GEOPHYSICS, VOL. 77, NO. 3 (MAY-JUNE 2012); P. WA45–WA63, 20 FIGS., 1 TABLE.
10.1190/GEO2011-0272.1

Quantitative log interpretation and uncertainty propagation of petrophysical properties and facies classification from rock-physics modeling and formation evaluation analysis

Dario Grana[1], Marco Pirrone[2], and Tapan Mukerji[3]

ABSTRACT

Formation evaluation analysis, rock-physics models, and log-facies classification are powerful tools to link the physical properties measured at wells with petrophysical, elastic, and seismic properties. However, this link can be affected by several sources of uncertainty. We proposed a complete statistical workflow for obtaining petrophysical properties at the well location and the corresponding log-facies classification. This methodology is based on traditional formation evaluation models and cluster analysis techniques, but it introduces a full Monte Carlo approach to account for uncertainty evaluation. The workflow includes rock-physics models in log-facies classification to preserve the link between petrophysical properties, elastic properties, and facies. The use of rock-physics model predictions guarantees obtaining a consistent set of well-log data that can be used both to calibrate the usual physical models used in seismic reservoir characterization and to condition reservoir models. The final output is the set of petrophysical curves with the associated uncertainty, the profile of the facies probabilities, and the entropy, or degree of confusion, related to the most probable facies profile. The full statistical approach allows us to propagate the uncertainty from data measured at the well location to the estimated petrophysical curves and facies profiles. We applied the proposed methodology to two different well-log studies to determine its applicability, the advantages of the new integrated approach, and the value of uncertainty analysis.

INTRODUCTION

In reservoir modeling and seismic-reservoir characterization, most of the physical models used are calibrated at the well location. However, well-log data do not provide direct measurements of the reservoir properties we account for in reservoir models; the desired reservoir properties are generally obtained from the measured well logs through formation-evaluation analysis. This process contains several sources of uncertainty. Similarly, log-facies classification (LFC) can be severely affected by the uncertainty of petrophysical curves performed in formation evaluation. Furthermore, in some cases, the classification could be not linked to seismically derived attributes if rock-physics modeling (RPM) results are not included in the classification methodology.

The first goal of this work is to present a new methodology for LFC based on petrophysical and acoustic/elastic properties to link log-facies to seismic inverted attributes. The second added value of the workflow is the introduction of Monte Carlo simulations to generate several realizations of petrophysical and elastic curves to obtain different log-facies profiles, which are used to infer facies probabilities and facies uncertainty.

The first step of the methodology is formation evaluation analysis. The petrophysical evaluation of subsurface formations requires the combined efforts of log measurements and core data together with a quantitative log interpretation (QLI) model. The main results obtained from the petrophysical interpretation of well logs are the volumes of certain formation components (solid matrix and fluids) at each data level that combine the measurements provided by several tools, such as well-log resistivity, acoustic, density, neutron, nuclear magnetic resonance, fluid sampling, coring, and imaging (for details we refer the reader to Darling [2005] and Ellis and Singer [2007]). The standard approach consists of simultaneously

Manuscript received by the Editor 25 July 2011; revised manuscript received 23 November 2011; published online 30 April 2012.
[1]Formerly Eni E&P; presently Stanford University, Stanford, California, USA. E-mail: dgrana@stanford.edu.
[2]Eni E&P, San Donato Milanese, Milan, Italy. E-mail: marco.pirrone@eni.com.
[3]Stanford University, Department of Energy Resource Engineering, Stanford, California, USA. E-mail: mukerji@stanford.edu.

optimizing equations described by one or more interpretation models. Formation evaluation is done by solving the so-called inverse problem, in which well-log measurements and response parameters are used together to compute volumetric results. An improved approach was introduced recently by Heidari et al. (2010). However, there always is some uncertainty in the inferred parameters obtained by the inversion process because errors in any of the measurements lead to errors in the final values (see Fylling, 2002; Verga et al., 2002; Kennedy et al., 2010; Viberti, 2010). Moreover, other sources of uncertainties in the determination of the volume fractions come from the theoretical models used in the interpretation and the input parameters characterizing them. Hence, a statistical QLI is performed to obtain reliable results and to assess the uncertainty for the characterization.

RPM represents the second step. Rock physics is mainly used to establish the link between petrophysical parameters and acoustic/elastic properties. In our approach, the uncertainty from formation evaluation is propagated through the well-calibrated rock-physics model (RPM) to provide a full probabilistic petro-elastic model. This paper only covers the application of the basic models used in rock physics and the statistical approach. For those who are interested in in-depth knowledge of rock physics, we refer to Bourbie et al. (1987), Nur and Wang (1989), Wang and Nur (1992, 2000), Avseth et al. (2005) and Mavko et al. (2009). Statistical rock physics was introduced by Mukerji et al. (2001a) and it is essentially based on Monte Carlo simulations. An extensive explanation is included in Doyen (2007).

Finally, an LFC is performed at each depth location of the well. Reservoir facies are usually defined from sedimentological information and core analysis. This classification must then be consistent with the one performed on well-log data to link petrophysical properties with reservoir facies. Traditionally, LFC is based on multivariate techniques (in particular, cluster analysis; see Kaufman and Rousseeuw, 1990) introduced to automatically identify common features within well-log data and computed curves (generally petrophysical properties). In this work, we also account for elastic properties to guarantee a clear discriminability in the petro-elastic domain, through RPM. This issue is of key importance for the following step of seismic facies classification performed on seismic data to obtain reliable reservoir models. In this paper, we do not focus on the classification of the seismic data but on the initial well-log analysis and LFC.

The probabilistic petro-elastic model yields a statistical LFC. Thus the probability of occurrence of different facies at the well is obtained. The concept of entropy (Shannon, 1948) is then exploited to quantify the uncertainty related to the discrete random variable describing the facies distribution along the well.

At each step of the workflow, namely formation evaluation analysis, RPM, and LFC, we account for the sources of uncertainty associated with measured data, numerical models, and natural heterogeneity to provide for each estimated curve, the corresponding confidence interval at each depth location along the well.

The crucial added value is the assessment and propagation of the uncertainty at each step to finally provide a reliable classification fully based on the amount of information characterizing the entire workflow. In general, implementations of formation evaluation, RPMs, and log-facies analysis commonly neglect the uncertainty involved in these processes and can thus lead to an erroneous final interpretation. This is the pitfall the probabilistic petro-elastic classification tries to avoid by quantifying the reliability of each step and of the final result. Moreover, the discriminated facies are automatically characterized from the petrophysical and acoustic/elastic point of view, which is a key factor in data integration in seismic reservoir characterization. Eidsvik et al. (2004) presented a probabilistic log-facies formulation using hidden Markov models to get the posterior probability distributions of the log-facies while accounting for uncertainties in the log data and RPM, but did not account for the uncertainties in the formation evaluation analysis.

In contrast to previously published works in formation evaluation analysis and LFC, the uncertainty propagation problem is addressed by using a full Monte Carlo approach for each step of the method. Monte Carlo methods allow us to numerically estimate the posterior probability density functions (PDFs) of the variables we are interested in: in formation evaluation we estimate the posterior PDFs by propagating, point by point, the uncertainty from acquired well logs; in RPM we estimate the posterior PDFs by propagating the uncertainty from the previously obtained petrophysical curves; finally both PDFs are combined in LFC where we sample from these distributions and combine the results to statistically classify the facies, point by point, and estimate the associated uncertainty. The so-obtained uncertainty represents the uncertainty associated with the degree of approximation of the physical models used in the method and the uncertainty related to error measurements. However, random errors are also taken into account to represent the natural variability and heterogeneity of the rocks.

The whole methodology is first illustrated by applying all the steps to the same well-log data set of an offshore West Africa field. We then perform the same study on another real case, located in the North Sea, where a complex sedimentological facies classification has been identified.

METHODOLOGY

The methodology can be divided into three parts:

- QLI;
- RPM computation;
- LFC.

The flowchart of the methodology is shown in Figure 1. In this specific case, it refers to a clastic environment, but it can be applied to different scenarios if suitable log interpretation and RPMs are available.

In the following, we analyze in detail these three steps and the uncertainty propagation through them.

QUANTITATIVE LOG INTERPRETATION

The petrophysical characterization of subsurface formations is an inverse problem and generally requires an integrated approach. Well-log measurements and core data are combined together with a QLI model. QLI, also known as formation evaluation, provides the volume fractions of interest (solid matrix and fluids) at each depth level by solving the corresponding inverse problem, in which well-log measurements and response parameters are reconciled with the theoretical interpretation models. In particular, the inverse problem is based on an opportune cost function that expresses the distance between the observed measurements and the predictions of the model chosen to describe the system. The final aim is to minimize the cost function and determine the solution (i.e., the volume

fractions of the formation components) that optimally mimics the observables.

Theoretical response-equations in QLI come from a simplified description of the physics involved in the well-log instruments and are considered together with several constraints. These are typically used to impose geological and/or petrophysical prior knowledge to avoid physically meaningless results and define the domain space of the problem. In details, QLI can be divided into three different phases:

- system parameterization, which defines the set of parameters (unknown volumes) which characterizes the formation;
- direct modeling, which involves the laws (linear and/or nonlinear) able to generate the synthetic values for the observables, once the parameters describing the model are fixed;
- inverse modeling, which plays its role in volumetric quantification.

System parameterization and direct modeling are quite straightforward; even so, it is worth mentioning a few points of discussion that need to be taken into account to properly define the domain space.

Mathematically, at a given depth x, the vector $\mathbf{d}(x)$ represents the observational log data (e.g., gamma ray, neutron, density, etc.). QLI model is determined by the components of a vector $\mathbf{p}(x)$ containing the depth-dependent unknown volumetric fractions, such as porosity ϕ, volume of clay v_{clay}, water saturation s_w, and others. For simplification purposes, from now on the dependency upon the spatial position x is disregarded and it is assumed that all the recorded log measurements are sampled with a constant depth step. Formally, we can write

$$\mathbf{p} = (\phi, v_{\text{clay}}, s_w, \ldots) \tag{1}$$

Once a well is drilled into the formation, some mud fluid is pumped into the borehole to maintain the pressure balance. Often, drilling is done at an overbalanced condition and the mud has a pressure slightly higher than the formation pore pressure. This pressure gradient induces mud filtrate to seep into the porous rock system. Hence, well-log measurements are affected, at least, by the existing step invasion profile due to the circulating fluid. The mud invasion divides the formation into two separated parts: the so-called flushed zone (usually labeled *XO*) and the unflushed zone, or deep zone (*DE*). This splits the set of unknowns into two subsets, and the vector $\mathbf{p}$ can be written as

$$\mathbf{p} = (\mathbf{p}_{XO}, \mathbf{p}_{DE}) \tag{2}$$

Shallow reading-log measurements are assumed to respond only to volumes of formation components in the first zone (vector $\mathbf{p}_{XO}$). Similarly, deep reading-logs only see the unflushed zone (vector $\mathbf{p}_{DE}$). Some tools, which have a medium depth of investigation, are assumed to be influenced by both zones, and their response-equations contain terms for all formation components, regardless of the zone (see Appendix A for details).

On the other hand, physically based assumptions constrain and reduce the domain space of the unknowns. For instance, the standard normalized-to-unity investigated total volume fractions (in the shallow and deep zones) need to be achieved. Moreover, it is common to assume lateral continuity: All the solid formation components extend infinitely from the borehole at 0° dip. Porosity in the flushed and undisturbed zone is the same, regardless of the type of fluids filling the pore space. So, the sum of the fluid volumes in the flushed zone is equal to the sum of fluid volumes in the unflushed zone. In a more compact form, all of the constraints can be collected in the function $u(\mathbf{p})$.

Synthetic log data $\mathbf{s}$ are obtained by applying the direct modeling operator $\mathbf{f}_{\text{QLI}}(\mathbf{p})$ for a given fixed values of the model parameters $\mathbf{p}$

$$\mathbf{s} = \mathbf{f}_{\text{QLI}}(\mathbf{p}). \tag{3}$$

The forward model is a set of equations, linear or nonlinear, which describes the link between volumetric curves ($\mathbf{p}$) and petrophysical properties measurable at the well ($\mathbf{d}$). The most common equations are summarized in Appendix A.

Inverse modeling tries to determine the components of the parameter vector $\mathbf{p}$, using the observational data $\mathbf{d}$. The optimization problem involves a cost function to be minimized to reduce the discrepancy $\mathbf{e}$ between the observed data and the synthetic data generated by the direct modeling

$$\mathbf{e} = \mathbf{s} - \mathbf{d} = \mathbf{f}_{\text{QLI}}(\mathbf{p}) - \mathbf{d}. \tag{4}$$

The definition of the cost function introduces the concept of a weighted two-norm for the vector $\mathbf{e}$. As a matter of fact, because the physical observables and measurements come from several well-log tools and involve absolute values that are in general very different (and have different units), we have to normalize the two-norm to treat all the quantities at the same level. This normalization mainly depends on typical tool accuracy. Moreover, an additional term based on log analyst experience, a multiplier, is incorporated to weight the different tools with respect to their influence in the final answer. The understanding of how a tool can affect the results and the knowledge of tool physics is required to select the multiplier values. Finally, normalizations and weight multiplier factors can be organized into a weighting matrix $\mathbf{W}$. So, the cost function reads

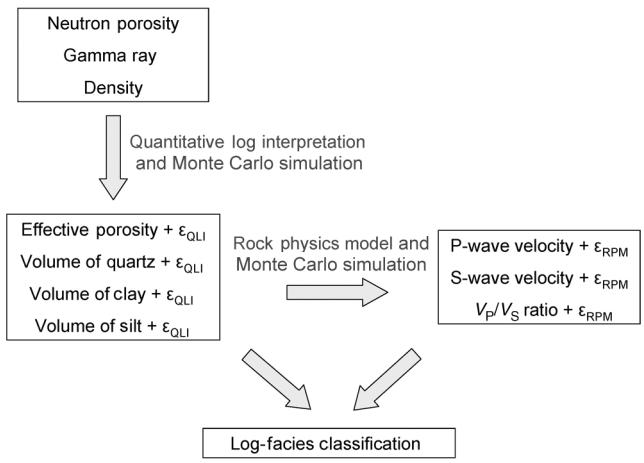

Figure 1. Flowchart of the methodology.

$$j(\mathbf{p}) = \left(\frac{1}{2}\mathbf{e}^T\mathbf{W}\mathbf{e}\right)_{|u(\mathbf{p})}, \quad (5)$$

which is a scalar and it can be represented as a function of the unknowns collected in the vector $\mathbf{p}$. The aim of optimization is to determine the optimal solution $\mathbf{p}^*$ that minimizes the above cost function, under the set of constraints $u(\mathbf{p})$ representing the solution space for the problem. The so-obtained vector $\mathbf{p}^*$ represents the output and solution of the standard QLI.

