Abstract

Methods for predicting mineralogy from logging tool measurements have been an active area of research for several decades. In spite of these efforts, methods for predicting quantitative mineralogy including clay types from well logging data were not fully achieved. The introduction of geochemical logging tools in the 1980s offered promise; however, early versions of geochemical logging tools did not measure elemental chemistry with enough accuracy and precision to enable reliable and quantitative determination of mineralogy. Recent advances in geochemical logging tool technology now enable accurate and robust measurements of the chemical elemental concentrations that are needed to determine continuous quantitative and detailed logs of mineralogy.

This paper presents a novel approach for determining more accurate, detailed mineralogy from an elemental spectroscopy logging tool. This work was made possible by three recent developments: the introduction of a new high-performance neutron-induced gamma ray spectroscopy logging tool, a new high-quality research database consisting of chemistry and mineralogy measured on cores acquired worldwide from conventional and unconventional reservoirs, and a new model-independent inversion method that overcomes the limitations of previous model-dependent methods.

The model-independent inversion makes use of the database which includes clean sands, shaly sands, shales, carbonates, and complex mixed lithologies. The database contains laboratory measurements of dry-weight elemental chemistry and mineralogy measured by transmission Fourier transform infrared spectroscopy. The database is used to derive a model-independent mapping function that accurately represents the complex functional relationship between the elemental concentrations and the mineral concentrations. Once the mapping function is determined from the database, it can be used to predict quantitative mineralogy from elemental concentrations derived from the logging tool measurements. Unlike previous inversion methods, the model-independent mapping function does not have any adjustable parameters or require any user inputs such as mineral properties or endpoints.

The mapping function is used to predict continuous logs of matrix densities plus concentrations of 14 minerals (i.e., illite, smectite, kaolinite, chlorite, quartz, calcite, dolomite, ankerite, plagioclase, orthoclase, mica, pyrite, siderite, and anhydrite) from eight dry-weight elemental concentrations derived from
the logging tool. The new method has been applied to well log data acquired worldwide in numerous conventional and unconventional reservoirs having a wide variety of complex mineralogies. The predicted mineralogies and matrix densities are generally found to be in good quantitative agreement with core-derived mineralogies and matrix densities.

Introduction

The early attempts in the late seventies and eighties at predicting mineralogy from well logging data preceded geochemical logging and mostly relied on nuclear tool measurements such as spectral natural gamma ray (Th, U, K), neutron, density, sonic, and photo electric absorption logs. The interpretation of these measurements provided good results in some environments and could quantify major lithological units such as shales, sands, and carbonates. However, these measurements did not have sufficient sensitivity to allow for the determination of complex mineralogy. Modern neutron-induced gamma ray spectroscopy or geochemical logging tools evolved from the C/O tools that were introduced in the 1970s to provide salinity-independent estimates of oil saturations (Culver et al. 1974). An important paper by Hertzog (1980) discussed an experimental inelastic scattering and capture neutron-induced gamma ray logging tool for measuring C/O as well as Si, Ca, and Fe yields to be used for lithology delineation. This paper signaled the birth of modern geochemical logging. Since 1980, geochemical logging tool technology has passed through multiple phases of development with continuing improvements in technology including improved neutron sources, electronics, and scintillation detectors (Radtke et al. 2012; Galford et al. 2009; Pemper et al. 2006). These technological advances resulted in more measured elemental concentrations with improved precisions and accuracies, thus providing an opportunity for more quantitative and detailed mineralogy analysis. This paper presents a model-independent inversion method developed to take full advantage of a new high-quality worldwide chemistry and mineralogy research core database and the more accurate and robust elemental concentrations measured by a new geochemical logging tool (Radtke et al. 2012).

Importance of Mineralogy

Detailed and accurate knowledge of mineralogy is essential for understanding the complexity of conventional and unconventional reservoirs. It provides vital information needed for the evaluation, completion, and production of hydrocarbons. The following examples illustrate the importance of knowing detailed and quantitative mineralogy.

- It is well-known that logging tool measurements are strongly affected by the different lithologies and minerals present in the rocks. For example, clay minerals such as smectite have a large cation exchange capacity, which provides an additional conductivity mechanism in clastic reservoir rocks. The clay conductance becomes a significant effect in the presence of low-salinity formation waters (Waxman and Smits 1968). The excess conductivity arising from the clays reduces the log-measured resistivities, which can result in underestimation of hydrocarbon saturations. Moreover, the equations and models used to derive key formation properties from tool responses require knowledge of rock matrix properties, such as, for example, matrix densities, hydrogen indices, neutron absorption cross sections, sonic wave velocities, and dielectric constants, which can be more accurately determined when the mineralogy is known.

