Place (HOIP) is based on the evaluation of the main petrophysical parameters, i.e., porosity and water saturation. These quantities cannot be directly measured, therefore, need to be estimated through an interpretation process consisting of combining log measurements and core data when available. The formation response is observed in the log recordings but the minerals and fluids contributing to this response are unknown. However, since each measured property can be described by a mathematical relationship (forward model) that is based on minerals and fluids fractional volumes, it is possible to couple these equations with the observed data and to solve the resulting inverse problem so as to infer the unknown quantities.

Each piece of information introduced in the interpretation process, i.e. log measurements, model coefficients, can be affected by an uncertainty that needs to be quantified and corrected whenever possible. Wellbore measurements are generally influenced by the presence of a mud invaded zone in the proximity of the well. Well drilling operations, in fact, require that mud is circulated in the well. The pressure gradient between the fluids in the wellbore and in the formation induces the mud fluid to invade a reservoir zone around the wellbore displacing the fluids originally in place (groundwater and/or hydrocarbons). Furthermore, different types of errors can affect the available
reservoir information. Errors can be due to inaccurate calibration of the measuring instruments, incorrect choice of the forward model parameters or the physics of the measurement system. Thus, some of the uncertainties affecting the input of log interpretation cannot be eliminated. It follows that the impact of these uncertainties on the final results of the interpretation process has to be evaluated through appropriate propagation methods which need to account for the interpretation models but also for the solution algorithms. In modern log analysis, the log interpretation process is treated as an inverse problem. The solution is obtained by adopting a constrained optimization algorithm to automatically minimize a cost (or objective, or error) function. This cost function represents the discrepancy between measured log data and simulated log response. Even if some optimization processes are partially probabilistic in their nature, they do not provide an evaluation of the uncertainty associated with the obtained results or require too strong assumptions on the uncertainty distributions characterizing the input data (Bertolini et al., 2009; Sen and Stoffa, 1995).

The purpose of this study is to present an algorithm obtained by the integration of a Lagrange-Newton optimization method along with a Monte Carlo approach to simultaneously find the solution of the log interpretation process, which is an inverse constrained problem and assess the uncertainties affecting the main petrophysical properties estimated from log measurements.

**MATERIALS AND METHODS**

**Problem and model definition:** The scientific procedure for studying a physical system can be divided into three steps: parameterization of the system, forward modeling and inverse modeling. The parameterization of the system consists in identifying a minimal set of model parameters whose values completely characterize the system. Forward modeling consists in determining the physical laws allowing predictions of the values of some observable quantities, for given values of the model parameters. Inverse modeling is the use of actual measurements to infer the values of the model parameters (Tarantola, 1987).

In the log interpretation process, the unknowns of the problem are the minerals and fluids fractional volumes. The forward model is mainly represented by the log response equations, i.e., the set of equations relating the log response to the fractional volumes of the formation components. In most log interpretation problems, the forward model is a non-linear operator since it is constituted by a set of equations containing both linear and non linear log response equations. During the development of the methodology, several log response equations were analyzed and implemented. Linear equations were implemented for the density, neutron, sonic and GR logs. Conductivity or resistivity log responses can be calculated by selecting the most suitable or preferred model: Archie, Indonesia and Nigeria equations are available. All three equations are non-linear, but the Indonesia and Nigeria models are more complicated due to the presence of the term expressing the shale content of the formation. For the purpose of this study, only the Indonesia formula is presented.

Let \( x \in \mathbb{R}^n \) be a model configuration, i.e. the vector of the fractional volumes of the formation components (solids and fluids). In the case of fluid components, subscripts \( w \) and \( h \) indicate water \( (x_w) \) and hydrocarbon \( (x_h) \), respectively, while subscripts \( x \) and \( u \) indicate the fractional volume of fluids in the invaded \( (x_{wu}) \) and undisturbed \( (x_{hu} \) and \( x_{nu} \) \) zones, respectively. Each component of vector \( x \) can range from 0-1. The function \( f(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m \) represents the system of the \( m \) forward model equations which relate the log responses to the fractional volumes of the formation components. Each equation represents the response of a single one of the \( m \) considered logs. The forward models of density, nuclear and sonic logs are represented by linear equations of the type:

\[
f_i(x) = \sum_{j=1}^{m} C_{ij} x_j + (1 - \alpha_i)(c_{wux_{wu}} + c_{ux_{nxu}}) + \\
+ \alpha_i(c_{wu}x_{wu} + c_{ux_{nxu}})
\]

(1)

where, \( i = 1, \ldots, m \) refers to the logs (Neutron, density), \( j = 1, \ldots, n \) refers to the solid components (for instance quartz, sand, clay,...), \( c_{ij} \) is the \( ij \)th log response in the pure component \( i \), \( \alpha_i \in [0,1] \) weighs the influence of the invaded zone on the measurement, \( \alpha_i = 1 \) indicates a measure influenced only by the invaded zone, while \( \alpha_i = 0 \) indicates a measure influenced only by the undisturbed zone).