We now set the scenario regarding the uncertainty analysis and evaluation for the QLI methodology. There always is some degree of uncertainty in the inferred volumes $\mathbf{p}^*$ obtained by the inversion process because each piece of information introduced in system parameterization and direct modeling can be affected by several errors. Quantification and propagation of these uncertainties throughout the inversion problem is considered and assessed here with appropriate methods.

The main sources of uncertainty in QLI come from well-log measurements, heterogeneity of rock systems, possible thin layered intervals, invasion effects, lithological and/or textural assumptions, environmental correction issues, simplified interpretation models, and input parameters. A complete analysis of the above uncertainties would involve a full knowledge of the individual contributions and such detailed preliminary evaluation is rarely, if ever, available. Thus, the practical quantification problem simplifies the classification of sources into three macrogroups: methodological, systematic and random errors (Theys, 1991, 1994, 1997).

Methodological errors are mainly due to an incorrect choice of interpretation models and related parameters. In general, model selection is one of the most critical steps in a reliable QLI and, for sure, it is mandatory to keep in mind all of the simplifying hypotheses of the real system it intends to describe. Because an inaccurate model selection could introduce significant errors to the interpretation, core measurement calibration and validation are fundamental at this particular stage. Petrophysical experience plays a role as well. Potential errors associated with an incorrect choice of the theoretical models are difficult to quantify and their magnitude may vary significantly. This is to state that methodological errors, once assessed, should be mainly used as a sort of guidelines on the applicability limits and constraints underlying each model.

Systematic errors are defined as reproducible inaccuracies of the measurement mainly due to imprecision of the instrumental system, processing of the data, environmental conditions, and so on. Because systematic errors cannot be removed by repeated runs, they need to be recognized and corrected before any calculation.

Random errors are those that need to be accounted for and propagated throughout the QLI process. Random errors are mostly associated with the physics of the well-log measurement system and cannot be corrected because they cannot be reproduced. Statistical variations in count rates or signal noise are examples of random errors. Also, the uncertainty associated with the parameters used to correct the measurements for any environmental effect has an important impact in the final petrophysical characterization and must be accounted for.

Given the broad spectrum of random uncertainty, a statistical treatment of the problem is suitable. Despite this apparent need, most formation evaluations in practice are based on a single deterministic description that can reproduce the well observations with a certain quality and confidence. The single deterministic model and its description make the adopted model rather unsuitable for uncertainty assessment.

On the other hand, the statistical approach combines the deterministic QLI with a Monte Carlo simulation to build a probabilistic framework to study the natural variability of the results (Figure 2). To avoid confusion, in what follows deterministic QLI means with no Monte Carlo simulation implemented.

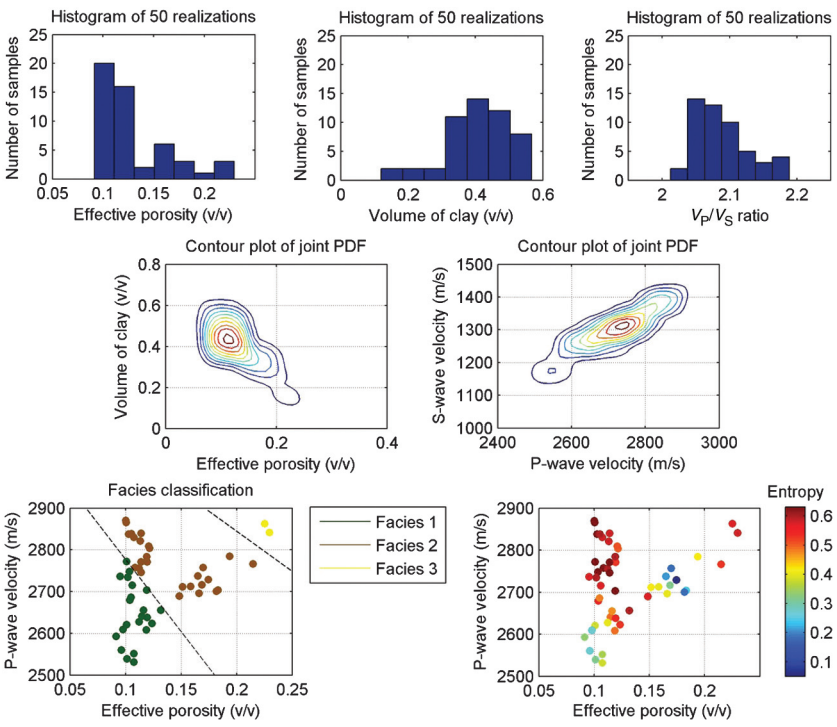

Figure 2. Schematic representation of petroelastic uncertainty estimation and LFC through Monte Carlo simulations of petrophysical and elastic properties.

We start from uncertain inputs representing log measurements. Uncertainty in each input variable is represented by a probability density function (PDF) whose range of values correspond to the nominal uncertainty provided with the measurement. In particular, Gaussian distributions are used to represent the random variable $\tilde{\mathbf{d}} \sim \mathcal{N}(\mathbf{d}, \boldsymbol{\Sigma}_\mathbf{d})$ and describe the uncertainty of acquired data. The measured value $\mathbf{d}$ represents the mean, the standard deviation is equal to the instrumental/processing error (collected in the covariance matrix $\boldsymbol{\Sigma}_\mathbf{d}$), and $\mathcal{N}$ is a compact notation for the normal distribution.

This set of distributions, estimated at each depth location along the well profile, provides the bandwidth of tolerance for the response variables of the forward model.

Model uncertainties can be added to the forward modeling $\mathbf{f}_{\mathrm{QLI}}$ by means of a random noise $\boldsymbol{\varepsilon}_{\mathrm{QLI}}$, distributed with zero mean and a prior covariance matrix, so that the forward model becomes

$$\mathbf{s} = \mathbf{f}_{\mathrm{QLI}}(\mathbf{p}) + \boldsymbol{\varepsilon}_{\mathrm{QLI}}. \tag{6}$$

In most of the cases, the forward model can be represented as

$$\mathbf{s} = [gr, nt, dn, rs] = \mathbf{f}_{\mathrm{QLI}}(\phi, v_{\min}, s_w) + \boldsymbol{\varepsilon}_{\mathrm{QLI}}, \tag{7}$$

where ϕ is porosity, $v_{\min}$ is the vector of the volumes of mineral fractions, s_w is water saturation, gr is gamma ray, nt is neutron porosity, dn is the acquired density, and rs is resistivity. It allows computing the volumetric fractions (porosity, volume fractions of mineral, and saturations) from the acquired logs (see Appendix A for details). In our approach, we also include velocity data to guarantee consistency with the RPM step that follows.

The simulation is performed by randomly sampling the values for each input log required for the interpretation model from the PDFs of $\tilde{\mathbf{d}}$ at each data level (Viberti, 2010). Suitable vertical correlated and conditioned Monte Carlo sampling can be implemented in a straightforward way to perform a geologically driven sampling and minimize the probability of the occurrence of physically meaningless scenarios.

Once a set of values has been sampled, the QLI cost function is optimized and the results are stored. When a statistically representative number of realizations have been drawn, the results can be sorted and histograms created to approximate the local PDFs of the uncertain output $\tilde{\mathbf{p}}$

$$\tilde{\mathbf{p}} = \mathbf{F}(\tilde{\mathbf{d}}), \qquad \mathbf{F} = (\mathbf{f}_{\mathrm{QLI}} + \boldsymbol{\varepsilon}_{\mathrm{QLI}})^{-1}. \tag{8}$$

Thus, the single vector $\mathbf{p}^*$ obtained in the deterministic QLI is now enhanced to account for the uncertainty in $\tilde{\mathbf{p}}$.

Once this goal is achieved, all inferences can be obtained from the posterior PDFs by computing statistics relative to individual parameters. Because the final result is presented as PDFs, this tool provides a unified framework for volume estimates and for the uncertainty associated with them.

ROCK-PHYSICS MODELING

An RPM is represented by a set of equations that transform petrophysical variables into acoustic/elastic variables. RPM can be a simple regression on well data or a more complex physical model characterized by parameters that need to be estimated, such as elastic moduli of matrix and fluid components, critical porosity, aspect ratio, and/or coordination number (Mavko et al., 2009).

Generally, the standard workflow starts from the model calibration from well-log data, or the estimation of the physical parameters involved, to obtain a good match between RPM predictions and well-log measurements. The model type depends on the geologic environment and the parameter estimation is given by the solution of an inverse problem.

In the traditional rock physics workflow, the generic RPM can be written in the concise form

$$\mathbf{r} = \mathbf{f}_{\mathrm{RPM}}(\mathbf{p}), \tag{9}$$

where the vector $\mathbf{p}$ represents the petrophysical input obtained from QLI and $\mathbf{f}_{\mathrm{RPM}}$ comprises the set of equations defining the model. The output vector $\mathbf{r}$ is the vector containing acoustic and elastic properties, typically P-wave velocity (V_{P}), S-wave velocity (V_{S}), and density (ρ)

$$\mathbf{r} = (V_{\mathrm{P}}, V_{\mathrm{S}}, \rho). \tag{10}$$

In literature, there are many different RPMs and they can be classified in three big classes (Avseth et al., 2005; Mavko et al., 2009):

- Empirical models: In this case, RPM usually is a simple regression using elastic moduli or directly velocities;
- Granular media models: These are based on Hertz-Mindlin contact theory, which assumes that the rock is represented by a random pack of spherical grains;
- Inclusions models: These models describe the rocks as a sequence of inclusions (typically ellipsoidal) until the desired pore fraction is achieved.

The various models differ for the calculation of the dry rock properties. The usual scheme of a RPM starts from the computation of fluid and solid phase (matrix) properties. Then, dry rock properties, i.e., the properties of the solid phase with its own porosity, are calculated with equations depending on the RPM chosen. Finally, the effect of fluids using Gassmann equation is included.

In general, once the RPM has been calibrated to well-log data, to evaluate the fluid effect, different saturation scenarios could be considered: a "brine" scenario, where the hydrocarbon column in the reservoir is substituted by water, and a "full hydrocarbon" scenario, where hydrocarbon saturation is constantly increased up to one-minus irreducible water saturation.

As described, our aim is to assess the uncertainty in RPM by propagating it from QLI probabilistic output. From the previous QLI analysis, we have realizations of the petrophysical variables representing basic volumetric properties of the formation: Uncertainty is represented by the obtained PDFs of $\tilde{\mathbf{p}}$. Thus, $\mathbf{f}_{\mathrm{RPM}}$ links the uncertain QLI inputs to multiple RPM output variables.

In the end, the output PDFs of elastic properties $\tilde{\mathbf{r}}$ are generated through a Monte Carlo simulation (Figure 2, top), extending the deterministic RPM to the probabilistic case

$$\tilde{\mathbf{r}} = \mathbf{G}(\tilde{\mathbf{p}}), \qquad \mathbf{G} = \mathbf{f}_{\mathrm{RPM}} + \boldsymbol{\varepsilon}_{\mathrm{RPM}}. \tag{11}$$

This means that the random errors of the first step (QLI) are propagated through RPM. To account for the uncertainty associated with the RPM approximation, a random noise $\boldsymbol{\varepsilon}_{\mathrm{RPM}}$ can be added to the $\mathbf{f}_{\mathrm{RPM}}$ direct modeling operator. Monte Carlo simulation still provides a simple way to propagate all uncertainties in the RPMs because the latter are in general nonlinear (except for linear empirical models).

As we previously mentioned, we will restrict our description to a clastic reservoir, without losing generality. For high-porosity clastic reservoirs, common RPMs are the so-called granular media models (see Appendix B and Mavko et al., 2009). These models are based on Hertz-Mindlin contact theory (Dvorkin et al., 1994; Dvorkin and Nur, 1996; Gal et al., 1998), which describes a rock as a random pack of spherical grains. The model equations are described in detail in the Appendix B. In the reservoir layer, in addition to input petrophysical curves, i.e., effective porosity, volumes of mineral components, and saturations, the applied RPM requires the following input data: fluid and matrix properties, reservoir pressure, and

temperature. Fluid parameters and reservoir condition information are generally available from PVT analysis and well tests measurements, whereas solid-phase properties are inferred from core mineralogical analysis. In reservoir characterized by low-porosity rocks, inclusion models (Avseth et al., 2005; Mavko et al., 2009) are more appropriate.

LOG-FACIES CLASSIFICATION

The goal of this section is to present a new methodology for LFC and the related uncertainty evaluation. The classification of facies at well location is a key important step in reservoir modeling: the static reservoir model essentially consists of stochastic simulations, which are driven by LFC at the well locations.

We propose here a classification based on petrophysical and acoustic/elastic properties to link log-facies to seismic inverted attributes that are often used as soft conditioning data in reservoir simulation. However, different sources of uncertainty affect high resolution LFC. We have classified the associated uncertainty in two main groups: the uncertainty related to petrophysical curves obtained in QLI and the uncertainty associated with elastic properties (velocities or impedances) recovered by RPM. We finally couple the so-obtained results by Monte Carlo simulations to generate several realizations of log-facies profiles that are used to infer facies uncertainty from the probabilistic analysis previously performed (Figure 2, bottom).