- Detailed knowledge of mineralogy is also needed for optimizing completions to achieve maximum well production. Knowing the concentrations of clays and other minerals present in a petroleum-bearing formation is vital for determining which completion and stimulation fluids should be used to optimize production rates and to avoid near-wellbore formation damage. For example, it is well known that acidizing formations that contain chlorite causes precipitation of iron, which plugs pore throats and greatly reduces the permeability in the vicinity of the wellbore. Similarly, during well completions the injection of low-salinity water into formations containing smectite should be avoided because smectite absorbs water and swells to a larger volume, thereby reducing the permeability. Accurate knowledge of the clay
types and their concentrations is needed to prevent the problem of fines migration. For example, if it is known that illite and kaolinite are present, the flow rates from the well can be lowered to minimize the chances of dislodging fine particles and damaging the formation.

**Shortcomings of Model-Dependent Inversion Methods**

Previous methods for predicting mineralogy are based on models. Model-dependent inversion methods require user inputs such as mineral compositions and mineral endpoints as well as subjective assumptions about which minerals are present and should be included in the model. Computerized log interpretation systems for predicting lithology and mineralogy (Mayer and Sibbit 1980; Quirein et al. 1986; Freedman and Puffer 1988) are based on simultaneous model-dependent inversions of suites of logging tool measurements. Theoretical tool response equations or forward models are used to represent the logging tool measurements. The tool response equations are functions of the reservoir properties including the lithologies and mineral concentrations to be determined. The predicted reservoir properties and mineral concentrations are determined by minimizing cost functions, which are weighted sums of the differences between the measured tool responses and the response equations. There are usually not enough independent measurements to solve for the many minerals that can be present. The model-based inversions attempt to circumvent this problem by requiring the user to select a “mineral model” that specifies a subset of minerals for which the minimization determines the concentrations. Clearly, assuming a mineral model is not a reliable method for predicting mineral concentrations when there are additional minerals present that are not in the assumed mineral model. Another shortcoming of model-based inversions is that the models require specification of mineral endpoints such as densities, hydrogen indices, dielectric constants, and thermal neutron absorption cross sections. Accurate mineral compositions and properties are not known because mineral compositions and properties can vary. This variability is especially notable for the more complex minerals such as the clays, micas, and feldspars. Requiring users to input mineral properties, compositions, and mineral models is one of the shortcomings of model-based inversion methods. It means that mineralogy predictions from model-dependent methods are user dependent because different log analysts can predict different lithologies and mineral concentrations from the same suite of well logging measurements.

Linear regression models have also been used to predict mineralogy. This approach is based on empirically developed linear regression equations for which the coefficients are found by fitting the equations to a core database of elemental chemistry and mineralogy (Herron and Herron 1996). This sequential approach uses different regression equations depending on the feldspar concentration. Another recently published method uses the bulk elemental composition to first determine lithology and constrain the mineral selection, and then to use a sequential mass balance approach to solve for selected minerals (Pemper et al. 2006; Jacobi et al. 2008).

**Developments Now Enabling Quantitative and Detailed Mineralogy Prediction**

This section discusses the confluence of three recent developments that enable quantitative and detailed mineralogy determination from elemental spectroscopy logging tool measurements.

**New High-Performance Spectroscopy Logging Tool**

The recent development of a new high-performance neutron-induced gamma ray spectroscopy geochemical logging tool (Radtke et al. 2012) is one of three critical developments that finally enable the prediction of detailed and quantitative mineralogy from well logs. The advanced technology in the new geochemical tool includes a high-output pulsed neutron generator (PNG), a high-resolution LaBr$_3$:Ce scintillation detector, and an advanced electronics acquisition system to measure very high gamma ray count rates. The output of the PNG exceeds $3 \times 10^8$ neutrons per second and the gamma ray count rate measured by the scintillation detector can exceed $2.5 \times 10^6$ counts per second. The high gamma ray count rate and the high
resolution of the LaBr₃:Ce scintillation detector make possible the improved accuracy and precision of the derived dry-weight elemental concentrations.

The new tool measures gamma rays emitted by nuclei after thermal neutron capture and after inelastic scattering of fast neutrons. The processing of the detected gamma ray spectra provides accurate and precise dry-weight elemental concentrations including Si, Ca, Mg, Al, K, Fe, S, Mn, and C. The high precision and accuracy of these derived dry-weight elemental concentrations are essential for determining detailed and quantitative mineralogy and also total organic carbon.

**New High-Quality Chemistry and Mineralogy Database**

Another key development that now makes it possible to determine quantitative and detailed mineralogy is the recent development of a high-quality research database containing 2279 samples. The database includes core samples from clean sands, shaly sands, shales, carbonates, and mixed lithologies from conventional and unconventional reservoirs from around the world. Thus, the mineral assemblages are representative of sedimentary formations commonly occurring in hydrocarbon-bearing environments. Major framework minerals, including quartz, calcite, dolomite, and anhydrite, are well distributed over the range of 0 to 100 wt%. Total clay content ranges from 0 to 92 wt%; illite is the dominant clay, averaging over half the total clay content. Total feldspar concentrations range from 0 to 53 wt%. Mica (predominantly muscovite), pyrite, and siderite have maximum values near 30 wt%, and ankerite has a maximum of 40 wt%.