As previously mentioned, the forward model of the conductivity measurements is represented by a non-linear equation. The following empirical Indonesia Formula (Poupon and Leveaux, 1971) is an example of a non-linear equation, where \( C_i \), indicates the total conductivity:

\[
f_{ci}(x) = \sqrt{C_i} = \left[ \sqrt{C_{cl}^{1-0.5x_{sl}}} + \frac{C_{ct} - \frac{0.5x_{sl}^{1-0.5x_{sl}}}{\phi_{c}}}{a} \right] \left( \frac{x_{wu}}{\phi_{c}} \right)\frac{z}{\phi_{c}}
\]

(2)
Where:

\( C_{\text{cl}} \) and \( x_{\text{cl}} \) = The clay conductivity and the clay fractional volume, respectively.

\( C_{\text{w}} \) and \( x_{\text{wu}} \) = The water conductivity and the water fractional volume, respectively.

\( a \) = The lithology coefficient.

\( m \) and \( m_{c2} \) = The cementing exponent and the correction factor for \( m \).

\( n \) = The saturation exponent.

\( \phi_e \) = The effective porosity calculated as:

\[
\phi_e = x_{\text{mu}} + x_{\text{wu}}
\]  

(3)

If \( r \in \mathbb{R}^n \) represents the vector of the observed log values at a single depth point, then the discrepancy between the measured and predicted response can be defined by the residual vector \( e(x) = r - f(x) \). It follows that the objective function is defined as:

\[
\text{obj}(x) = \frac{1}{2} \left[ \text{We}(x) \right]^T \left[ \text{We}(x) \right]
\]  

(4)

where the diagonal matrix \( \text{W} \in \mathbb{R}^{n \times n} \) has been introduced to define the weight to be assigned to each log. Weights are calculated by taking into account the normalization factors and the noise affecting each log measurement.

The minimum of the objective function often corresponds to a physically meaningless solution such as negative values for fractional volumes of formation components. Therefore, the problem has to be formulated so as to take into account some constraints classified in equality constraints \( c_i(x) \) and inequality constraints \( c_i(x) \):

\[
\begin{align*}
\text{Minimize} & \quad \text{obj}(x) \\
\text{Subject to} & \quad c_i(x) = b_i \\
& \quad c_i(x) \geq 0
\end{align*}
\]  

(5)

The main equality constraints are: fractional volume balance (6) and horizontal continuity of porosity (7):

\[
\sum_{i=1}^{n} x_{\text{cl}} + x_{\text{wu}} + x_{\text{mu}} = 1
\]  

(6)

\[
x_{\text{wu}} + x_{\text{mu}} = x_{\text{wu}} + x_{\text{mu}}
\]  

(7)

Note that Eq. 7 follows from Eq. 3 by assuming a homogeneous value of porosity along the investigation depth of the whole log.

The main inequality constraints impose that each volume component can only have non-negative values (8):

\[
x_i \geq 0 \quad i = 1, \ldots, n
\]  

(8)
terminated before reaching the solution, the terminating point is feasible.

Therefore, the solution to equation (9) is determined by the exploration of the feasibility region \( \Omega \), whose boundaries \( \partial \Omega \) are defined by the set of constraints which hold exactly. Since \( \Omega \) can change at each step of the exploration, the minimum of the objective function strictly depends on the chosen set of binding constraints. In other words, the inequality constraints are divided in two dynamic subsets \( i \{ C, C \} \). Subset \( i C \) is the matrix of the inactive constraints and subset \( i C \) represents the matrix of the active ones. It follows that, at each step of the exploration, the matrix \( C \) changes and is defined as

\[
C = \begin{bmatrix} C_i \\ C_i \end{bmatrix}
\]

Since it is reasonable to assume that the petrophysical properties at different depth points are independent, problems of the form of (10) will be solved at each depth point independently.

**Problem solution:** The solution of the problem as shown in (10), depends on the choice of the optimal set of active constraints, which define the feasibility region and on the minimization of the objective function. Then, an Active Set method (Luenberger, 2003) coupled with the Lagrange-Newton method (Fletcher, 1987) needs to be applied to solve the linearly-constrained minimization problem under the hypothesis that the system is well-determined. The optimal solution can thus be theoretically obtained in a finite number of steps (Fletcher, 1987).