Here, we first present the standard approach to set the basis of the statistical methodology. LFC usually combines a multivariate statistics technique (cluster analysis) and interpretation of prior sedimentological information. Cluster analysis is a well-known technique that helps to group objects according to their mutual similarity. Petrophysical and elastic logs can be statistically processed to find clusters, by using, for example, an unsupervised hierarchical clustering algorithm. This algorithm groups the log data based on their statistical similarity into hierarchically ordered set of clusters. During the characterization phase, an initial number of clusters is selected. The identified clusters are compared to quantitative and qualitative information derived from cores (routine and special core analyses, and lithological, sedimentological, and petrographical descriptions). The integration of information from different sources (logs, core measurements, lithology, sedimentology, etc.) is a key step in the characterization of the different clusters. In particular, a suitable training set $\mathbf{t}$ is selected from QLI and RPM outputs ($\mathbf{p}^*$, and $\mathbf{r}$, respectively) and LFC is run including the data all along the well, providing a depth-dependent and discrete class vector, collecting the results of the classification. In detail, the log-facies vector gives the classified litho-classes at a given depth value. Eventually, some of the clusters are grouped to identify and classify log-facies for use in the 3D geological model. In the conventional workflow, LFC does not consider the uncertainty associated with input curves (petrophysical and acoustic/elastic ones).

To evaluate the uncertainty related to LFC, we use the Monte Carlo approach. From the previous steps of the methodology (QLI and RPM), we have obtained PDFs of petrophysical and acoustic/elastic curves ($\tilde{\mathbf{p}}$ and $\tilde{\mathbf{r}}$), and these define the probabilistic training set $\tilde{\mathbf{t}}$. In particular, from this set of PDFs collected in $\tilde{\mathbf{t}}$, N realizations can be sampled and then used to perform N LFC profiles through Monte Carlo simulations, and estimate the posterior probability on facies classification, in addition to the most probable facies profile. In other words, cluster analysis is applied to the different QLI and RPM realizations to obtain the related scenarios of log-facies. Then, from the so-generated profiles we can count, at each depth location, the frequency of the occurrence of each log-facies, and infer a depth-by-depth posterior probability distribution.

The results are collected in a vector $\mathbf{c}$ whose entries are the probability of each classified log-facies, so that if have a set of M possible discriminated facies, their probabilities of occurrence are $\mathbf{c} = (c_1, \ldots, c_M)$. Sedimentological information can be further integrated by multiplying the probability vector $\mathbf{c}$ by the confusion matrix obtained from the sedimentological classification (e.g., the matrix accounting for the degree of confusion among the different facies with those provided by sedimentologists). We finally observe that, from vector $\mathbf{c}$, for instance, the most likely facies scenario can be estimated.

We can take a further step to quantify the uncertainties associated with the probabilistic LFC result. In fact, the most probable scenario, as the output of a probabilistic workflow, should be coupled with a confidence value showing the amount of information brought by the discrete classification. For this reason, the concept of entropy (Shannon, 1948) can be exploited. This function is a statistical parameter able to provide some numerical insights on the intrinsic variability of the discrete variable studied with respect to all the outcomes of the classification.

The scalar information entropy (Shannon, 1948; Mavko and Mukerji, 1998; Mukerji et al., 2001b) associated with the probability vector $\mathbf{c}$ is defined as follows

$$h = -\mathbf{c}^T \log_M \mathbf{c} = -\sum_{i=1}^{M} c_i \log_M c_i, \tag{12}$$

where the logarithm is computed in base M (number of facies) because we have classified M different facies; this restricts and normalizes the entropy to the range [0,1].

As a simple example, assume that $M = 3$ and, at a particular point of the depth profile, the three facies are equally probable: $\mathbf{c} = (1/3, 1/3, 1/3)$. Hence, the entropy is $h = 1$, namely the maximum entropy value. This means that uncertainty is at its maximum. On the other hand, if $\mathbf{c} = (1, 0, 0)$ we get $h = 0$, giving a minimum entropy scenario and no uncertainty. In-between situations are characterized by entropy values that quantify the degree of disorder brought by the classification: $0 < h < 1$. Entropy measure of uncertainty goes beyond measures such as variance and covariance, and can be used for categorical variables (e.g. facies).

The information provided by entropy analysis on LFC, comprising uncertainty from QLI and RPM, will be discussed in two real examples.

APPLICATION: FIRST EXAMPLE

The methodology has been applied to a real data set from an exploration well (here called A), offshore West Africa. Well A is vertical and drilled in a deep offshore clastic reservoir made by soft sandstones of a turbidite channel complex. The prospect is a typical combined structural-stratigraphic trap. It is composed of a large turbidite sandstone channel that drapes over the north flank of a large salt diapir. Sparse shale beds are present and analysis of core samples reveals that the clay is a mixture of illite and kaolinite (respectively 70% and 30%). The reservoir fluid is oil. A comprehensive set of wireline logs (standard and high resolution) has been acquired to provide a reliable formation evaluation: gamma ray (*gr*),

resistivity (*rs*), sonic (*sn*), density (*dn*), neutron (*nt*), and nuclear magnetic resonance (*nmr*) logs. Globally, the quality of the acquired data is very good (Figure 3). Following our notation, the depth-dependent data vector (see also Appendix A) is

$$\mathbf{d} = (gr, rs, sn, dn, nt). \qquad (13)$$

Deterministic QLI via a straightforward $\mathbf{f}_{\mathrm{QLI}}$ resulted in the generation of mineralogical volumes (v_{sand}, v_{silt}, and v_{clay}), effective porosity ϕ, and water saturation s_w

$$\mathbf{p} = (v_{\mathrm{sand}}, v_{\mathrm{silt}}, v_{\mathrm{clay}}, \phi, s_w). \qquad (14)$$

In well A, invasion effects were negligible so that $\mathbf{p}_{XO} = \mathbf{p}_{DE}$.

In Figure 4, the results of our local minimization, $\mathbf{p}^*$, in other words the final set of petrophysical curves performed in QLI, are shown. In Figure 5, we give an example of comparison between our result and a standard commercial software output, which shows a satisfactory match between the two estimated set of data.

An RPM is then used to establish the link between petrophysical parameters $\mathbf{p}^*$ and acoustic properties $\mathbf{r}$. For this reservoir study, we used the soft sand (uncemented) model (Appendix B), based on Hertz-Mindlin contact theory (Dvorkin et al., 1996).

Petroelastic property uncertainty assessment and facies classification has been performed in the lower reservoir whose top is located at approximately 2510 m. In the reservoir layer, the applied RPM requires the following input data from QLI: effective porosity, volumes of mineral components (sand, silt, and clay), and fluid saturations. The parameters include effective pressure of 35 MPa, a critical porosity of 0.4 while the coordination number is 9. In this particular case, effective pressure has been assumed to be constant within the reservoir, but, in general, pressure effects on velocities should be taken into account. Several models have been developed to account for pressure effect on velocities (e.g., Eberhart-Phillips et al., 1989) or on dry rock elastic moduli (Macbeth, 2004); however, we point out that these methods must be calibrated (typically using lab measurements at different pressure regimes) to determine the empirical parameters. The RPM is calibrated at well location, by comparing the velocity and density estimated from the elastic parameters with the velocity and density from the recorded logs in the borehole. The calibration is performed in wet condition (meaning that a preliminary fluid substitution is performed on well logs) to avoid the effect of fluid in rock parameters calibration. Focusing on the depth interval including the lower reservoir layer (2480–2600 m), the results can be seen in Figure 6. The calibration is performed with a trial and error method; optimization techniques could be used as well, but these methods do not guarantee that the optimized parameters still preserve their

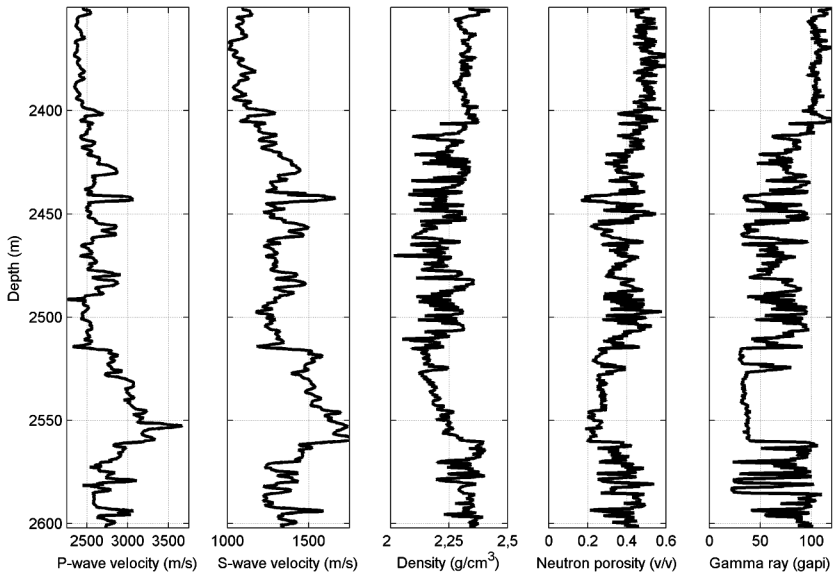

Figure 3. Well-log data set interval of well A, from left to right: P-wave and S-wave velocity, density, neutron porosity, and gamma ray.

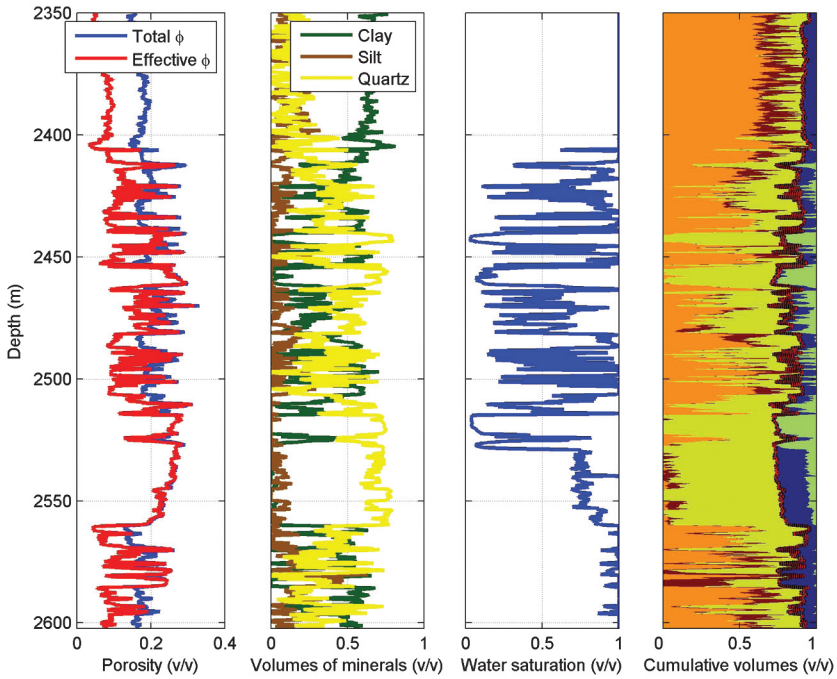

Figure 4. Petrophysical curves performed in QLI, from left to right: porosity (total porosity in blue, effective porosity in red); volumetric fraction curves (clay in green, quartz in yellow, and silt in brown); water saturation; and the cumulative volumetric display (shale in orange, silt in brown, quartz in yellow, water in blue, and oil in green; red dashed line represents one-minus porosity and separates solid and fluid phase).

physical meaning. RPM calibration, in many practical applications, can be quite complex because the physical-mathematical model cannot account for heterogeneity and natural variability of the rock. In our application, the model overpredicts sonic log velocities at 2520 m depth and does not match the high peak of P-wave velocity at 2550 m depth. The mismatch at 2520 m is due to lack of accuracy of the model, whereas the velocity peak at 2550 m could be a measurement error or represent a different lithology. The lack of accuracy of the RPM in some small depth intervals should be partially compensated by the variability introduced in the following Monte Carlo simulations.

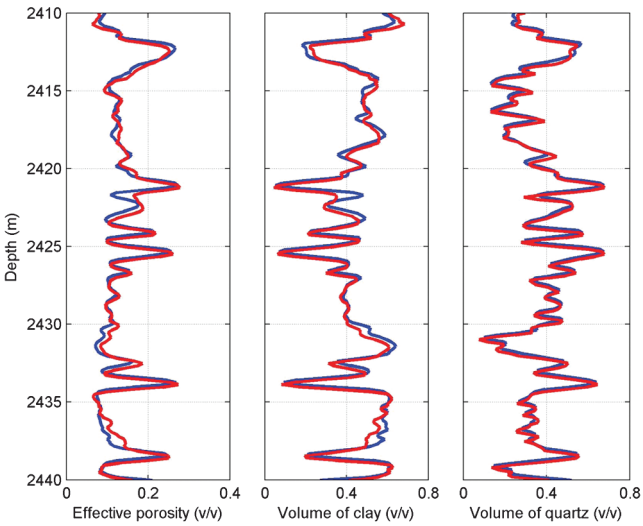

Figure 5. Petrophysical outputs comparison, from left to right: porosity, volume of clay, and volume of quartz (red curves represent our results; blue curves represent standard commercial software outputs).

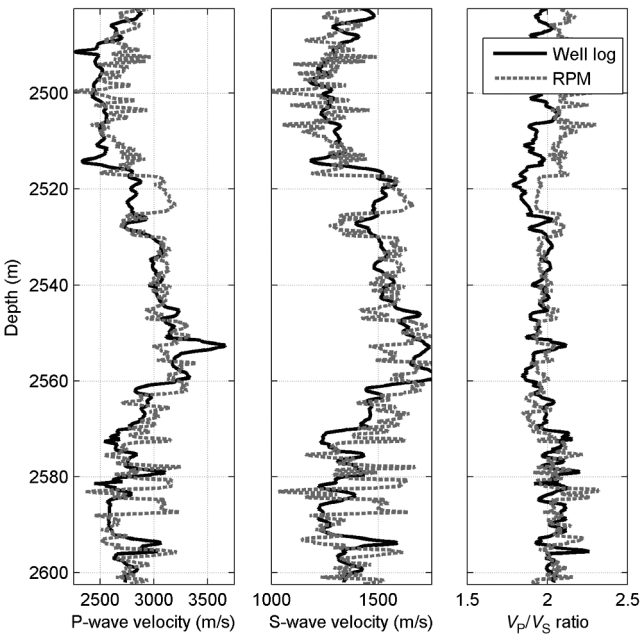

Figure 6. Calibration of the RPM, from left to right: P-wave and S-wave velocity and V_P/V_S ratio (black curves represent the actual sonic log and grey dashed curves represent the predicted RPM). The RPM has been calibrated in wet condition and applied to the well scenario by Gassmann fluid substitution.