Careful quality control was exercised in construction of the database, according to procedures described by Herron et al. (2014). Samples were crushed, homogenized, and split using a rotary splitter, and identical splits were used for chemical and mineral analysis. Chemical concentrations were measured by a combination of XRF, ICP/MS, and LECO by SGS. The accuracy of these measurements was vetted by blind analyses of Certified Reference Materials from NIST, USGS, and other similar organizations. The mineralogy was determined by transmission dual-range Fourier transform infrared spectroscopy (FTIR) (Herron et al. 1997). The quality of these measurements was monitored using the QCMin procedure in which the mineral concentrations determined for each sample are multiplied by the chemical compositions of those minerals and summed to produce a mineral-based estimate of the multiple-element sample’s chemical composition. This estimated chemical composition is then compared, element by element, with the measured chemical composition.

**Model-Independent Inversion Method**

The model-independent inversion method discussed in this paper was first proposed by Freedman (2006; 2007) as a new approach for solving complex reservoir characterization problems for which accurate forward models are not known. It has already been successfully used to solve a number of challenging problems (e.g., Anand et al. 2011; Gao et al. 2011; Freedman et al. 2012; Freedman et al. 2013). Our paper is the first application of the method to the mineralogy problem. The model-independent inversion method relies on having a comprehensive calibration database of measurements. The database is used to derive a model-independent mapping function that accurately represents the functional relationship between the dry-weight elemental concentrations and the dry-weight mineral concentrations for all of the samples in the database. The mapping function is expressed as a weighted sum of normalized Gaussian radial basis functions (RBFs). The weights and the widths of the Gaussian functions are determined from the database.

The model-independent inversion method has a strong and well-established mathematical foundation. Applied mathematicians have proven that RBF mapping functions provide more accurate interpolations for multi-dimensional functions of many variables (Powell 2001) than do other interpolation methods. One of the attractive features of RBF interpolation is that neither densely populated nor large databases are required. Mathematicians have established that accurate results can be obtained using sparsely populated databases with scattered data (i.e., non-evenly distributed). Moreover, mathematicians have shown that RBF mapping functions have superior generalization properties for samples not included in the
database. Once the mapping function has been derived from the database it is used to predict continuous logs of mineral concentrations and matrix densities from dry-weight elemental concentrations measured by the new gamma ray spectroscopy logging tool.

The model-independent approach has many attractive features, including the following:

- It is easy to implement.
- It is applicable to both linear and non-linear problems.
- It works well with sparsely populated databases.
- Mineral compositions and endpoints are not needed.
- No mineral model is assumed.
- Number of outputs can exceed number of inputs.
- There are no free parameters or subjective user inputs.

Log examples discussed later in this paper show predicted logs of 14 mineral concentrations (i.e., illite, smectite, kaolinite, chlorite, quartz, calcite, dolomite, ankerite, plagioclase, orthoclase, mica, pyrite, siderite, and anhydrite) plus matrix densities. The mineralogy logs are predicted from eight elemental concentrations (i.e., Si, Al, Ca, Mg, K, Fe, S, and Mn) measured by the new geochemical logging tool (Radtke et al. 2012). The log mineralogies are shown to agree very well with core mineralogies measured in our laboratory using FTIR.

**Derivation of the RBF mapping function** This section summarizes the derivation of the RBF mapping function. To save space some of the details are omitted, such as we do not display the normalized Gaussian functions that can be found in the Freedman (2006) paper. The latter paper also provides additional intuitive insight as to how the method works. Our objective here is to construct a mapping function from the database that can be used to predict mineralogy and matrix densities from an n-dimensional vector \( \tilde{x} \) of dry-weight elemental concentrations (e.g., \( n = 8 \)) measured by the logging tool. The mapping function, \( \tilde{f}(\tilde{x}) \), is an m-dimensional vector whose elements are the dry-weight mineral concentrations and the matrix densities (e.g., \( m = 15 \)) predicted from the log-derived elemental concentrations. The number of RBF outputs exceeds the number of measurements. This does not violate any rules of algebra because the mapping function outputs are not determined by solving a system of equations but by interpolation in the multi dimensional mineral output space. On the contrary, model-dependent inversions with constraints require about as many measurements as there are outputs because to determine the outputs, one must solve a set of algebraic equations that arise from fitting the measurements to the model equations. This is the reason that model-dependent methods require use of a mineral model because there are seldom enough measurements to solve for all the minerals that can be present.