Let’s suppose that the set of active constraints, i.e. the working set, is fixed.

Let’s now introduce the Lagrangian function:

\[
L(x, \lambda) = \text{obj}(x) = x^T (C x - b)
\]

where, \( \lambda \in \mathbb{R}^p \) and \( \lambda_i \) is the \( i \)-th Lagrange multiplier \((1 \leq i \leq p)\). Then the Lagrangian relaxation of (10) is:

Minimize \( L(x, \lambda) \)

Let’s now define a local quadratic model of \( L \) in a neighborhood \( N_0 \) of \([x_0, \lambda_0]\), that is a generic feasible solution:

\[
q_0(\delta) = L_0 + \nabla L_0^T \delta + \frac{1}{2} \delta^T H_0 \delta
\]

with

\[
\delta = \begin{bmatrix} \delta_x \\ \delta_\lambda \end{bmatrix} = \begin{bmatrix} x - x_0 \\ \lambda - \lambda_0 \end{bmatrix}
\]

Where:

\[
L_0 = L(x_0, \lambda_0)
\]

\[
\nabla L_0 = \begin{bmatrix} g_0 - C^T \lambda_0 \\ -C x_0 + b \end{bmatrix}
\]

\[
H_0 = \begin{bmatrix} H_0 & -C^T \\ -C & 0 \end{bmatrix}
\]

and:

\[
g_0 = J_0^x W^T e(x_0)
\]

\[
H_0 = J_0^x W^T J_0 + Q_0
\]

\[
J_0 = \nabla f(x_0)
\]

\[
Q_0 = \sum_{i=1}^{n_0} w_i^2 H_{ii}(x_0) e_i(x_0)
\]

It is worth noting that the objective function is twice continuously differentiable in a neighborhood of a feasible solution \( x_0 \) since the log responses are linear with respect to \( x \). Even if the conductivity model is non-linear, it is also continuously differentiable. Since the Lagrangian relaxation depends linearly from \( \lambda \), \( L_0 \) can also be supposed twice continuously differentiable in \( N_0 \): \( \nabla q_0 = 0 \Rightarrow H_1 \delta_\lambda = -\nabla L_0 
\]

\[
H_0 \delta_x - C \delta_\lambda = -g_0 + C^T \lambda_0
\]

\[
C \delta_\lambda = -C x_0 + b
\]

Since the optimization algorithm moves only from one feasible point to another, then there is no loss of generality if it is assumed that \( x_0 \), which is the result of a generic iteration, is a feasible solution for the original problem. Thus \( C x_0 = b \). Moreover let \( d = \delta_x \) and \( \lambda = \lambda_0 + \delta_\lambda \), then system (21) becomes:

\[
\begin{bmatrix} H_0 & -C^T \\ -C & 0 \end{bmatrix} \begin{bmatrix} d \\ \lambda \end{bmatrix} = \begin{bmatrix} g_0 \\ 0 \end{bmatrix}
\]

In order to reduce the computational cost, the term \( Q_0 \) in the definition of \( H_0 \) was neglected in the Gauss-Newton approach. This approximation makes the second order operator dependent only on the first order derivatives (\( J_0 \)). As a consequence, the Hessian matrix
can be decomposed as \( H_0 = A^T A \) with \( A = W J_0 \) real and non-singular. This implies that \( H_0 \) is symmetric and positive definite \( \forall x \in N_0 \) (Ayres, 1962).

Since \( C \) is full rank and \( H_0 \) is positive definite, it follows that the coefficient matrix in \( (22) \) is non-singular (Luenberger, 2003). Thus, the solution \( [d, \lambda] \) can be obtained in a closed form concluding the following:

\[
x_* = x_0 + d, \lambda = \lambda
\]

(23)

Note that the solution of the linear system \( (22) \) requires that the inversion of the Hessian matrix \( H_0 \) is performed. In general, the inversion is not directly computed and the use of a factorization strategy is preferred. In the analyzed problem, taking advantage of the symmetry of \( H_0 \), the following unique decomposition can be calculated: \( H_0 = LDL^T \) where \( L \) is a lower triangular matrix and \( D \) is a diagonal matrix (Gill et al., 1991).