Traditional LFC is achieved through cluster analysis, carried out using key-curves (defining the training set **t**) selected to be effective porosity, clay content, and V_P/V_S ratio

$$\mathbf{t} = (\phi, v_{\text{clay}}, V_P/V_S). \tag{15}$$

In particular, V_P/V_S ratio refers to brine condition to avoid any fluid effect as a result of hydrocarbon presence. In this case, a preliminary analysis showed that V_P/V_S ratio and clay content have a good correlation; however, the training set could include only V_P, or V_S (or both of them), if these logs are identified as good lithological indicators for the reservoir (see second example).

The training set **t** was statistically processed by means of a hierarchical agglomerative clustering algorithm with Ward's minimum variance linkage method (Ward, 1963). Hierarchical algorithms find successive clusters using previously established clusters (i.e., the resultant classification has an increasing number of nested classes). Thus, given a data set consisting of N_O objects, agglomerative clustering methods generate outputs where objects are gradually partitioned. Ward's minimum variance linkage method makes use of squared Euclidean distances to define the dissimilarity among clusters. The most common and practical way to visualize the results is through a plot called dendrogram. This allows reconstructing the merging history of the N_O objects of the studied data set from the beginning (each object forms a cluster of its own) to the end of the clustering process (all objects are in the same cluster). The dendrogram for our case application is plotted in Figure 7.

In traditional log-facies analysis, the number of the classes (facies) that can be identified from sedimentological information is often higher than the number of classes of interest in reservoir models because the quality of the data at the well location and the indirect measured data far away from the well (seismic and electromagnetic data) do not allow us to distinguish sedimentological features with the same accuracy. Using the dendrogram that shows a very stable clustering (Figure 7) and core analysis calibration, a three-facies classification, based on sand concentration, is determined (Figure 8). It consists of:

- low-concentration turbidite (LCT) facies in green;
- mid-concentration turbidite (MCT) facies in brown;
- high-concentration turbidite (HCT) facies in yellow.

The proposed facies classification is based on the percentage of quartz and clay in the facies as derived from interpreted log curves. This classification is simplified compared to sedimentological models, but is strongly consistent with the facies classification of the static reservoir model, where porosity and net-to-gross are distributed following distributions estimated from well logs in each facies. Because thin interbeddings of hemipelagic shales have been observed in the reservoir at the well location, facies LCT could be subdivided into two subclasses, namely low and very low concentration turbidite (the latter including nonreservoir shales). This behavior can be also observed in the dendrogram. However, these two classes are not distinguishable from the petro-elastic point of view within the reservoir.

Monte Carlo simulations are then exploited to provide the uncertainty propagation and evaluation. The starting point are the normal PDFs associated with the set of wireline logs ($\tilde{\mathbf{d}}$) used in QLI with standard deviations that are typical of the log measurements and the processing of raw data. Table 1 collects common standard deviations (Viberti, 2010).

In this example, a sensitivity analysis shows that 100 realizations are enough to provide a stable solution and reliable results. Realizations of input curves sampled from $\tilde{\mathbf{d}}$ provide the set of cost functions to be minimized and the related set of petrophysical curves defining $\tilde{\mathbf{p}}$. In Figure 9, we plot the different realizations and the median of the petrophysical sets of curves.

From the previous QLI analysis, we have realizations of the petrophysical variables representing basic properties of the formation, such as porosity and mineral volumetric contents. Thus, we can generate a set of realizations of elastic properties logs through a Monte Carlo simulation by applying the RPM to the set of petrophysical curves realizations from QLI. In the Monte Carlo simulation, we add a random error to account for the uncertainty associated with RPM. From the PDFs of $\tilde{\mathbf{p}}$ and the previously calibrated RPM, the set of acoustic and elastic curves and the corresponding PDFs of velocities $\tilde{\mathbf{r}}$ are then computed (Figure 10). In Figure 11, we show the full probability distributions of porosity and P-wave velocity, and, as an example, we extract the corresponding histograms at two given depth locations.

On the base of the set of 100 realizations of effective porosity, volume of clay, and V_P/V_S ratio, we can obtain 100 training sets sorted in $\tilde{\mathbf{t}}$, to be used to generate 100 profiles of facies. From the previous steps of the methodology, we have obtained 100 realizations of petrophysical curves (QLI results, Figure 9) and 100 realizations of rock physics curves (RPM results, Figure 10). We now perform 100 LFCs through Monte Carlo simulations (Figure 12, top) and estimate the posterior probability on facies classification, in addition to the most probable facies profile. In particular, cluster analysis is applied to the 100 realizations to obtain 100 profiles of log-facies; then the frequency of the facies, at each depth location of the profile, provides the posterior probability of facies occurrence and the most likely facies estimation.

In Figure 12 (bottom), we show, as an example, five realizations of facies obtained from Monte Carlo simulations: the thick high-porosity sand layer is well-identified in all the profiles, whereas the thin layers at the top and at the bottom of the reservoir are more uncertain. This means that, in the upper and lower part of the reservoir, the LFC can vary as a function of RPM and QLI input parameters in a traditional deterministic workflow and the underestimation of the uncertainty could lead to a misclassification of the facies.

From the so-generated 100 profiles, we can count, point by point along the profile, the frequency of occurrence of each log-facies and infer the posterior probability distribution $\mathbf{c}$, by normalizing, at each data level, the frequencies by the total number of realizations (100). In Figure 13, we show the probability of log-facies obtained from petroelastic Monte Carlo realizations and the final most probable classification. We notice that in the mid part of the profile the probability of having a HCT facies is very high and the uncertainty is small, whereas in the thin layers zones the uncertainty is much higher.

This also is reflected in the entropy curve that we plot together with the most likely facies profile (Figure 13) to quantify the uncertainty in the discrete probabilistic scenario. Entropy is computed in base three because we have discriminated three different facies, to

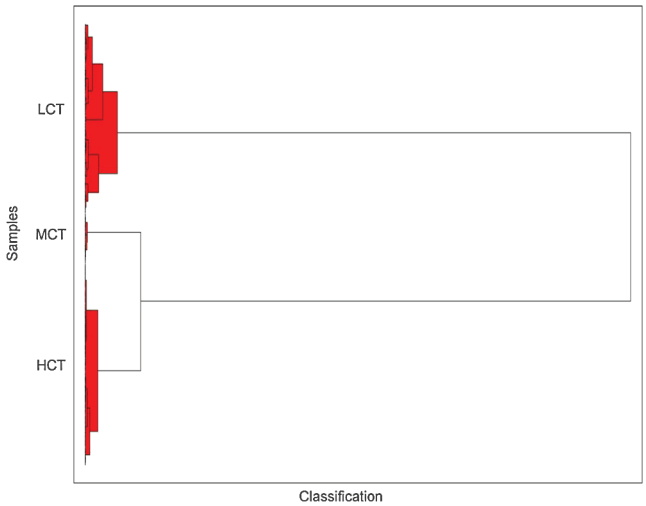

Figure 7. Dendogram associated with LFC. A dendrogram consists of many U-shaped lines connecting objects in a hierarchical tree. The stem of each U represents the distance between the two objects being connected. Red clusters refer to the connecting histories of the three recognized facies: low-concentration turbidite (LCT), mid-concentration turbidite (MCT), and high-concentration turbidite (HCT).

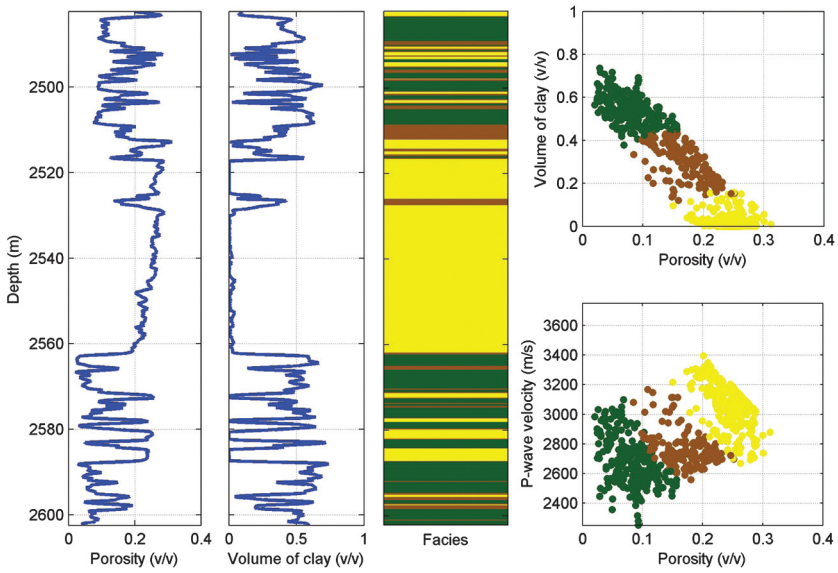

Figure 8. LFC performed at well location: LCT in green, MCT in brown, and HCT in yellow. LFC is derived by using petrophysical curves (porosity and clay content) and velocity data (V_P/V_S ratio). On the right, we show two crossplots in petrophysical (top right) and petroelastic (bottom right) domain, color-coded by facies classification.

get values between zero (no uncertainty) and one (maximum uncertainty).

To interpret entropy results, we comment here on the surface described by the variation of the probabilities related to the three discriminated facies. Because the three probabilities add up to one, we can take any two as independent variables to map the behavior of entropy for different values of the facies probabilities. Specifically,

$$h(c_{\mathrm{HCT}}, c_{\mathrm{LCT}}) = -[c_{\mathrm{HCT}} \log_3(c_{\mathrm{HCT}}) + c_{\mathrm{LCT}} \log_3(c_{\mathrm{LCT}}) + (1 - c_{\mathrm{HCT}} - c_{\mathrm{LCT}}) \times \log_3(1 - c_{\mathrm{HCT}} - c_{\mathrm{LCT}})] \quad (16)$$

with $c_{\mathrm{HCT}} + c_{\mathrm{LCT}} \leq 1$, where c_{HCT} is the probability of obtaining HCT at a given depth and c_{LCT} is the probability associated with LCT; whereas $c_{\mathrm{MCT}} = 1 - c_{\mathrm{HCT}} - c_{\mathrm{LCT}}$. As previously explained, the three probabilities are collected in the vector $\mathbf{c} = (c_{\mathrm{LCT}}, c_{\mathrm{MCT}}, c_{\mathrm{HCT}})$. When the three facies are equiprobable, we have maximum entropy, whereas when the probability of a given facies is close to 1, the entropy tends to 0. The two crossplots on the right of Figure 13 help in clarifying the information provided by entropy analysis. With Figure 8 in mind (right), it can be easily seen that the transition areas between the different facies are characterized by high entropy (yellow clouds in Figure 13) as expected. This fact mainly affects the generalized high entropy associated with MCTs. On the other hand, the extreme LCT and HCT facies show low entropy (blue clouds).

Table 1. Standard deviations associated with log measurements.

Log	Standard deviation
Neutron	7%
Density	0.015 g/cm^3
Gamma ray	5%
Resistivity	10%
Sonic	7%

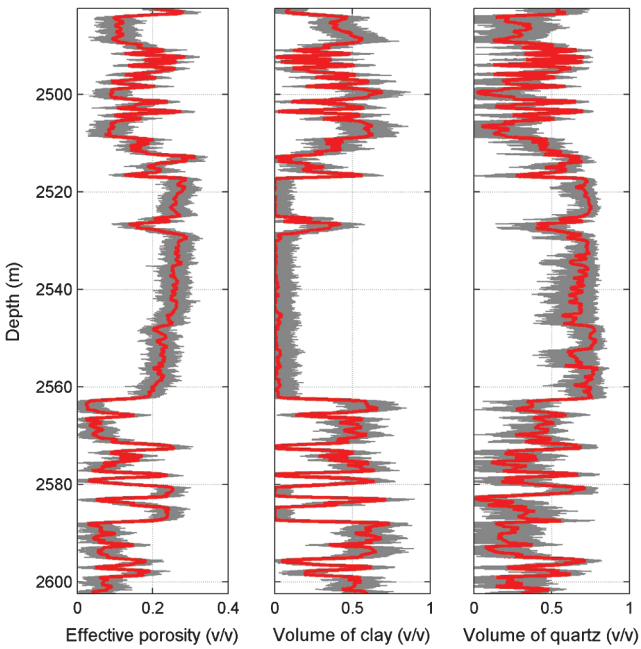

Figure 9. Set of 100 realizations of petrophysical curves (gray curves), from left to right: effective porosity, volume of clay, and volume of quartz (volume of silt is computed by difference 1 minus the sum of effective porosity, clay, and quartz). The pointwise median curve (P50) is displayed in red.

APPLICATION: SECOND EXAMPLE

The second case study is in the North Sea: The reservoir is located at approximately 1750 m, has an average thickness of about 80 m and is part of a complex fluvio-deltaic system with sequences of sandstone and shale. Sandstone and shale layers are relatively thin compared to the seismic resolution. Several works have been published on nearby fields in the North Sea (Mukerji et al., 2001a, 2001b; Avseth et al., 2005).

A complete set of well logs is available for one well of the area. The well is slightly deviated and passes through the main reservoir layer filled with oil. The set of sonic logs, density, and a set of preliminary petrophysical curves performed with commercial software are shown in Figure 14. We point out that three main mineralogical fractions have been identified: clay (mainly illite), muscovite, and quartz. Porosity is relatively high in clean sand in the upper part of the reservoir. The quality of the sonic log is generally good except for a few data samples where some high peaks in P-wave velocity are recorded, whereas no changes are recorded by S-wave velocity, generating unrealistic values of Poisson ratio and $V_{\mathrm{P}}/V_{\mathrm{S}}$ ratio.

Furthermore, a preliminary facies classification has been performed based on sedimentological information, accurate depositional models, and core sample analysis. A set of eight facies has been identified in this reservoir: (1) marine silty-shale; (2) prodelta; (3) flood plain; (4) mouth bar; (5) distributary channel; (6) crevasse splay; (7) tidal deltaic lobes; (8) tight. The distribution of porosity and clay content color coded by facies classification (Figure 15) shows that some of these sedimentilogical facies cannot be discriminated by petrophysical properties. For example, marine silty-shale, prodelta, and flood plain have similar distributions in terms of porosity and clay content (Figure 15). As a consequence, the elastic response of some facies is approximately the same. For this reason, we applied the proposed methodology with two different classifications. We first used the eight-facies sedimentological definition, followed by a simplified three-facies classification, namely sand, silty-sand (mixed), and shale, to classify facies that can be recognized at seismic scale.