Consider a database of \( N \) samples where the database measurements include dry-weight elemental concentrations, dry-weight mineral concentrations, and matrix densities. The mapping function is constructed from the database as follows. The mapping function can be expressed as a weighted sum of \( N \) normalized Gaussian functions \( (\phi) \) called the RBFs,

\[
\tilde{f}(\tilde{x}) = \sum_{i=1}^{N} \tilde{c}_i \phi(\|\tilde{x} - \tilde{x}_i\|).
\]  

The summation in Eq. 1 is over the \( N \) samples in the database where \( \tilde{x}_i \) is a vector with components that are the dry-weight elemental concentrations (or a subset of the measured elemental concentrations) for the \( i \)th database sample. The weighting coefficient \( \tilde{c}_i \) in Eq. 1 is an \( m \)-dimensional vector associated with the RBF centered at the \( i \)th database sample. The widths of the Gaussian RBFs can be determined from the Euclidean nearest neighbor distances of \( \tilde{x}_i \) in the input measurement space. The mapping function predictions are not overly sensitive to the Gaussian function widths as discussed by Freedman.
(2006), and good results have been achieved by choosing the widths to be equal to or of the order of the nearest neighbor distances.

The arguments of the Gaussian RBFs in Eq. 1 are proportional to the squared Euclidean norms in the $n$-dimensional input space, e.g.,

$$\| \mathbf{x}_i - \mathbf{x}_j \|^2 = \sum_{p=1}^{n} (x_{ip} - x_{jp})^2.$$  

If the database inputs and outputs are required to satisfy Eq. 1, then the weighting coefficients $\hat{c}_i$ in Eq. 1 for fixed Gaussian widths are determined from the matrix equation,

$$C = \Phi^{-1} \cdot Y,$$  

where the $N \times N$ matrix $\Phi$ is positive definite. Its matrix elements are normalized Gaussian RBFs for which the arguments are the Euclidean norms of all pair-wise differences of the database elemental concentrations:

$$\Phi_{ij} = \phi(\| \mathbf{x}_i - \mathbf{x}_j \|) \quad 1 \leq i,j \leq N.$$  

The $N \times m$ matrix $C$ contains the weighting coefficients and the $N \times m$ matrix $Y$ contains the measured database mineral concentrations and matrix densities for the database samples.

A few comments about the mapping function in Eq. 1 will be instructive. It follows from Eqs. 1 and 3, that if the mapping function in Eq. 1 is evaluated at any database $\mathbf{x}_i$ then the mapping function output exactly reproduces, to within computer precision, the database mineral concentrations and matrix densities for that sample. Another useful and perhaps non-obvious property of the mapping function in Eq. 1 is that the sum of the RBF-predicted mineral weight percentages automatically sums to 100 because the database mineral concentrations sum to 100%.

The RBF matrix ($\Phi$) in Eq. 4 is positive definite for Gaussian RBFs so that the inverse matrix in Eq. 3 always formally exists and is non-singular; however, if the condition number (ratio of the largest to the smallest eigenvalue) of the $\Phi$ matrix is large then the weighting coefficients can contain numerical noise arising from the matrix inverse in Eq. 3. This noise can propagate via Eq. 1 to the predicted mineralogy logs.

We observed this problem in a few of the many well logging datasets we processed. Some of the predicted mineralogy logs had a spurious high-frequency character that was not consistent with either the vertical resolution of the tool or thin laminations in the reservoir. Plots of the weighting coefficients for each of the minerals confirmed that some of the coefficients were noisy. The condition number of the $\Phi$ matrix for our database was approximately 36,000. The noise can be suppressed by using mathematical regularization to reduce the condition number and filter out the noise that arises from the small eigenvalues in the $\Phi$ matrix. This was discussed by Freedman (2006). It leads to the following modification of Eq. 3,

$$C = (\Phi + \alpha \cdot I)^{-1} \cdot Y,$$  

where $\alpha$ is a non-negative “regularization parameter.” The regularization parameter in Eq. 5 multiplies the $N \times N$ identity matrix ($I$). The default value is $\alpha = 0$; however, if the predicted mineralogy logs appear noisy then a small positive value of $\alpha$ can significantly reduce the condition number of the $\Phi$ matrix and filter out any numerical noise in the weighting coefficients. The regularization reduces noise and improves the repeatability of the logs. It can, however, introduce some bias and therefore the tradeoff between accuracy and precision determines how much regularization should be used. We have found that for our database a regularization parameter equal to 0.5 significantly reduces the noise on the mineralogy logs without adversely affecting the fidelity of the predicted mineral concentrations and matrix densities.
It should be noted that if a nonzero regularization parameter is used then the RBF-predicted mineral percent concentrations no longer sum to 100% because the regularization tends to slightly suppress them. The predicted concentrations can be renormalized so that they sum to 100%. Likewise, if Eq. 5 is used to compute the weighting coefficients in Eq. 1, then the mapping function evaluated at database inputs \( \tilde{\mathbf{x}} \) closely approximates but no longer exactly reproduces the measured database mineral concentrations and matrix densities.