Integration of uncertainty estimation: Theoretically, the Monte Carlo method provides the possibility of simulating a high number of measurements that, due to economic reasons, cannot be performed in common practice. In all Monte Carlo simulations, it is necessary to draw statistically representative samples from given probability distributions (Spanier and Gelbard, 1969). The results of the process are computed for each sampling so that, after a reasonable number of samples, a probability distribution of results can be defined. Thus, a suitable number of samples needs to be selected in order to obtain stable and reliable results in a reasonable elaboration time.

A flexible statistical tool for Monte Carlo simulation was implemented and integrated with the interpretation algorithm. The methodology provides the possibility of taking several sources of uncertainties into account, Statistical distributions and random sampling can be associated not only to the input log, but also to the main model parameters such as: conductivity model parameters, fluid density and water resistivity as a function of salinity.

Previous studies (Verga et al., 2001) demonstrated that the uncertainties associated to log measurements can be represented by different kinds of statistical distributions. The implemented methodology provides the possibility to associate a different statistical distribution to each source of uncertainty: normal, lognormal, uniform and triangular statistical distributions were considered in this study.

The steps of the algorithm can be summarized as follows.

1. **Associate an uncertainty distribution to the log measures and/or to the model parameters:**
2. **Generate a set \( S \) of random configurations \( s_i = \log \) measures \( \cup \) model parameters
3. **Determine \( V, \phi \) and \( S_w \) distributions by solving the inverse problem associated for each configuration \( s_i \in S \)
4. **Compute mean and confidence al \([p_{10}, p_{90}]\) of the resulting porosity, \( \phi \) and water saturation, \( S_w \), distributions**

Sensitivity analyses showed that a limited number of random samples (approximately 100) is sufficient to reach stable values of the statistical parameters associated to each result.

RESULTS

Application to a real case: The previously described methodology was applied to a real case (Viberti et al., 2007) represented by a deep-water exploration well intercepting a shaly sand oil bearing formation. The application has to be considered as an example used to show the potential of the methodology and should not be taken as an exhaustive demonstration of its effectiveness. No laboratory tests were available to characterize the conductivity equations parameters. An interpretation with a commercial software was performed and the values of model parameters were imposed by assuming a petrophysical analogy with the some nearby fields. However, the information available was different for each of the nearby fields even though the lithology was the same; therefore, there wasn’t a set of parameters that could be used as a reference but rather a range for each parameter. This made even more necessary the use of an approach capable of handling uncertainties. The interpretation model is constituted by two solid components: sand (Vsand) and clay (Vclay). The volume of clay is represented mainly by a mixture of illite and montmorillonite. Mud invasion phenomena were taken into account by splitting the fluids (oil and water) in the flushed zone and in the undisturbed zone and considering the invasion factor \( \alpha \) showed in Eq. 1.

The measured curves available for the interpretation were the density, neutron, GR and sonic logs along with a deep resistivity log assumed to be representative of the true resistivity of the formation. The Indonesia equation was adopted as conductivity model.

Initially, a base case interpretation scenario was performed assuming Archie parameters \( n = 2 \) and \( m = 2 \).
The interpretation obtained for the base case scenario, expressed in terms of formation component fractional volumes and reconstructed logs versus measured logs, is shown in Fig. 2.

The evaluation of the uncertainty associated to porosity and water saturation was performed adopting the Monte Carlo method coupled with the inversion methodology while assuming the three simulation scenarios summarized in Table 1.

The uncertainties associated to GR, Neutron, Sonic and Density log measurements were represented by normal statistical distributions where the measured value represents the mean value and the standard deviation is equal to the instrumental error. Based on previous studies (Verga et al., 2001), the uncertainty associated to the resistivity log was represented by a lognormal statistical distribution. The instrumental error associated to each log type is given in Table 2. The uncertainty associated to the conductivity equation parameters was represented by a uniform statistical distribution having a standard deviation of 10% of the mean values (Table 3).

Table 1: Simulation scenarios for uncertainty evaluation

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Uncertain parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Input Log</td>
</tr>
<tr>
<td>2</td>
<td>m, n</td>
</tr>
<tr>
<td>3</td>
<td>Input Log, m, n</td>
</tr>
</tbody>
</table>

Table 2: Instrumental errors

<table>
<thead>
<tr>
<th>Log</th>
<th>GR</th>
<th>Neutron</th>
<th>Resistivity</th>
<th>Sonic</th>
<th>Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instrumental error</td>
<td>5%</td>
<td>7%</td>
<td>10%</td>
<td>7%</td>
<td>0.015</td>
</tr>
<tr>
<td>Statistical distribution</td>
<td>Normal</td>
<td>Normal</td>
<td>Log normal</td>
<td>Normal</td>
<td>Normal</td>
</tr>
</tbody>
</table>