As in the previous example, we show the different steps of the methodology. First, an RPM is calibrated at the well location. In this case, we applied a stiff sand model (Mavko et al., 2009). Soft sand and stiff sand models belong to the group of granular media models and are based on Hertz-Mindlin contact theory (Appendix B): The soft sand model extrapolates elastic property values to low porosities by using a modified Hashin-Shtrikman lower bound; the stiff sand model uses a modified Hashin-Shtrikman upper bound, resulting in an increase of elastic properties values compared to the soft sand model. Similarly to the previous case, the calibration is performed in wet conditions by comparing sonic-logs data and

rock-physics model predictions (Figure 16). The model parameters are effective pressure $P = 27$ MPa, critical porosity of $\phi_0 = 0.42$, and coordination number $n = 9$. The overestimation of S-wave velocity in the stiff sand model is well-known (Mavko et al., 2009) and, for this reason, we applied a factor of $3/4$ to reduce S-wave velocity predictions and match the data. This correction is purely heuristic but it often is applied in real cases (a physical explanation of this reducing factor for shear wave can be found in Bachrach and Avseth, 2008). Another method to achieve a good fit is to reduce the values of shear moduli of mineralogical fractions; however, this could lead to unrealistic values for this mineralogy and would require a recalibration of P-wave velocity. The direct comparison of RPM predictions and sonic logs is shown in Figure 17: Generally, we have a good agreement except for the data samples with high peaks in V_P velocity. A slight overestimation of velocity in the upper shaly layer is observable but this could be due to a different mineralogical composition of clay in the overcap layer on top of the reservoir.

We then apply the uncertainty propagation methodology to petrophysical properties and rock physics elastic attributes. We generate a set of 50 realizations of effective porosity, volume of clay, and volume of quartz (volume of muscovite is computed by difference to guarantee that the sum of the mineralogical fraction is 100%); we then compute the corresponding set of P-wave and S-wave velocities and density (Figure 17). The parameters used for the probabilistic formation evaluation analysis are the same as those used in the previous case and summarized in Table 1.

Rock physics diagnostics (Figure 15) showed that, for this specific case, the use of V_S velocity could improve facies classification in the combined petroelastic domain. For this reason, we decided to use in this case an extended data set made by effective porosity, volume of clay, volume of quartz, P-wave, and S-wave velocities. We then performed the facies classification with eight sedimentological facies. In such an application, we expect more variability and

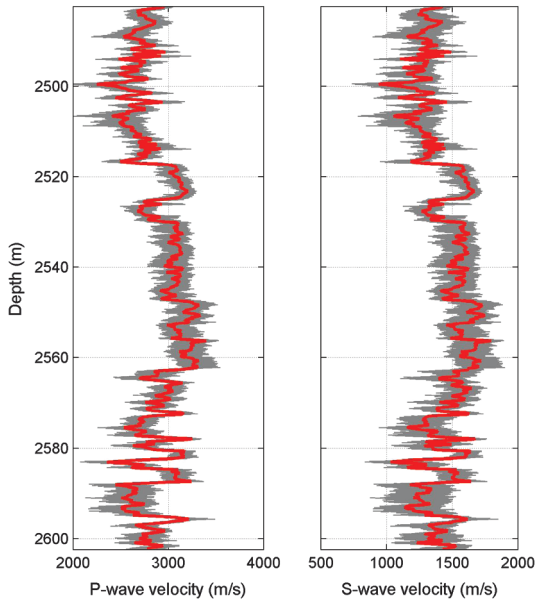

Figure 10. Set of 100 realizations of elastic curves (gray curves), from left to right: P-wave and S-wave velocity. The pointwise median curve (P50) is displayed in red.

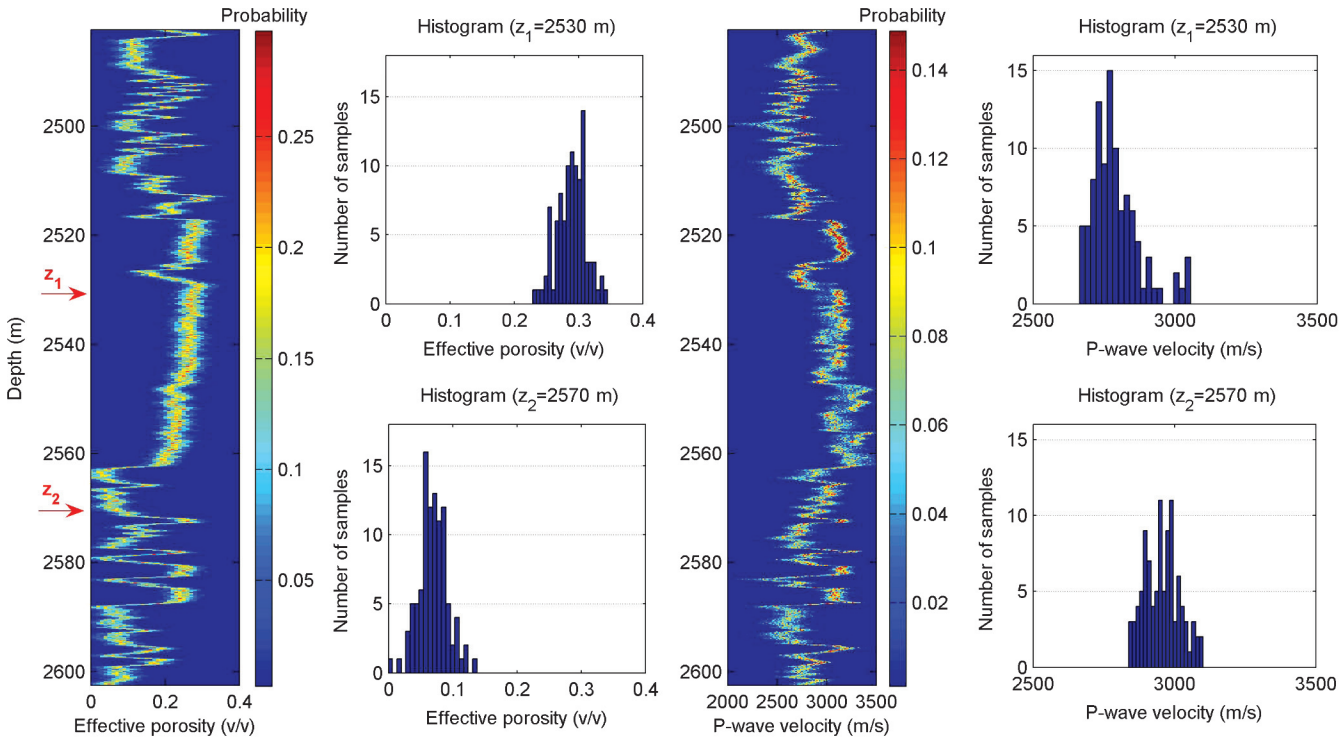

Figure 11. Posterior PDFs of porosity (left side) and P-wave velocity (right side). For two depth locations z_1 and z_2, we also show the histogram of the simulated values.

higher entropy compared to the previous example as the number of facies is higher and some facies cannot be discriminated from the petrophysical and/or elastic point of view. The full set of 50 realizations is shown in Figure 18: By comparing all the realizations we can detect some similar features (such as layers made by facies one, two, and three, and layers made by facies four and five in the upper part) but, as expected, it is hard to discriminate between facies with similar petroelastic properties. The extracted statistics, in particular the e-type (ensemble average) and the maximum a posteriori (MAP) of facies, help to interpret the results (Figure 18). The e-type is the ensemble average of the set of models and is a continuous variable; the maximum a posteriori represents the facies that maximizes the posterior probability estimated from the ensemble of realizations as the facies frequency divided by the total number of realizations (Figure 19). In the upper part of the profile, we can recognize the shaly overcap made by facies 1, 2, and 3 and the thicker layer of the main reservoir made by facies 4 and 5. In the lower part, the variability in the classification is higher, which results in a sequence of thin layers. However, if we look at the whole set of realizations, we observe that even in the upper part of the reservoir the variability is quite high; this is confirmed by the entropy profile (Figure 19). Differently from the previous case, the overall entropy is generally high; this can be due to the large number of facies in the classification but also to the vertical heterogeneity of petrophysical and elastic properties (Figure 17).

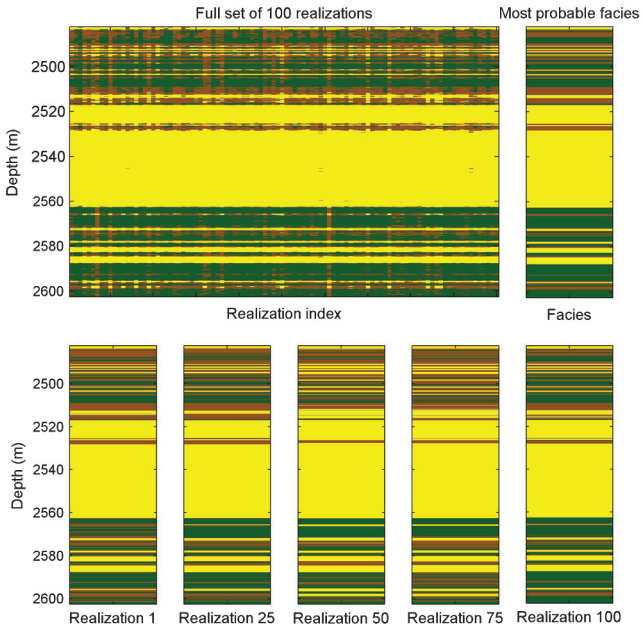

Figure 12. Set of 100 realizations of facies (top left) at well location and estimated most likely facies profile (top right). On the bottom five realizations are shown: LCT in green, MCT in brown, and HCT in yellow.

To improve the facies classification at the well location, we introduced a new approach based on Markov chain methods (Krumbein and Dacey, 1969). Markov chains are a statistical tool that have been used in geophysics to simulate facies sequences to capture the main features of the depositional process. Markov chains are based on a set of conditional probabilities that describe the dependency of the facies value at a given location on the facies values at the locations above (upward chain) or the locations below (downward chain). The chain is said to be first-order, if the transition from one facies to another depends only on the immediately preceding facies. The transition matrix is built by assuming (or estimating from other wells in the field or nearby fields) prior proportions and transition probabilities. The matrix in our case is eight by eight elements: the terms on the diagonal of the transition matrix are related to the thickness of the layers. The height of the numbers on the diagonal corresponds to the strength of the probability of observing no transition, and, as a consequence, also corresponds to the thickness of the layer.

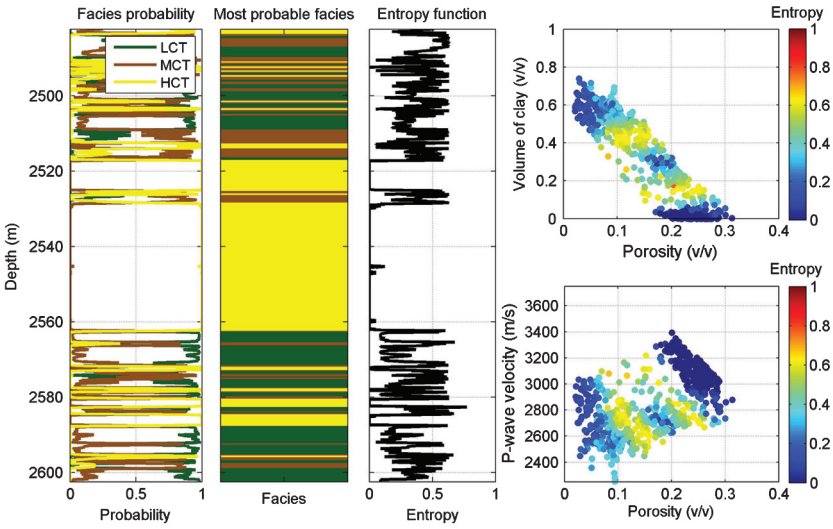

Figure 13. Posterior probability distribution of facies estimated by Monte Carlo simulation (left), most likely facies profile (middle), and associated entropy function (right): LCT in green, MCT in brown, and HCT in yellow. On the right, we show two crossplots in petrophysical (top right) and petroelastic (bottom right) domain, color-coded by the associated entropy given by the probabilistic facies classification.

In our application, the posterior probability of facies estimated from logs of petroelastic properties is combined with the probabilities of the transition matrix (for mathematical details, see Grana and Della Rossa, 2010) by a simple integration, and facies values are sampled from the resulting distribution. In Figure 19, we show one realization of facies obtained by this technique: the use of Markov chain allows us to account for vertical continuity and produces a more realistic classification in cases where a high number of facies is identified. However, the main limitation of such a choice is a result of the assumptions required to assemble the transition matrix. In our case, for example, we do not have any information about the distribution and thickness of facies seven and eight, and the results we obtained only depend on the information assumed from the depositional and sedimentological model.

Finally, we repeated the same application with a simplified classification: we assumed three facies in the reservoir that, for the sake of simplicity, we called shale, silty-sand, and sand. The

reservoir characterization study performed on this field had shown that these facies can be discriminated in the petroelastic domain and represent the facies classification recognizable at seismic scale. Figure 20 shows the comparison of the results obtained with the different classification. The methodology clearly works better when a limited number of facies is identified (also by comparing the entropy profiles of Figures 19 and 20); however we notice that there is a good agreement between the two classifications as facies one, two, and three (in eight facies definition) correspond to shaly layers, and facies four and five (in eight facies definition) correspond to sandy layers.

DISCUSSION

The proposed methodology classifies log-facies at the well location by integrating three different disciplines (formation evaluation, RPM, and cluster analysis) by means of a statistical approach that allows us to understand how plausible the obtained interpretation is, or, in other words, how large the uncertainties are in the obtained solution. The assessment of uncertainty in geophysical inverse problems has been addressed in several works (Jackson, 1972, 1979; Tarantola, 2005), but a standard application to well-log analysis is, in general, neglected. Classical deterministic workflows in formation evaluation analysis and, in particular, LFC, do not allow us to propagate the uncertainty in measured data to the property estimates in a robust and comprehensive way. Statistical methods, on the other hand, can incorporate all available information on the studied system (observational data, theoretical predictions, and prior knowledge) and this information can be represented by probability distribution functions. According to the statistical approach, the solution is not unique, but rather it consists of a posterior probability distribution function over the model domain space that describes the probability of a given model being the closest to the true one.