**Accuracies of Mapping Function Predictions** The database itself can be used to assess the accuracies that can be expected from the mapping function predictions. This is done using the so-called “leave-one-out” (LOO) method. In the LOO method a sample is removed from the database and the mapping function for the reduced database with \((N - 1)\) samples is computed. The mapping function from the reduced database is used to predict the mineral concentrations and the matrix density for the sample that was removed using the database elemental concentrations for the removed sample. The sample that was removed is restored to the database, another sample is removed, and this process is repeated for all \(N\) database samples. The results of the LOO method are the RBF-predicted mineral concentrations and matrix densities for the \(N\) database samples. The deviations of each of the RBF predictions from the database values can then be computed. The LOO method is time consuming because it requires computing \(N\) mapping functions. The LOO method needs to be done only to assess the accuracy of the mapping function predictions.

The results of the LOO method applied to the worldwide database consisting of 2279 samples are summarized in Table 1 and shown in Fig. 1. The LOO results shown in Fig. 1 were computed using a regularization parameter equal to 0.5. The LOO was also applied without regularization and the results (not shown here) showed very slightly higher average absolute deviations and slightly lower average deviations, as expected. The plots in Fig. 1 show RBF-predicted mineral concentrations on the y-axis and the measured values on the x-axis. On each plot we show three statistical quantities characterizing the deviations (i.e., RBF predictions minus the measured values). These quantities are the average absolute deviation (aad), average deviation (ad), and correlation coefficient (cc).

The cross plots in Fig. 1 show comparisons of predicted and measured values for all 2279 database samples. For most of the 14 predicted mineral concentrations, the average absolute deviations, a measure of their scatter, is less than 1.0 wt%. Also, for the matrix density the average absolute deviation is less than 0.01 g/cm³. The average deviations, which are a measure of bias, are generally very small. Finally, the correlation coefficients are very good for most of the predicted mineral concentrations. The lower correlation coefficients for ankerite and mica attest to the difficulty in predicting these complex minerals.

### Table 1—SUMMARY OF DATABASE MINERAL CONCENTRATIONS AND AVERAGE ABSOLUTE DEVIATIONS FOR RBF PREDICTIONS

<table>
<thead>
<tr>
<th>Mineral</th>
<th>Database Range (wt %)</th>
<th>RBF Average Absolute Deviation (wt %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Illite</td>
<td>0 – 55</td>
<td>2.0</td>
</tr>
<tr>
<td>Smectite</td>
<td>0 – 34</td>
<td>0.9</td>
</tr>
<tr>
<td>Kaolinite</td>
<td>0 – 68</td>
<td>0.9</td>
</tr>
<tr>
<td>Chlorite</td>
<td>0 – 25</td>
<td>0.9</td>
</tr>
<tr>
<td>Quartz</td>
<td>0 – 99</td>
<td>1.3</td>
</tr>
<tr>
<td>Calcite</td>
<td>0 – 100</td>
<td>1.1</td>
</tr>
<tr>
<td>Dolomite</td>
<td>0 – 100</td>
<td>1.0</td>
</tr>
<tr>
<td>Ankerite</td>
<td>0 – 40</td>
<td>0.6</td>
</tr>
<tr>
<td>Orthoclase</td>
<td>0 – 44</td>
<td>0.8</td>
</tr>
<tr>
<td>Plagioclase</td>
<td>0 – 27</td>
<td>1.3</td>
</tr>
<tr>
<td>Mica</td>
<td>0 – 29</td>
<td>1.0</td>
</tr>
<tr>
<td>Pyrite</td>
<td>0 – 22</td>
<td>0.3</td>
</tr>
<tr>
<td>Siderite</td>
<td>0 – 32</td>
<td>0.3</td>
</tr>
<tr>
<td>Anhydrite</td>
<td>0 – 100</td>
<td>0.4</td>
</tr>
<tr>
<td>Matrix density (g/cm³)</td>
<td>2.63 – 3.09</td>
<td>0.008</td>
</tr>
</tbody>
</table>

**Log Examples in Cored Wells**

**Processing Summary**

This section discusses RBF processing of three unconventional wells for which core-measured mineralogy and matrix densities are available for comparison. The three wells are from very different areas and have very different lithologies and mineralogies. The RBF mapping function in Eq. 1 was used to process the
elemental chemistry logs from the new geochemical logging tool. No other data were used to obtain the results shown in this section. RBF processing is simple and very fast because only the summation in Eq. 1 is needed and hundreds of feet of elemental chemistry data from the new geochemical tool can be processed in a few seconds to produce continuous logs of mineral concentrations and matrix densities. It is worth repeating again that the RBF method does not require any user inputs or have any tunable parameters. All of the log data shown in this section were processed with the same RBF mapping function.

Fig. 2 shows the eight elemental weight fractions that are the logging tool inputs and the RBF mapping function outputs, which are continuous logs of 14 mineral weight fractions and the matrix density.

**South Texas Eagle Ford Shale Well** The first log example is from a well drilled in the Eagle Ford shale formation in South Texas, which is one of the most prolific oil-producing reservoirs currently in North America. To the right of the depth track in Fig. 3 are continuous logs of RBF-predicted mineral weight fractions and matrix densities over the cored interval shown as solid curves. The solid circles are the core measurements. A color-coded lithology column is shown to the left of the depth track. The sum of the 14 RBF predicted mineral fractions is equal to one.