Table 3: Uncertainty associated to conductivity equation parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Standard deviation</th>
<th>Statistical distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>10%</td>
<td>Uniform</td>
</tr>
<tr>
<td>n</td>
<td>10%</td>
<td>Uniform</td>
</tr>
</tbody>
</table>

DISCUSSION

Preliminary analyses of the results provided by the application of the Monte Carlo method showed that the statistical distribution associated to porosity and water saturation assumes different shapes, often not symmetrical, at different depths. Therefore, in order to provide a detailed uncertainty characterization, the results of the statistical approach were analyzed in terms of four statistical parameters: median, standard deviation and percentiles 10 and 90. Figure 3-5 show the results obtained for simulation scenarios 1, 2 and 3, respectively.

The uncertainties associated to the input logs (scenario 1) have an impact on both porosity and water saturation (Fig. 3). The uncertainty associated to porosity is quite constant and ranges from 1-2 porosity units for all the porosity values. If expressed in terms of percentages, these values correspond to an uncertainty varying from approximately 2-15% of the calculated porosity value. On average, the uncertainty associated to water saturation is 3% of the deterministic value and, as expected, progressively reduces when water saturation approaches unity.

The uncertainties associated with the conductivity equation parameters m and n, (scenario 2) have a negligible impact on porosity and a relatively strong impact on water saturation (Fig. 4). The uncertainty associated to porosity has an average value of about 0.3 porosity units and therefore, it is in accord with the results obtained from the sensitivity analyses previously discussed. The uncertainty associated to water saturation and expressed in terms of percentage of the calculated water saturation, approximately ranges from 50% for low water saturation to 10% for high water saturation.

In scenario 3, the impact of the uncertainty associated to both input log and conductivity equation parameters on porosity and water saturation was evaluated (Fig. 5). The uncertainty associated to porosity is quite constant and ranges from 1-2 porosity units, which show no significant differences with respect to the results obtained for scenario 1. The uncertainty associated to water saturation and expressed in terms of percentage of the calculated water saturation, approximately ranges from 48% for low water saturation to 10% for high water saturation.

In scenario 3, the impact of the uncertainty associated to both input log and conductivity equation parameters on porosity and water saturation was evaluated (Fig. 5). The uncertainty associated to porosity is quite constant and ranges from 1-2 porosity units, which show no significant differences with respect to the results obtained for scenario 1. The uncertainty associated to water saturation and expressed in terms of percentage of the calculated water saturation, approximately ranges from 48% for low water saturation to 10% for high water saturation, which show no significant differences with respect to the results obtained for scenario 2. Therefore, in the considered case, the combination of two sources of uncertainty, i.e., m, n and input logs did not introduced any additional uncertainty in the water saturation estimation with respect to scenario 2.
CONCLUSION

The goal of Formation Evaluation is to provide the information necessary for reserves evaluation and thus for defining the most suitable technical and economic strategies for reservoir exploitation. Any source of errors affecting the results obtained from the log interpretation process should be identified and possibly removed to ensure that reliable petrophysical parameters are obtained, especially during the exploration and appraisal phases of a reservoir.

The present study describes the details of the mathematical background behind the developed methodology of coupling the robust constrained optimization process and the Monte Carlo approach.

When the probabilistic approach is adopted, not only the uncertainty associated to the rock petrophysical characteristics are quantified, but also the most critical sources of uncertainty affecting porosity and water saturation can be identified for a given formation. Therefore, a consistent reduction of the uncertainty associated with the reservoir petrophysical parameters can be expected upon acquisition of specific additional information.

The Monte Carlo method is fully integrated into the solution algorithm to provide the uncertainty affecting porosity and water saturation based on the uncertainty associated to any input variable, log measurements and model parameters. The developed approach was applied to an offshore exploration oil well where a poor characterization of the shaly sand formation complicated the interpretation.

Results showed that the uncertainty on porosity was mainly due to errors potentially affecting log measurements and not particularly critical. On the other hand, the uncertainty on water saturation ranges approximately from 10-50%. This is mainly due to the uncertainty affecting the parameters in the water saturation model. It is quite evident that the large uncertainty of water saturation could make the difference on the economic reliability of a subsequent development project. Therefore, laboratory analyses to determine the model parameters and a new, more reliable interpretation would be recommended before any decision is taken on the field.

ACKNOWLEDGEMENT

The researcher would like to express their sincere gratitude to Prof. Francesca Verga, Cristina Serazio and Eloisa Salina Borello for their valuable contribution to the development of this research.
REFERENCES