The aim of this paper is to enhance the classical approach by introducing Monte Carlo simulations providing a proper way to account for a statistical view. Because each step of the methodology strongly depends on the previous one, a robust uncertainty analysis and classification from the very beginning (initial acquired data set) is a mandatory requirement for a safe and trustworthy probability distribution definition. The input uncertainties represent the critical point of the methodology. Formation evaluation is the aforementioned starting point and so it drives the remaining modeling. As we have already pointed out, the prior characterization of input uncertainties for QLI relies on separating the different sources of errors associated with tool measurements, environmental corrections, preprocessing computations, and model parameterizations. Systematic errors need to be corrected because their propagation is meaningless. Random errors may not be neglected when the number of measurements to be averaged is low and/or the measurement precision is very poor.

Monte Carlo simulations try to address this uncertainty issue, but a robust and straightforward way to account for all the sources of uncertainty is still lacking. However, other open problems exist and two of the major ones are the possible correlation existing between the uncertainty affecting the different log acquisitions and the various reference volumes of investigation involved (vertical and lateral resolution). The assignment of probability distribution functions to log measurements assumes a fixed and certain value of the given nominal depth. A more robust approach should also

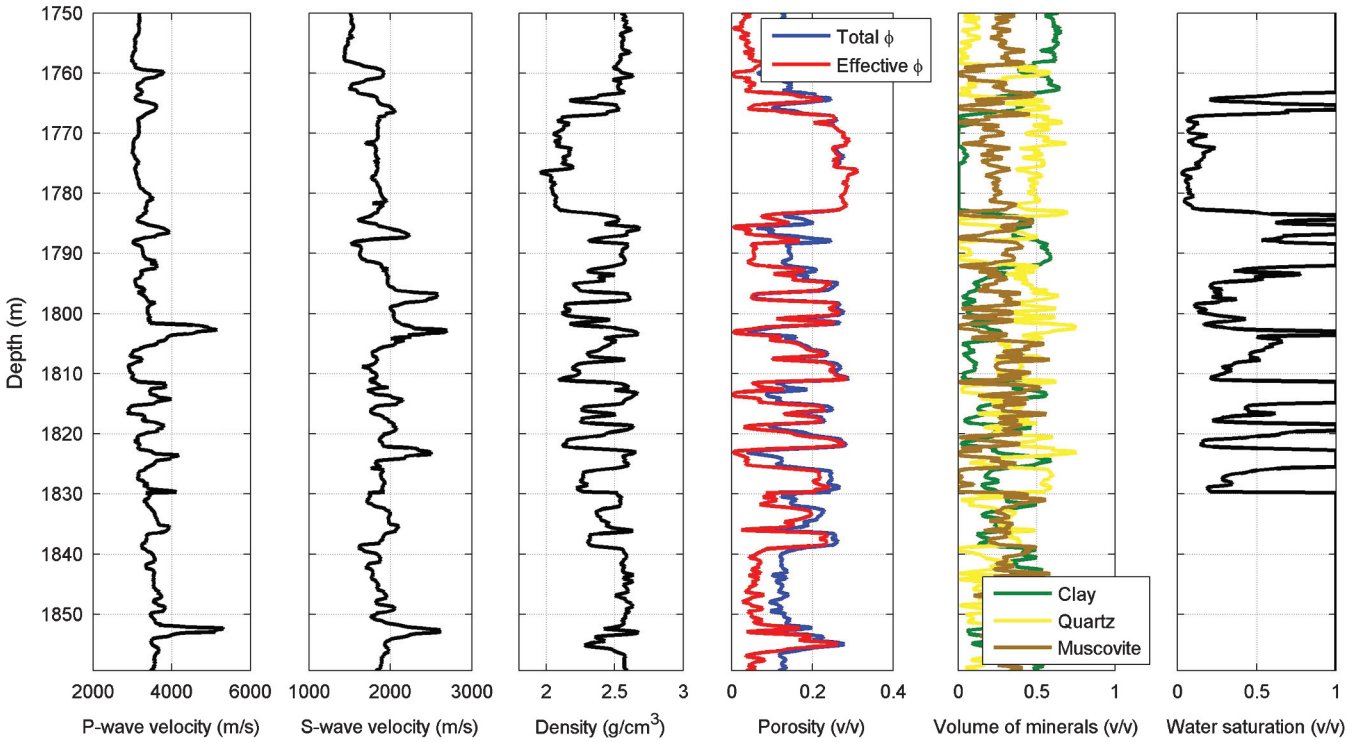

Figure 14. Well-log data set and preliminary petrophysical curves at well location, from left to right: P-wave and S-wave velocity; density; porosity (total porosity in blue, effective porosity in red); volumetric fraction curves (clay in green, quartz in yellow, and muscovite in brown); and water saturation.

consider the vertical uncertainty related to the recorded depth of any log measurements.

An important point of discussion is also related to the constrained number of facies in the probabilistic workflow. In particular, the methodology quantifies and propagates the uncertainty through LFC once the number of facies has been a priori fixed. Monte Carlo simulations are applied to the multivariate statistical technique chosen to discriminate a given number of log-facies. Different

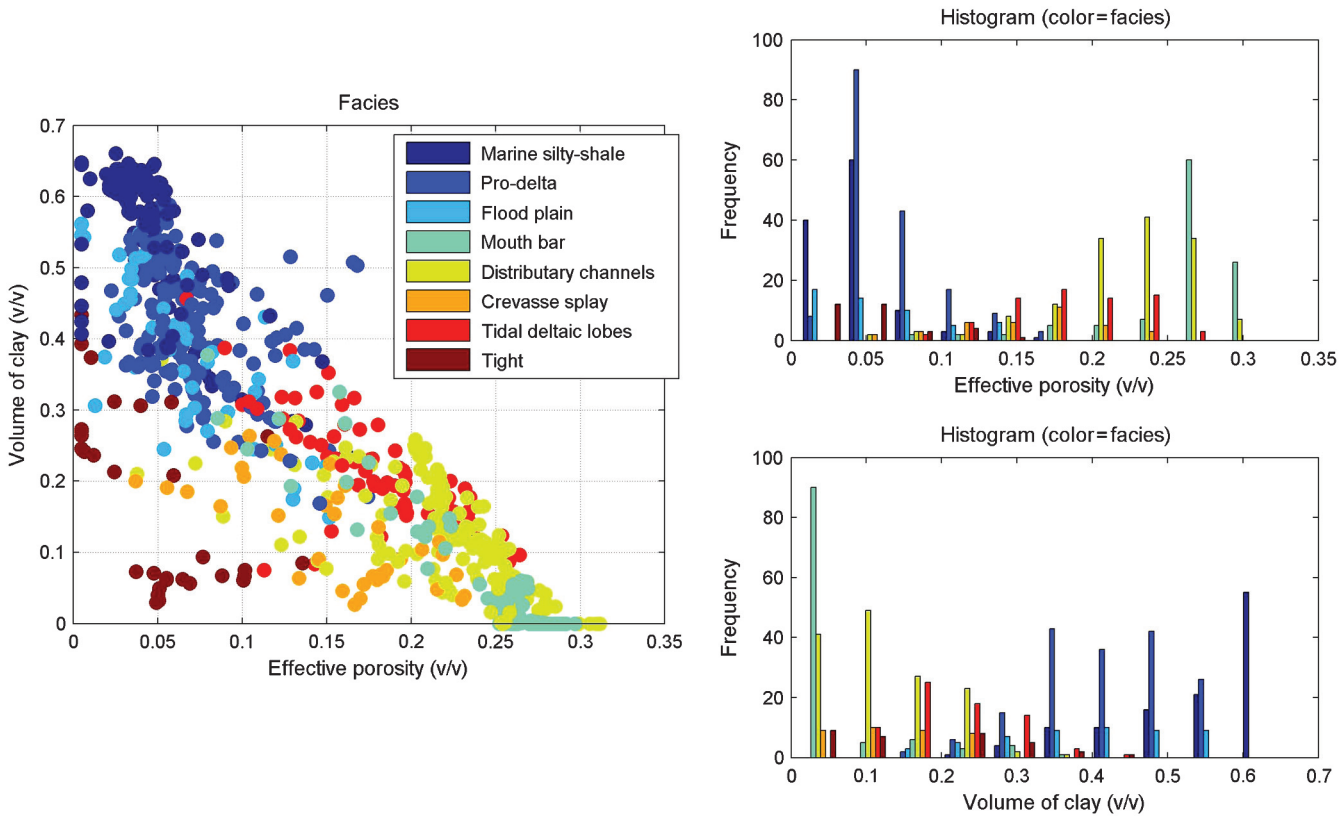

Figure 15. Preliminary facies classification: a set of 8 facies has been identified in this reservoir: (1) marine silty-shale; (2) prodelta; (3) flood plain; (4) mouth bar; (5) distributary channel; (6) crevasse splay; (7) tidal deltaic lobes; (8) tight (left). Grouped histogram of effective porosity (top right) and clay content (bottom right) as a function of facies classification.

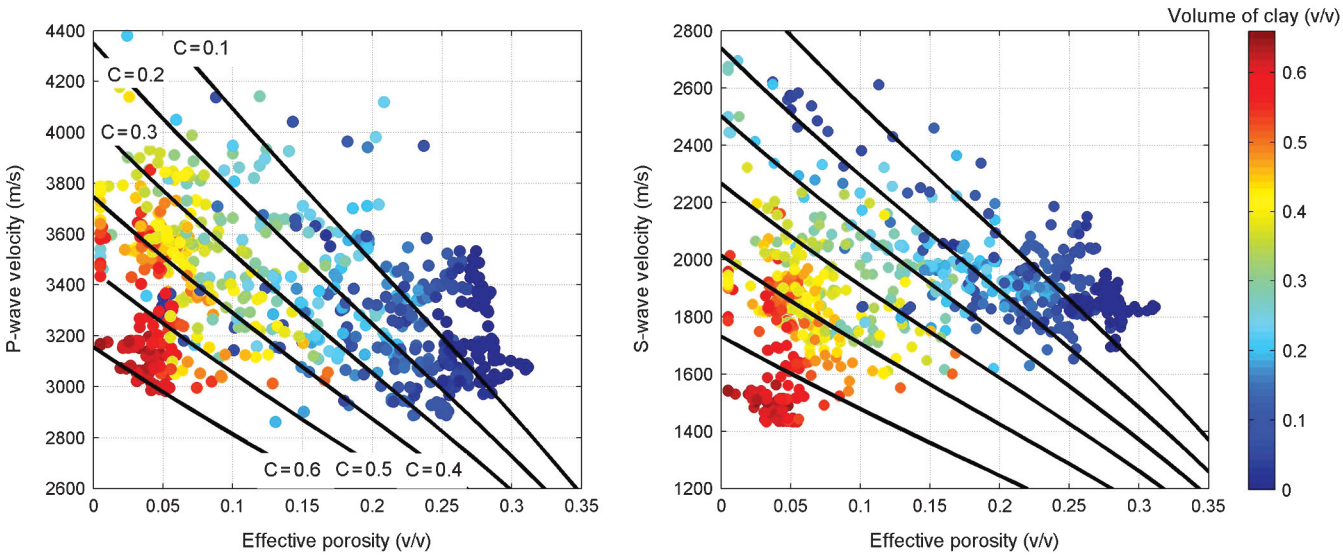

Figure 16. Rock physics crossplots: (left) P-wave velocity versus effective porosity; (right) S-wave velocity versus effective porosity color-coded by volume of clay. Black lines represent stiff sand model for different clay contents (from top to bottom: 10%, 20%, 30%, 40%, 50%, and 60%).

petrophysical and elastic scenarios in the probabilistic training set could be better described by a number of facies that is higher or lower with respect to the a priori fixed one. A generalization of the methodology to take into account this issue also could help for a more rigorous characterization.

The application of the entire workflow to real data sets seems to confirm its feasibility and reliability. Log-facies probability and entropy evaluation quantify the uncertainty in the classification and represent a probabilistic approach from the very beginning of borehole data interpretation for seismic reservoir characterization. It is worth mentioning that even more complex systems need to be tested and studied to finally state all the potentiality of the methodology here developed.

It is also worth mentioning the choice of Monte Carlo simulations for accounting the uncertainty propagation problem. First of all, Monte Carlo techniques calculate the probability density of any functions of random variables and do not require any a priori assumption on the type of distribution for the results. In fact, QLI and RPM can be highly nonlinear and the uncertainty propagation problem cannot be solved analytically. In this light, Monte Carlo simulations provide a simple means by which uncertainties in inputs can be translated into uncertainties in the calculated output properties. Moreover, this methodology is very flexible, allowing different interpretation models to be built and uncertainties tested in a robust way. Correlated as well as independent parameters can be handled and all the possible constraints defining the problem can be taken into account in a straightforward way. The downside to Monte Carlo simulations is that a large number of simulations are

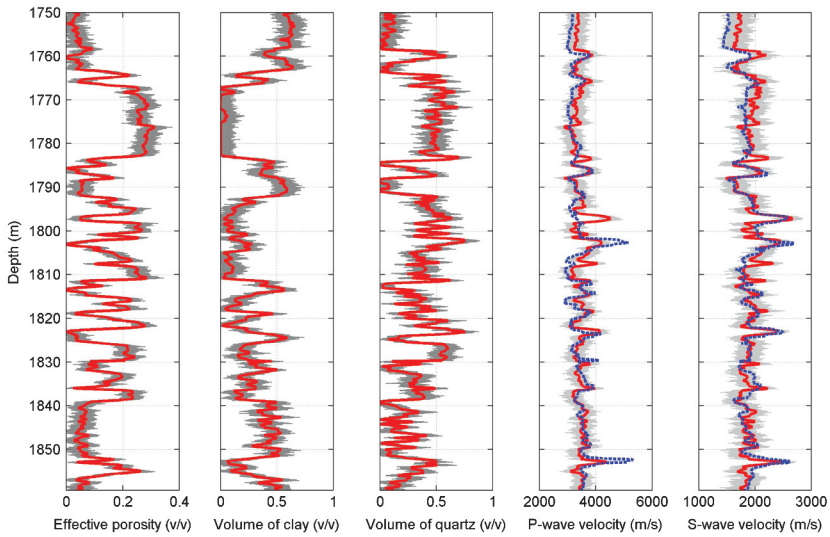

Figure 17. Set of 50 realizations of petrophysical and elastic curves (gray curves), from left to right: effective porosity, volume of clay, volume of quartz (volume of silt is computed by difference 1 minus the sum of effective porosity, clay, and quartz), and P-wave velocity and S-wave velocity predictions. The pointwise median curve (P50) is displayed in red. The dashed blue line represents sonic-log data.