This well has a complex mixed lithology composed of significant amounts of calcite, quartz, and clay minerals. While illite is by far the dominant clay mineral, there are measurable quantities of both kaolinite and smectite. Observe the excellent agreement between the RBF-predicted logs and the core measurements. In particular, note that the method correctly evaluates minerals that are quantitatively significant such as illite, kaolinite, smectite, quartz, and calcite. Equally important is the fact that the RBF does not
Figure 2—A schematic of the RBF mapping function processing of the geochemical logging tool data. The elemental concentrations from the geochemical logging tool are inputs to the RBF mapping function (Eq. 1), which is shown here. The outputs of the mapping function are the 14 minerals and the matrix density (MDens) shown.

\[
\tilde{F}(\tilde{x}) = \sum_{i=1}^{\bar{N}} c_i \cdot \phi(\| \tilde{x} - \tilde{x}_i \|)
\]

Figure 3—RBF processing of the geochemical tool data from an unconventional well in the Eagle Ford shale formation in South Texas. The curves in the tracks to the right of the depth track show continuous logs of RBF-predicted dry-weight matrix densities and mineral concentrations. The circles are the dry-weight matrix densities and mineral concentrations derived from core measurements. The core mineral concentrations were measured by FTIR. A color-coded lithology track is shown to the left of the depth track. Observe the good quantitative agreement between the log-predicted mineral concentrations and the core-measured concentrations in this complex formation.
predict or overestimate mineral fractions that are absent or present in only trace amounts, such as chlorite, muscovite, feldspars, and other carbonate minerals. It has not previously been possible to quantitatively determine these complex minerals from well logging data. The smectite concentration is particularly important due to its adverse impact on both conventional and unconventional resources, and here the predicted smectite concentrations are in good agreement with the measured values. The RBF-predicted dry-weight matrix densities shown in Fig. 3 are also in excellent agreement with dry-weight densities measured on the cores.

As a check on the integrity of the RBF-predicted logs in Fig. 3, we removed the 20 core samples in our database from this well. A reduced mapping function was computed from the remaining 2259 database samples. The reduced mapping function was used to reprocess the elemental chemistry. The mineralogy and matrix density results from the reduced database processing are essentially identical to the results in Fig. 3, which demonstrates the robustness of our worldwide database and the excellent generalization properties of RBF mapping functions. Similar tests done with other wells also confirm this conclusion.

**West Texas Wolfcamp Shale Well**  The second log example is from a well drilled in the Wolfcamp shale in west Texas. To the right of the depth track in Fig. 4 are continuous logs of RBF-predicted mineral weight fractions and matrix densities over the cored interval shown as solid curves. The solid circles are the core measurements. A color-coded lithology column is shown to the left of the depth track. The sum of the 14 RBF-predicted mineral fractions is equal to one.

The Wolfcamp shale is currently one of the most prolific unconventional oil reservoirs in the Permian Basin. There are three distinct cored intervals in this well with non-cored intervals separating them, as can be seen from the log in Fig. 4. The dominant minerals in this complex lithology are siliciclastics with only very small amounts of carbonate present. There are significant amounts of quartz and feldspars, for which the RBF-predicted mineral weight fractions are in quantitative agreement with the core weight fractions. There are also present in this well all four clay types, for which the RBF-predicted weight fractions are in good overall agreement with the core-measured values. The small amounts of ankerite predicted by the RBF processing agree very well with the small weight fraction measured on the cores. It is worth noting
again that the RBF-predicted weight fractions do not overestimate the weight fractions for minerals such as pyrite, dolomite, and anhydrite, which are present only in trace amounts. The RBF dry-weight matrix densities are also seen to be in very good quantitative agreement with the dry-weight matrix densities measured on cores.

Canadian Montney Shale Well  The third log example is from a well drilled in the Montney shale in Canada. To the right of the depth track in Fig. 5 are continuous logs of RBF-predicted mineral weight fractions and matrix densities over the cored interval shown as solid curves. The solid circles are the core measurements. A color-coded lithology column is shown to the left of the depth track. The sum of the 14 RBF-predicted mineral fractions is equal to one.

This well has a complex mixed lithology with significant amounts of quartz, calcite, dolomite, plagioclase (Na feldspar), orthoclase (K feldspar), illite, and smectite, as shown in Fig. 5. There is good quantitative overall agreement between the RBF and core mineralogy. Particularly noteworthy is the quantitative prediction of quartz, clay minerals, and mica, which is very challenging in the presence of so much feldspar, although the RBF-predicted feldspar is underestimated relative to that from core. Particularly noteworthy once again is the very good prediction of smectite concentrations. The calcite and dolomite predictions are also reasonably good. The RBF-predicted matrix densities are observed to be in very good agreement with the measured values.

Repeatability of Log-Derived Mineralogy
As discussed previously, the precision of the RBF-predicted mineral concentrations and matrix density logs is improved by using a nonzero regularization parameter (\( \alpha \)) to compute the weighting coefficients in Eq. 1. Applying nonzero regularization reduces the condition number of the matrix inversion in Eq. 3 and serves to filter out numerical noise that arises from inverting the matrix. For all of the results shown in this paper we have used a mapping function computed using a regularization parameter of 0.5 in Eq. 3. This reduces the condition number of the matrix inversion in Eq. 3 from 36,000 for \( \alpha = 0 \) to 4 for \( \alpha = 0.5 \). The regularization filters out the effects of the small eigenvalues in the \( \Phi \) matrix. This causes no
loss of information because the small eigenvalues contribute negligible information to the weighting coefficients.

**Fig. 6** shows to the left of the depth track the RBF-predicted mineral concentrations for a main and a repeat pass over an interval in a heavy oil well in the San Joaquin Valley, California. To the right of the depth track are logs of elemental concentrations. The mineral concentrations and matrix densities predicted for the two passes are in good overall agreement. The larger differences between the main and repeat mineralogy logs at some depths are caused by differences in the main and repeat log elemental concentrations, especially in the Al and K weight fractions. It should be clear from this example that quantitative log predictions of mineral fractions and matrix densities are possible only when the elemental concentrations can be accurately measured with high precision. This requires a high-performance geochemical tool with accurate data processing algorithms, including borehole corrections that can output accurate and precise elemental concentrations.

**Quality Checks on RBF Mineralogy Predictions** This section addresses the important question of how one detects when the input elemental chemistry data are not well represented by the calibration database samples and therefore when the RBF predictions should be flagged as not trustworthy. The question of what to do when we are “outside of the database” is not specific to the RBF method but is relevant to all empirically derived methods that are based on databases. For example, Archie used a core database of clean Gulf Coast sandstone formations to derive his famous saturation equation. It is well-known that the equation does not account for clay conductance effects in freshwater shaly sand formations, where it underestimates oil saturations.

Two of the RBF quality control checks are based on a level-by-level analysis of the elemental concentrations vector that is input from the logging tool. The first quality check is to test whether or not each of the eight input elemental concentrations (e.g., Si, Al, Ca, Mg, K, Fe, S, and Mn) are within the range of the database chemistry. If a single input elemental concentration is outside of the database range for that element then a warning flag is turned on at the depths for which this occurs.
The second quality check can detect a situation where the input elemental chemistry falls within the database range but might be located near the edge or boundary of the database. In this case there might be too few nearby database samples for accurate RBF interpolation. To address this case we compute the Euclidean distances of the input elemental concentrations vector from the database elemental concentrations. A “proximity flag” is turned on if there are less than a specified number of database samples (e.g., four) within a specified Euclidean radius in the elemental chemistry space. This radius is of the order of the average of the database nearest neighbor distances used to determine the widths of the Gaussian RBFs. The optimal number of nearby database neighbors and the optimal radius can be determined by simulation and confirmed using field data.

The third quality check uses the RBF-predicted dry-weight mineral concentrations to compute reconstructed or theoretical dry-weight elemental concentrations. The elemental reconstruction computation assigns fixed elemental compositions to each mineral. The theoretical elemental concentrations computation is conducted for the major elements present in sedimentary rocks (i.e., Si, Al, K, Fe, S, Ca, Mg, and Na). Incorrectly predicted mineral concentrations can be detected in intervals where there are significant deviations between the reconstructed elemental concentrations and the log-derived elemental concentrations. The usage and interpretation of reconstructed elemental concentrations to quality check mineralogy is discussed in much more detail by Herron et al. (2014). Deviations between the reconstructed and tool-derived elemental concentrations can be used to flag intervals where the predicted mineral concentrations are likely to be incorrect.

Although our existing database contains 2279 samples including cores from around the world of clean sands, shaly sands, shales, carbonates and mixed lithologies, it could be expanded in future developments and made more comprehensive by adding additional core data for which the mineralogy is absent from or not well represented by the existing database. Of course, any added core chemistry and mineralogy data has to be of the same high quality as the data in the existing database to enable quantitative mineralogy predictions. After the database is augmented with the new data, it takes only a few minutes of computer time to compute a new and more universal RBF mapping function. Continuing investment in the database expansion might ultimately achieve a more universal database that can predict accurate mineralogy anywhere in the world.

Conclusions and Summary

This paper presents a robust, general interpretation procedure to quantitatively derive accurate, detailed mineral concentrations from geochemical tool logging data. This major breakthrough solves a long-standing important problem in formation evaluation. It was made possible by the confluence of the three recent developments of a new high-performance geochemical logging tool, a new high-quality worldwide core database of measured dry-weight elemental concentrations and dry-weight mineral concentrations, and a new model-independent inversion method. All three of these recent developments are equally important and are essential enablers for quantitative mineral predictions.