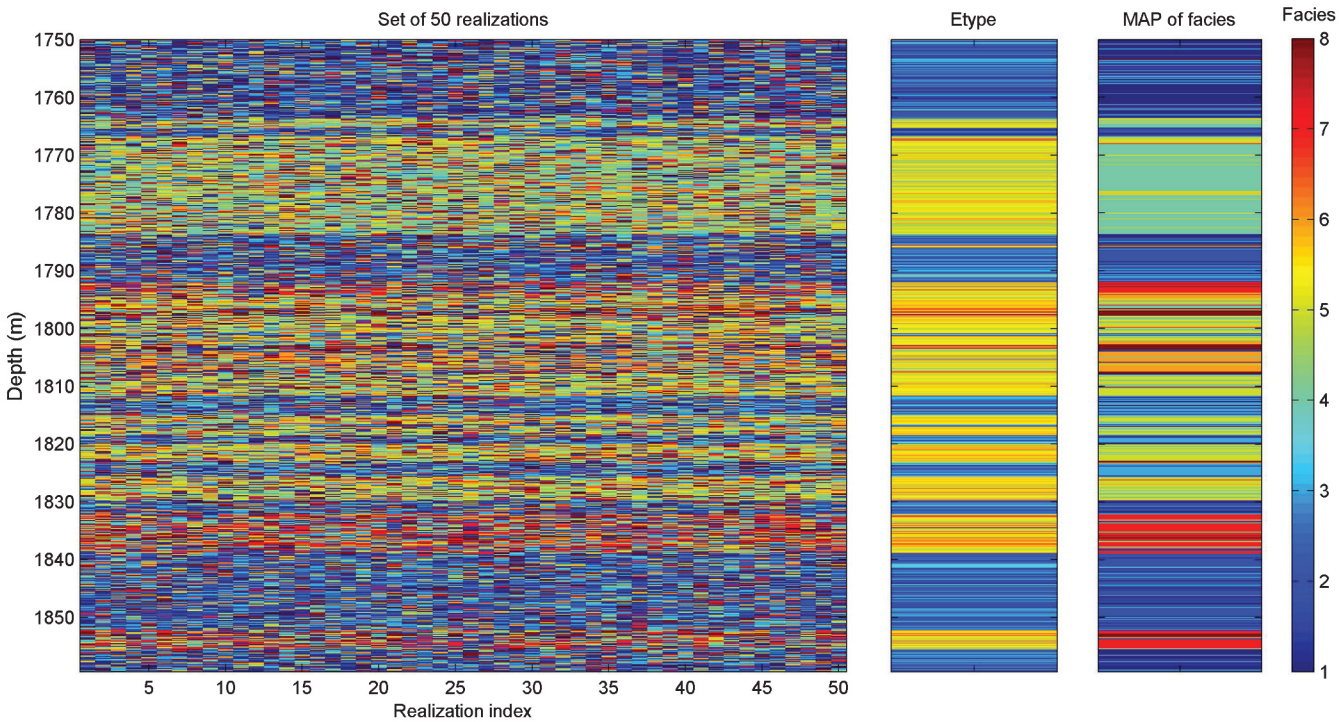

Figure 18. Set of 50 realizations of facies (left) at well location and e-type (ensemble average) and estimated most likely facies profile (top right). Color codes are the same as in Figure 15.

typically required for meaningful statistics to be developed. So, it is fundamental to understand the number of samplings needed to get stable results. To conclude, with Monte Carlo simulations, we have a powerful tool to propagate uncertainties in the different steps of the methodology: the technique can be applied to different studies, independently of the physical models (defining formation evaluation and rock physics interpretation) and mathematical techniques (for LFC) chosen for the specific case.

As described in the application sections, in this work, we performed LFC following a hierarchical agglomerative clustering algorithm with a Ward's minimum variance linkage method. Other methods can be used to classify facies, such as linear discriminant analysis, partitioning cluster algorithms, or neural networks. In particular, in petrophysics, neural networks are used in unsupervised learning mode. The different classification algorithms were previously investigated to rank the obtained results, undertaking tests on known synthetic data sets of variable complexity. The objective was to understand which methods are the most effective in generating an LFC suitable for reservoir characterization purposes. In particular, the proposed and here used hierarchical agglomerative Ward's technique proved by far the best, as it correctly classified most of the synthetic data sets. K-means clustering, unsupervised neural networks, centroid linkage methods, and unweighted average (UPGMA) with Euclidean distance algorithms showed reasonably good classification rates, even though they were not suitable for arbitrary-shaped data clouds. Moreover, no significant advantages were observed when Manhattan or Mahalanobis distance were used in substitution for the Euclidean distance in the UPGMA algorithm. Other clustering algorithms (e.g., single linkage, complete linkage, and weighted average) proved unsuccessful because they were characterized by high misclassification rates. Given the above considerations on the ranking and availability of the different classification algorithms, we focused on Ward's hierarchical agglomerative method. However, the overall methodology works with any appropriate classification algorithm. As previously mentioned, the main advantage of the presented workflow is that the Monte Carlo simulations allow us to extend the training data set and to propagate the uncertainty from the input measured data to the facies classification.

Uncertainty evaluation is important from a qualitative point of view to assess the reliability of the estimated properties, but it can be used in quantitative modeling by using the full probability distributions of the property. In reservoir modeling, many geostatistical simulations methods (for example, sequential indicator simulation and sequential Gaussian simulation) require input prior distributions of the properties we want to model, i.e., facies proportions and the prior distributions of rock properties (such as porosity and net-to-gross). These distributions are generally assumed from prior geological knowledge of the field or nearby fields, but, in our case, the distributions can be extracted directly from Monte Carlo simulations. Nevertheless, we point out that the PDFs derived from Monte Carlo simulations are depth-dependent; stationary conditions should be verified before extrapolating marginal distributions. In a practical reservoir modeling workflow (Doyen, 2007), we first generate reservoir facies models with sequential indicator simulations (or multipoint geostatistics if a suitable training image is available). Then, we generate porosity and other reservoir property models with sequential Gaussian simulation where the prior distributions of reservoir properties are facies-dependent and are estimated from the marginal distributions derived from Monte Carlo simulations in each facies.

An important point of our methodology is the integration of rock-physics modeling. The main advantage of this step is the link between facies classification and elastic properties, which is provided by the model calibrated at the well location. In fact, this link

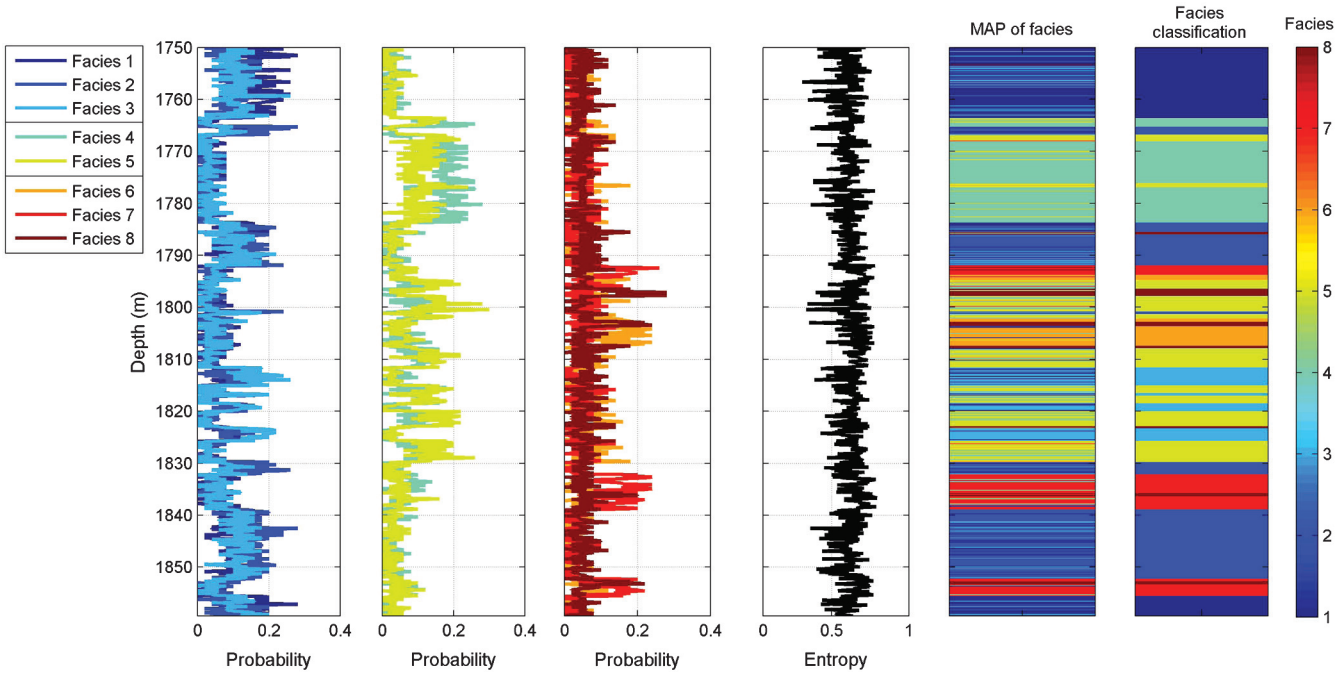

Figure 19. Facies probability estimated from ensemble and extracted statistics, from left to right: probability of facies (first three plots); entropy, maximum a posteriori of facies distribution, and new facies classification performed by Markov chain integrated approach. Color codes are the same as in Figure 15.

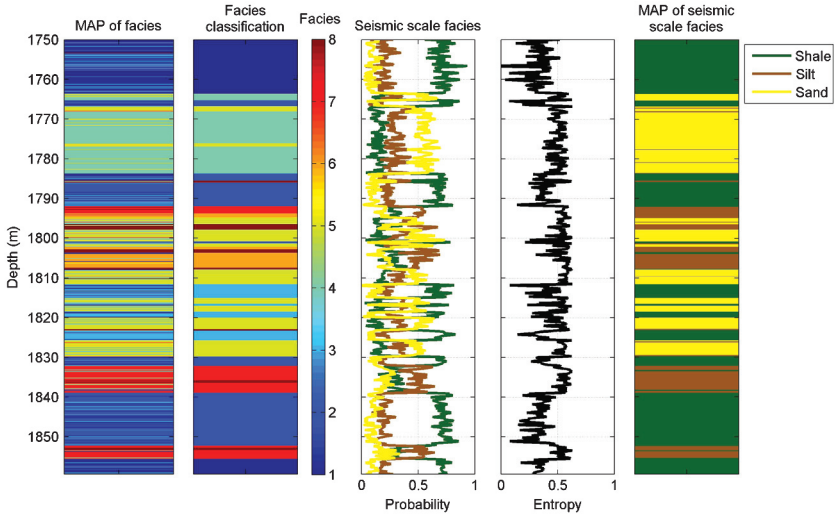

Figure 20. Comparison of two different classifications. On the left (first two plots), we show the eight-facies classification: maximum a posteriori of facies and Markov chain classification (color codes are the same as in Figure 15). On the right, we show the simplified three-facies classification (namely seismic facies classification, last three plots): posterior probability distribution, entropy, and maximum a posteriori of facies (shale in green, silty-sand in brown, sand in yellow).

guarantees that facies classification estimated from seismic amplitudes in seismic characterization studies is consistent with LFC at well location. Instead of rock physics predictions, sonic logs could be used as well, but the correlation between sonic logs and petrophysical curves is not always optimal in well-log analysis. Furthermore, resolution of sonic logs is generally lower than the resolution of well logs used in formation evaluation analysis. If the sonic logs are reliable and correlated to petrophysical logs, the lower resolution of sonic logs would reduce the entropy but, at the same time, increase the uncertainty in the output estimation. On the other side, RPMs are applied to a set of petrophysical curves with similar resolution, which results in a complete consistent petroelastic data set. The drawback of such a choice is that the RPM approximation could be not accurate in some interval layers, as pointed out in our applications, but this lack of accuracy is partially compensated by Monte Carlo multiple realizations.

Finally, we point out that in our methodology, we tried to include different sources of uncertainty: log measurements, RPM approximations, heterogeneity, and natural variability of the rocks. However, other factors could influence formation evaluation analysis, RPM, and LFC, such as anisotropy, laminar and/or dipping layering, and well deviation. These issues should be carefully evaluated in each case and a sensitivity analysis should be performed to verify if they can be neglected or not.

CONCLUSIONS

The proposed probabilistic LFC workflow is based on the integration of three different disciplines: QLI, RPM, and multivariate statistical techniques (cluster analysis) for LFC. Furthermore, this workflow focuses on the uncertainty propagation through the three different steps and is based on Monte Carlo simulation method. The application of the proposed methodology allows us to generate several log-facies profiles at well location and to estimate the most probable facies classification and the associated uncertainty via entropy computation. The so-obtained classification includes petrophysical properties and acoustic/elastic attributes to directly link log-facies to seismic velocities used in facies simulation constrained by seismic information. This workflow represents a rigorous starting point to propagate the well-facies classification and related uncertainty to the entire reservoir modeling. These obtained uncertainties are more realistic inputs to reservoir modeling, in contrast to taking log data as fixed and certain (so-called "hard data" in geostatistics). The method has been illustrated on real well-log data sets in clastic environments but can be applied to different scenarios.

ACKNOWLEDGMENTS

We thank Eni E&P management for the permission to publish the paper and Stanford University, Stanford Rock Physics and Borehole project, and Stanford Center for Reservoir Forecasting for the support. We also would like to thank Livio Ruvo (Eni E&P) for helpful discussion and comments.