We discussed the high-output neutron generator, high-resolution scintillation detector, and advanced electronics in the new geochemical logging tool. These advances in technology together with improvements in data processing make it possible to derive, from the measured gamma ray spectra, robust and accurate elemental concentrations including Si, Al, Ca, Mg, K, Fe, S, and Mn. These elements are absolutely essential for quantitative and detailed mineralogy prediction in sedimentary rocks.

The new high-quality research database with 2279 samples includes cores from conventional and unconventional reservoirs worldwide. The measurements on the database cores include dry-weight elemental concentrations, dry-weight mineral concentrations measured by Fourier transform infrared spectroscopy, and dry-weight matrix densities. We also discussed database quality checks that were used to ensure the consistency of the measured elemental concentrations and the measured mineral concen-
trations. The quality checks are essential to ensure the construction of an accurate calibration database, which is essential for developing and testing methods for quantitative and detailed mineralogy predictions.

The new model-independent inversion method uses an RBF mapping function derived from the calibration database to predict quantitative and detailed mineral concentrations and matrix densities. The shortcomings of traditional model-dependent inversion methods are discussed. The features of the new model-independent method that overcome the limitations of the model-dependent methods are discussed. The derivation of the RBF mapping function from the database is explained. The LOO method by which the expected accuracies of the RBF mapping function can be determined from the database is explained and results of the LOO method are shown and discussed.

We processed a subset of the elemental concentrations data acquired by the new geochemical logging tool in three unconventional shale wells. The same RBF mapping function was used to process the data from all three wells. The inputs to the RBF mapping function were continuous logs of the eight dry-weight elemental concentrations (i.e., Si, Al, Ca, Mg, K, Fe, S, and Mn) derived from the logging tool measurements. The outputs from the RBF mapping function were continuous logs of 14 dry-weight mineral concentrations (i.e., illite, smectite, kaolinite, chlorite, quartz, calcite, dolomite, ankerite, orthoclase, plagioclase, mica, pyrite, siderite, and anhydrite) plus the matrix density. The RBF-predicted minerals and matrix densities were compared with core results. The core mineralogy was measured using Fourier transform infrared spectroscopy. The comparisons of the RBF-predicted and core-measured dry-weight mineral concentrations and matrix densities are shown to be in very good quantitative agreement. These predictions were obtained without using any adjustable parameters or user inputs.

We discussed three quality checks on the RBF-predicted mineralogy. The input elemental concentrations are tested at each depth to determine if they are within the range of the database concentrations. Another test determines the proximity of the input elemental concentrations at each depth to the database samples to determine if there are enough nearby samples to provide for accurate RBF interpolation. If these tests show that the input data are “outside of the database” at some depth, then a warning flag is used to indicate that the predicted mineralogy and matrix density are questionable at that depth. The third quality check compares elemental concentrations from the logging tool at each depth with those reconstructed from the predicted mineralogy. Significant deviations between the elemental concentrations from the logging tool and the reconstructed concentrations are indicative of errors in the predicted mineralogy.

**Nomenclature**

\[ \hat{c}_i = m \text{-dimensional vector of weighting coefficients for the } i \text{th database sample for } i = 1, \ldots, N \]

\[ m = \text{dimensionality of the RBF mapping function (i.e., equal to the number of mineral concentration outputs plus one)} \]

\[ n = \text{number of independent variables in the mapping function (i.e., equal to the number of elemental concentration inputs)} \]

\[ \tilde{x} = n \text{-dimensional vector with elements that are the elemental concentrations input from the logging tool} \]

\[ \tilde{x}_i = n \text{-dimensional vector with elements that are the elemental concentrations for the } i \text{th database sample} \]

\[ C = N \times m \text{ weighting matrix with rows that are the vectors, } \hat{c}_i \]

\[ N = \text{number of samples in the database} \]

\[ \mathbf{F}(\tilde{x}) = m \text{-dimensional RBF mapping function defined in Eq. 1} \]

\[ Y = N \times m \text{ matrix whose elements are the mineral concentrations and matrix densities of the database samples} \]
\[ \alpha = \text{non-negative regularization parameter in Eq. 3} \]

\[ \phi(\|\bullet\|) = \text{normalized Gaussian RBF} \]

\[ \Phi = N \times N \text{ RBF matrix with elements, } \Phi_{i,j} = \phi(\|\mathbf{x}_i - \mathbf{x}_j\|) \]

\[ \|\bullet\| = \text{double bars denoting the Euclidean norm of the vector argument} \]

**Acknowledgments**

It is a pleasure to thank Jim Grau and Jack Horkowitz for developing the algorithms used to derive the elemental concentrations from the measured gamma ray spectra. Also, thanks to the Litho Scanner project team for their support and to the SDR Petrolabs staff for core analyses. We also owe thanks to Aria Abubakar, Farid Hamichi, Ridvan Akkurt, Daniel Codazzi, and Vlad Tertychnyi for managerial support and to Schlumberger for permission to publish this paper.

**References**