APPENDIX A

FORMATION EVALUATION ANALYSIS EQUATIONS

We recap here the basic equations for formation evaluation analysis (Darling, 2005; Ellis and Singer, 2007). Starting from well-log measurements, formation evaluation computes volumetric curves of the corresponding formation components. Assuming a unitary investigated total volume,

$$\sum_{h=1}^{N} v_h = 1, \tag{A-1}$$

where $0 < v_h < 1$ is the volumetric fraction of the single hth component and N is the total number of minerals and fluids defining the modelled formation.

Fundamental in dealing with response-equations is the knowledge of how clay is handled and the concept of wet versus dry clay. Here, we take the clay as wet, so the volume fraction v_{clay} is intended to consider this phase composed of dry clay and associated bound water. As already mentioned, well-log measurements could be affected by the existing step invasion profile due to the circulating mud fluid. The mud invasion divides the formation into two separated parts: (1) the flushed zones (XO) and (2) the unflushed deep zone (DE). Well-logs tools have different depths of investigation and can be influenced only by one zone (shallow reading-log tools and deep reading logs) or by both zones; then, their response can contain terms for only one or for all formation components. Mathematically this is taken into account by a special factor $0 < \psi_i < 1$ called the invasion factor, which controls how much influence for each log i (such as neutron porosity or density) comes from the XO zone. The remaining influence $(1 - \psi_i)$, comes from the DE zone. In well-log interpretation processes it is common to assume a lateral homogeneity and, as a consequence, to assume that

the sum of the fluid volumes in the flushed zone is equal to the sum of fluid volumes in the unflushed zone

$$\sum_{k=1}^{N_{\text{fluid}}} (v_{XO_k} - v_{DE_k}) = 0, \tag{A-2}$$

where the sum only considers the fluid involved.

Other constraints involve the feature of the mud used. For example, when oil-based mud is used, the invading fluid is hydrocarbon. Therefore, we must impose a constraint that limits the volume of water in the flushed zone to be less than or equal to the volume of water in the undisturbed zone on the basis of the normal response-equations. The reverse is true for water-based mud.

Finally, it is also possible to limit the sum of the fluid volumes to be less than a particular critical porosity. This critical porosity is the value that marks the limit between a saturated rock and a suspension and will also be relevant for the RPM (Appendix B).

In general, a response equation is a mathematical description of how a given measurement m_i (where i is the generic index to identify a well-log type) varies with respect to each formation components. The simplest linear response-equations are of the form

$$m_i(\mathbf{v}) = \sum_{j=1}^{N_{\text{solid}}} \tilde{m}_j v_j + (1-\psi_i) \sum_{k=1}^{N_{\text{fluid}}} \tilde{m}_{XO_k} v_{XO_k} + \psi_i \sum_{k=1}^{N_{\text{fluid}}} \tilde{m}_{DE_k} v_{DE_k}, \tag{A-3}$$

where $\tilde{m}_j$, $\tilde{m}_{XO}$, and $\tilde{m}_{DE}$ are the response parameters for the corresponding formation components. This form is typical for neutron, density, and sonic logs. As gamma ray log only considers solid components, it is not affected by invasion and the response equation does not involve the invasion factor.

On the other hand, the resistivity/conductivity model used is nonlinear and, in our formation, we consider two different log measurements: one related to the flushed zone (shallow reading log) and the other in the undisturbed zone (deep reading). The response-equations are the same for both situations, but the parameters involved depend on the investigated zone.

For conductivity C, we consider the theoretical form of the Indonesian equation (Poupon and Leveaux, 1971) because it is widely used in shaly sand formations:

$$\sqrt{C} = \left[\sqrt{C_{\text{clay}}}\, v_{\text{clay}}^{(1-\frac{1}{2}v_{\text{clay}})} + \frac{\sqrt{C_w}}{\sqrt{a}} \phi^{\frac{1}{2}(m+\frac{m_2}{\phi})} \right] \left(\frac{v_w}{\phi}\right)^{\frac{n}{2}}, \tag{A-4}$$

where C_{clay} and C_w are the specific conductivities of clay and water, respectively; v_w is the volume of water; and the parameters m, m_2, n, and a come from empirical observations.

APPENDIX B

ROCK PHYSICS EQUATIONS

Granular media models are based on Hertz-Mindlin contact theory (Mavko et al., 2009) and they provide estimates for the bulk K_{HM} and shear moduli μ_{HM} of the dry rock, assuming that the frame is a dense random pack of identical spherical grains subject to an effective pressure P with a given critical porosity ϕ_0 and an average number of contacts per grain n. Critical porosity is the value which marks the porosity limit between a saturated rock and a suspension. Here, we recap the set of equations of the soft sand model used in the application case.

First Hertz-Mindlin equations are given by

$$K_{\text{HM}} = \sqrt[3]{\frac{P[n(1-\phi_0)\mu_{\text{mat}}]^2}{18[\pi(1-\nu)]^2}} \tag{B-1}$$

$$\mu_{\text{HM}} = \frac{5-4v}{5(2-v)} \sqrt[3]{\frac{3P[n(1-\phi_0)\mu_{\text{mat}}]^2}{2[\pi(1-\nu)]^2}}, \tag{B-2}$$

where ν is the grain Poisson's ratio, namely

$$v = \frac{3K_{\text{mat}} - 2\mu_{\text{mat}}}{2(3K_{\text{mat}} + \mu_{\text{mat}})}, \tag{B-3}$$

and K_{mat} and μ_{mat} are the solid phase (zero porosity) elastic moduli.

For porosity values ranging between zero and the critical porosity, the soft sand RPM connects the grain elastic moduli K_{mat} and μ_{mat} with the elastic moduli K_{HM} and μ_{HM} of the dry rock at critical porosity. This is done by interpolating these two end members in the intermediate porosity values by means of a heuristic-modified Hashin-Shtrikman lower bound:

$$K_{\text{dry}} = \left(\frac{\frac{\phi}{\phi_0}}{K_{\text{HM}} + \frac{4}{3}\mu_{\text{HM}}} + \frac{1-\frac{\phi}{\phi_0}}{K_{\text{mat}} + \frac{4}{3}\mu_{\text{HM}}} \right)^{-1} - \frac{4}{3}\mu_{\text{HM}} \tag{B-4}$$

$$\mu_{\text{dry}} \left(\frac{\frac{\phi}{\phi_0}}{\mu_{\text{HM}} + \frac{1}{6}\xi\mu_{\text{HM}}} + \frac{1-\frac{\phi}{\phi_0}}{\mu_{\text{mat}} + \frac{1}{6}\xi\mu_{\text{HM}}} \right)^{-1} - \frac{1}{6}\xi\mu_{\text{HM}} \tag{B-5}$$

$$\xi = \frac{9K_{\text{HM}} + 8\mu_{\text{HM}}}{K_{\text{HM}} + 2\mu_{\text{HM}}}. \tag{B-6}$$

On the other hand, the stiff-sand model connects the two end-members with the modified Hashin-Shtrikman upper bound (Mavko et al., 2009). Gassmann's equation is then used for calculating the effect of fluid

$$K_{\text{sat}} = K_{\text{dry}} + \frac{\left(1 - \frac{K_{\text{dry}}}{K_{\text{mat}}}\right)^2}{\frac{\phi}{K_{\text{fl}}} + \frac{1-\phi}{K_{\text{mat}}} - \frac{K_{\text{dry}}}{K_{\text{mat}}^2}} \tag{B-7}$$

$$\mu_{\text{sat}} = \mu_{\text{dry}}, \tag{B-8}$$

where K_{fl} is the fluid bulk modulus, and K_{sat} and μ_{sat} are the rock-saturated moduli. Fluid properties can be obtained by means of Batzle-Wang formulas (Batzle and Wang, 1992), a set of empirical equations that allows calculating bulk moduli and densities of fluid components starting from fluid analysis parameters (such as gas gravity, oil gravity, gas-oil-ratio, temperature, and pressure). Finally, velocities are obtained as follows:

$$V_P = \sqrt{\frac{K_{sat} + \frac{4}{3}\mu_{sat}}{\rho}} \quad \text{(B-9)}$$

$$V_S = \sqrt{\frac{\mu_{sat}}{\rho}}, \quad \text{(B-10)}$$

where ρ is the fluid-saturated bulk density.

REFERENCES

Avseth, P., T. Mukerji, and G. Mavko, 2005, Quantitative seismic interpretation: Cambridge University Press.

Bachrach, R., and P. Avseth, 2008, Rock physics modeling of unconsolidated sands: Accounting for nonuniform contacts and heterogeneous stress fields in the effective media approximation with applications to hydrocarbon exploration: Geophysics, **73**, no. 6, E197–E209, doi: 10.1190/1.2985821.

Batzle, M. L., and Z. Wang, 1992, Seismic properties of pore fluids: Geophysics, **57**, 1396–1408, doi: 10.1190/1.1443207.

Bourbie, T., O. Coussy, and B. Zinszner, 1988, Acoustics of porous media: CRC Press.

Darling, T., 2005, Well logging and formation evaluation: Elsevier Inc.

Doyen, P., 2007, Seismic reservoir characterization: EAGE.

Dvorkin, J., and A. Nur, 1996, Elasticity of high-porosity sandstones: Theory for two north sea data sets: Geophysics, **61**, 1363–1370, doi: 10.1190/1.1444059.

Dvorkin, J., A. Nur, and H. Yin, 1994, Effective properties of cemented granular materials: Mechanics of Materials, **18**, 351–366, doi: 10.1016/0167-6636(94)90044-2.

Eberhart-Phillips, D., D.-H. Han, and M. D. Zoback, 1989, Empirical relationships among seismic velocity, effective pressure, porosity, and clay content in sandstone: Geophysics, **54**, 82–89, doi: 10.1190/1.1442580.

Eidsvik, J., T. Mukerji, and P. Switzer, 2004, Estimation of geological attributes from a well log: An application of hidden markov chains: Mathematical Geology, **36**, 379–397, doi: 10.1023/B:MATG.0000028443.75501.d9.

Ellis, D. V., and J. M. Singer, 2007, Well logging for earth scientists: Springer.

Fylling, A., 2002, Quantification of petrophysical uncertainty and its effect on in-place volume estimates: Numerous challenges and some solutions: SPE 77637, doi: 10.2118/77637-MS.

Gal, D., J. Dvorkin, and A. Nur, 1998, A physical model for porosity reduction in sandstones: Geophysics, **63**, 454–459, doi: 10.1190/1.1444346.

Grana, D., and E. Della Rossa, 2010, Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion: Geophysics, **75**, no. 3, O21–O37, doi: 10.1190/1.3386676.

Heidari, Z., C. Torres-Verdin, and W. E. Preeg, 2010, Improved estimation of mineral and fluid volumetric concentrations in thinly-bedded and invaded formations, Transactions of the SPWLA 51st Annual Symposium.

Jackson, D. D., 1972, Interpretation of inaccurate insufficient and inconsistent data: Geophysical Journal of the Royal Astronomical Society, **28**, 97–109, doi: 10.1111/gji.1972.28.issue-2.

Jackson, D. D., 1979, The use of a priori data to resolve non-uniqueness in linear inversion: Geophysical Journal of the Royal Astronomical Society, **57**, 137–157, doi: 10.1111/gji.1979.57.issue-1.

Kaufman, L., and P. J. Rousseeuw, 1990, Finding groups in data. An introduction to cluster analysis: Wiley.

Kennedy, J., Pujiyono, A. Cox, and R. Aldred, 2010, Using quantified model based petrophysical uncertainty to aid in conflict resolution: Transactions of the SPWLA 51st Annual Symposium.

Krumbein, W. C., and M. F. Dacey, 1969, Markov chains and embedded Markov chains in geology: Mathematical Geology, **1**, no.1, 79–96, doi: 10.1007/BF02047072.

MacBeth, C., 2004, A classification for the pressure-sensitivity properties of a sandstone rock frame: Geophysics, **69**, 497–510, doi: 10.1190/1.1707070.

Mavko, G., and T. Mukerji, 1998, A rock physics strategy for quantifying uncertainty in common hydrocarbon indicators: Geophysics, **63**, 1997, doi: 10.1190/1.1444493.

Mavko, G., T. Mukerji, and J. Dvorkin, 2009, The rock physics handbook: Tools for seismic analysis in porous media: Cambridge University Press.

Mukerji, T., P. Avseth, G. Mavko, I. Takahashi, and E. F. Gonzalez, 2001b, Statistical rock physics: Combining rock physics, information theory, and geostatistics to reduce uncertainty in seismic reservoir characterization: The Leading Edge, **20**, 313–319, doi: 10.1190/1.1438938.

Mukerji, T., A. Jorstad, P. Avseth, G. Mavko, and J. R. Granli, 2001a, Mapping lithofacies and pore-fluid probabilities in a north sea reservoir: Seismic inversions and statistical rock physics: Geophysics, **66**, 988–1001, doi: 10.1190/1.1487078.

Nur, A., and Z. Wang, 1989, Seismic and acoustic velocities in reservoir rocks: Volume **1**, Experimental studies: SEG.

Poupon, A., and J. Leveaux, 1971, Evaluation of water saturation in shaly formations, Transactions of the SPWLA 12th Annual Symposium.

Shannon, C. E., 1948, A mathematical theory of communication: Bell System Technical Journal, **27**, 379–423.

Tarantola, A., 2005, Inverse problem theory: SIAM.

Theys, P., 1991, Log data acquisition and quality control

Theys, P., 1994, A serious look at repeated sections: Transactions of the SPWLA 30th Annual Symposium.

Theys, P., 1997, Accuracy. Essential information for a log measurement: Transactions of the SPWLA 38th Annual Symposium.

Verga, F., D. Viberti, and M. Gonfalini, 2002, Uncertainty evaluation in well logging: Analytical or numerical approach?: Transactions of the SPWLA 43rd Annual Symposium.

Viberti, D., 2010, A rigorous mathematical approach for petrophysical estimation: American Journal of Applied Sciences, **7**, no. 11, 1509–1516, doi: 10.3844/ajassp.2010.1509.1516.

Wang, Z., and A. Nur, 1992, Seismic and acoustic velocities in reservoir rocks: Volume **2**, Theoretical and model studies: SEG.

Wang, Z., and A. Nur, 2000, Seismic and acoustic velocities in reservoir rocks: **3**, Recent development: SEG.

Ward, J. H., Jr., 1963, Hierarchical grouping to optimize an objective function: Journal of the American Statistical Association, **58**, 236–244.